NOTES ON LECTURES ON THE SPECTRAL THEORY OF EISENSTEIN SERIES

PAUL D. NELSON

Contents

1. Introduction and motivation

1.1. Overview. We will review, through several examples, the general problem of *spectral decomposition*, which may be summarized as follows: given a "natural" Hilbert space H, find a "nice basis" $(e_j)_j$ of H. A typical example of such a Hilbert space is $H = L^2(X, \mu)$ for some nice topological space X and some nice reference measure μ. "Nice" typically refers to being an eigenfunction of some operator of interest.

1.2. Fourier series. Consider the circle group

$$X = \mathbb{R}/\mathbb{Z},$$

equipped with the Lebesgue measure μ. A "nice basis" $(e_n)_{n\in\mathbb{Z}}$ for this space is given by the complex exponentials

$$e_n(x) := e(nx), \qquad \text{where } e(x) := e^{2\pi i x}.$$

What do we mean by "basis"? The theory of Fourier series implies that the e_n furnish a Hilbert space basis of

$$H = L^2(X).$$

In other words, we have the direct sum decomposition

$$L^2(X) = \oplus_{n\in\mathbb{Z}}\mathbb{C}e_n$$

in the Hilbert sense, meaning that the summands are orthogonal and their sum is dense.

What do we mean by "nice"? There are a couple natural meanings:

Date: May 4, 2021.

- each e_n is an eigenfunction for the Laplacian $\Delta = \partial_{xx}$:
$$\Delta e_n = -(2\pi n)^2 e_n.$$
- For $x \in \mathbb{R}/\mathbb{Z}$, let $\rho(x) : H \to H$ denote the "translation by x" map: for $v \in H$,
$$\rho(x)v(y) := v(x+y).$$
Then each e_n is an eigenfunction for each $\rho(x)$:
$$\rho(x)e_n = e(nx)e_n,$$
as follows readily from the fact that $e_n : \mathbb{R}/\mathbb{Z} \to \mathbb{C}^\times$ defines a homomorphism. Said another way, each e_n spans a one-dimensional, hence irreducible, subrepresentation $\mathbb{C}e_n \subset H$ for the regular representation $\rho : \mathbb{R}/\mathbb{Z} \to \mathrm{GL}(L^2(\mathbb{R}/\mathbb{Z}))$. Informally, each e_n transforms "as simply as possible" under translation.

1.3. Equidistribution and Weyl's criterion.

Why care about "nice bases" as in §1.2? One reason is that they provide a natural tool for equidistribution problems.

Definition 1.1. Let (X, μ) be a Borel probability space. We say that a sequence $(x_j)_{j \geqslant 1}$ in X *equidistributes* (with respect to μ) if for each $\psi \in C_c(X)$, we have as $J \to \infty$
$$\frac{1}{J}\sum_{1}^{J} \psi(x_j) \to \int_X \psi \, d\mu. \tag{1.1}$$

If a sequence equidistributes, then it is dense. This implication gives a powerful tool for establishing the density of certain sequences.

Theorem 1.2 (Weyl's criterion). *Suppose that $X = \mathbb{R}/\mathbb{Z}$, equipped with Lebesgue measure. The following are equivalent for all sequences (x_j) in X:*

(i) (x_j) equidistributes

(ii) for each $n \in \mathbb{Z}$, the convergence (1.1) holds for $\psi = e_n$, i.e.,
$$\frac{1}{J}\sum_{1}^{J} e(nx_j) \to \int_{\mathbb{R}/\mathbb{Z}} e(nx)\, dx = \begin{cases} 1 & \textit{if } n = 0, \\ 0 & \textit{otherwise.} \end{cases} \tag{1.2}$$

We note that the condition (1.2) is automatic when $n = 0$, so the main point in applying this criterion is to estimate the sums $\sum_1^J e(nx_j)$ for $n \neq 0$.

Corollary 1.3. *For $\alpha \in \mathbb{R} - \mathbb{Q}$, the fractional parts $[j\alpha] \in \mathbb{R}/\mathbb{Z}$ ($j = 1, 2, 3, \dots$) equidistribute: for each $\psi \in C_c(\mathbb{R}/\mathbb{Z})$,*
$$\frac{1}{J}\sum_{1}^{J} \psi([j\alpha]) \to \int_{\mathbb{R}/\mathbb{Z}} \psi(x)\, dx.$$
In particular, $\{[j\alpha] : j = 1, 2, 3, \dots\}$ is dense in $\mathbb{R}/\mathbb{Z}$.

Exercise 1.1. Let N be a large positive integer. Define $f_0 : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ by
$$f_0(0) := N, \quad f_0(x) := 0 \text{ for } x \neq 0.$$
Define f_n for $n \geqslant 1$ inductively by
$$f_{n+1}(x) := \frac{f_n(x-1) + f_n(x+1)}{2}.$$

Suppose that N is odd. Show that there exist positive constants c_0, c_1 (not depending upon N or n) so that

$$\sum_{x \in \mathbb{Z}/N\mathbb{Z}} |f_n(x) - 1|^2 \leqslant c_0 N^4 \exp(-c_1 n/N^2).$$

Thus if n is a bit larger than N^2, then f_n is quite uniformly distributed. What happens if n is even?

[Hints: Define $\psi : \mathbb{Z}/N\mathbb{Z} \to \mathrm{U}(1)$ by $\psi(x) := e(x/N) = e^{2\pi i x/N}$. For $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, define the Fourier transform $\hat{f} : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ by

$$\hat{f}(\xi) = \frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} f(x)\psi(-x\xi),$$

so that the inversion formula

$$f(x) = \sum_{\xi \in \mathbb{Z}/N\mathbb{Z}} \hat{f}(\xi)\psi(x\xi)$$

holds. Compute that $\hat{f}_n(\xi) = \cos(2\pi\xi/N)^n$. For the estimation, it may be useful to note that $|\cos(2\pi x/N)| \leqslant 1 - c_1/N^2 \leqslant \exp(-c_1/N^2)$ for some constant $c_1 > 0$ whenever x is an integer with $2x \neq N$.]

1.4. Fourier transform. We briefly describe the main examples of Fourier analysis on *non-compact abelian* groups. (The general picture is described by *Pontryagin duality.*)

Example 1.4. Let $G = \mathbb{R}$, regarded as the additive group. A Schwartz function v on $\mathbb{R}$ has a Fourier transform $v^\wedge(\xi) := \int_{x \in \mathbb{R}} v(x) e(-\xi x)\, dx$, which satisfies the Fourier inversion formula $v = \int_{\xi \in \mathbb{R}} v^\wedge(\xi) e_\xi \, d\xi$, where $e_\xi(x) := e(\xi x)$. Remarks:

(1) The e_ξ are eigenfunctions of $\Delta = \partial_{xx}$, with eigenvalues $-(2\pi\xi)^2$.
(2) The e_ξ are eigenfunctions under the translation operators $\rho(x)$, defined as above, with eigenvalues given by $\rho(x)e_\xi = e(\xi x)e_\xi$.
(3) We have $e_\xi \notin L^2(\mathbb{R})$.
(4) The eigenfunctions e_ξ come in *continuous* families indexed by $\xi \in \mathbb{R}$ (unlike the discrete families encountered in the $\mathbb{R}/\mathbb{Z}$ case).
(5) We have the Parseval identity for $L^2(\mathbb{R})$:

$$\langle v_1, v_2 \rangle = \int_{\xi \in \mathbb{R}} \langle v_1, e_\xi \rangle \langle e_\xi, v_2 \rangle \, d\xi.$$

A convenient way to summarize some of the properties noted above is through the language of direct integrals of Hilbert spaces: we have $L^2(\mathbb{R}/\mathbb{Z}) = \oplus_{n \in \mathbb{Z}} \mathbb{C} e_n$ (Hilbert direct sum) but $L^2(\mathbb{R}) = \oint_{\xi \in \mathbb{R}} \mathbb{C} e_\xi \, d\xi$ (Hilbert direct integral).

Example 1.5. The multiplicative group $G = \mathbb{R}^\times$ is isomorphic to $\mathbb{R}^\times_+ \times \{\pm 1\}$, and $\mathbb{R}^\times_+ \cong \mathbb{R}$ via the logarithm map. Thus the discussion for $\mathbb{R}$ applies also to $\mathbb{R}^\times$, where it is usually formulated in the language of Mellin transforms.

1.5. The space of lattices.

Definition 1.6. A *lattice* $L \subseteq \mathbb{R}^n$ is a subgroup of the form $L = \oplus_{i=1}^n \mathbb{Z} v_i$ for some basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$.

For example, there is the standard lattice $\mathbb{Z}^n \subseteq \mathbb{R}^n$ generated by the standard basis elements $e_1, \dots, e_n$.

We denote by X_n the space of lattices in $\mathbb{R}^n$.

We regard $\mathbb{R}^n$ as a space of row vectors. The group $\mathrm{GL}_n(\mathbb{R})$ acts on $\mathbb{R}^n$ by right matrix multiplication: $(v, g) \mapsto vg$. This induces an action on the space X_n of lattices. The stabilizer of the standard lattice $\mathbb{Z}^n$ is the group $\mathrm{GL}_n(\mathbb{Z})$, so we may identify

$$X_n = \mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$$

$$\mathbb{Z}^n g \leftrightarrow \mathrm{GL}_n(\mathbb{Z}) g.$$

We equip X_n with the quotient topology coming from the surjective map $(v_1, \dots, v_n) \mapsto X_n$ sending an ordered basis for $\mathbb{R}^n$ to the lattice it generates, or equivalently, with the quotient topology coming from the surjective map $\mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$. Thus a sequence $\mathcal{L}^{(j)}$ of lattices converge to $\mathcal{L}$ if there are bases $v_1^{(j)}, \dots, v_n^{(j)}$ of $\mathcal{L}^{(j)}$ and $v_1, \dots, v_n$ of $\mathcal{L}$ so that $v_i^{(j)} \to v_i$.

Definition 1.7. We say that a lattice $L \in X_n$ is *unimodular* if $\mathrm{vol}(\mathbb{R}^n / L) = 1$.

We denote by $X_n^{(1)} \subseteq X_n$ the space of unimodular lattices. A lattice L is unimodular if and only if it may be written $L = \mathbb{Z}^n g$ for some $g \in \mathrm{SL}_n(\mathbb{R})$. Thus we may identify

$$X_n^{(1)} = \mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R}).$$

The group $\mathbb{R}^\times$ acts on the space of lattices by scaling. Scaling a lattice by t multiplies the volume by $|t|^n$. We may thus identify

$$X_n / \mathbb{R}^\times \cong X_n^{(1)}.$$

Here the LHS consists of scaling classes of lattices, while the RHS consists of unimodular lattices; the bijection assigns to each scaling class the unique unimodular representative.

One advantage of working with the space $X_n^{(1)}$ of unimodular lattices rather than the larger space X_n of all lattices is that $X_n^{(1)}$ admits a (unique) $\mathrm{SL}_n(\mathbb{R})$-invariant probability measure. (We will verify this fact soon.) We may thus meaningfully speak of a "random unimodular lattice."

Many problems in number theory, Diophantine analysis, ... may be recast in terms of the space of lattices. We record a typical example:

Example 1.8. The *Littlewood conjecture* asserts that for all real numbers α and β,

$$\liminf_{n \to \infty} n \|n\alpha\| \|n\beta\| = 0,$$

where $\|x\|$ denotes the distance from x to the closest integer.

Let $A \leqslant \mathrm{SL}_3(\mathbb{R})$ denote the diagonal subgroup. It is known that the Littlewood conjecture would follow from the following conjectural implication concerning lattices $L \in X_3^{(1)}$: if the orbit $LA \subseteq X_3^{(1)}$ is precompact, then that orbit is closed.

The basic idea underlying the the relationship between these two statements is that $n\|n\alpha\|\|n\beta\|$ is the smallest number of the form $n(n\alpha + m)(n\beta + \ell)$. Such numbers are values taken by the trilinear form $(x, y, z) \mapsto xyz$ on the lattice $L_{\alpha\beta}$ generated by $(1, 0, 0), (\alpha, 1, 0), (\beta, 0, 1)$. The group A stabilizes that trilinear form.

1.6. Goals for the course. The main goal is to describe "nice bases" for the spaces of functions on the spaces of X_n and $X_n^{(1)}$. Here

- "nice" may be characterized by analogy to what happened for $L^2(\mathbb{R}/\mathbb{Z})$, either in terms of eigenfunctions of certain differential operators or irreducibility under the action of the groups $\mathrm{GL}_n(\mathbb{R})$ or $\mathrm{SL}_n(\mathbb{R})$, and
- "basis" refers to some mixture of what happened for $L^2(\mathbb{R}/\mathbb{Z})$ (discrete decomposition) and $L^2(\mathbb{R})$ (continuous decomposition).

One of the main tools for doing so is the *Eisenstein series* construction, which consists of a way to produce functions on X_n from functions on simpler spaces (e.g., X_m for smaller m).

Example 1.9. Let $f \in C_c(\mathbb{R}^n)$. By summing f over vectors in a given lattice, we obtain a function $\mathrm{Eis}[f] : X_n \to \mathbb{C}$:

$$\mathrm{Eis}[f](L) := \sum_{v \in L} f(v).$$

(Many variants of this definition are possible: summing only over nonzero vectors, summing only over primitive vectors, etc.) In this way functions on the simpler space $\mathbb{R}^n$ may be used to construct functions on the space X_n.

Example 1.10. Write $n = n_1 + n_2$. Let X_{n_1,n_2} denote the space of triple (V_1, L_1, L_2), where

- V_1 is an n_1-dimensional subspace of $\mathbb{R}^n$,
- L_1 is a lattice in V_1, and
- L_2 is a lattice in V/V_1.

We may regard X_{n_1,n_2} as a twisted product of the Grassmannian of n_1-dimensional subspaces of $\mathbb{R}^n$ and the space $X_{n_1} \times X_{n_2}$ of pairs of lattices in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$.

Let $L \in X_n$ be a lattice. Let $L_1 \subseteq L$ be a submodule of rank n_1, thus $L_1 \cong \mathbb{Z}^{n_1}$. We say that L_1 is *primitive* if it admits a complement, i.e., if there is a submodule $L_2 \subseteq L$ (of rank n_2) so that $L = L_1 \oplus L_2$.

Given $f \in C_c(X_{n_1,n_2})$, we may define $\mathrm{Eis}[f] : X_n \to \mathbb{C}$ by

$$\mathrm{Eis}[f](L) := \sum_{L_1} f(\mathrm{span}_{\mathbb{R}}(L_1), L_1, L_2),$$

where the sum is taken over primitive rank n_1 submodules L_1 of L, with L_2 as above.

In this special case $n_1 = 1$, this construction generalizes one of the variants mentioned in Example 1.9.

We will explain how, e.g., $L^2(X_n^{(1)})$ decomposes into

- a space spanned by Eisenstein series as above attached to functions on spaces like X_{n_1,n_2}, and
- its orthogonal complement, called the *cuspidal subspace.*

The cuspidal subspace decomposes like $L^2(\mathbb{R}/\mathbb{Z})$ does, while the Eisenstein space may be described in terms of the cuspidal subspaces for smaller X_m's.

These notes will likely be incomplete, and in particular short on the history; we refer to any of the course references for details.

2. Basics on the space of lattices

We continue to regard $\mathbb{R}^n$ as a space of row vectors.

References. Some expositions are Borel [3], Morris [10], and Platonov–Rapinchuk [11], while historical references include Siegel [12] and Borel–Harish-Chandra [4, 5].

2.1. Module-theoretic preliminaries. (Not presented in lecture.) Here we summarize some standard facts from abstract algebra and their consequences, to be used implicitly in what follows. (The proofs, omitted here, boil down to the fact that any n-tuple $(a_1, \dots, a_n) \in \mathbb{Z}^n$ of integers with relatively prime entries extends to a basis of $\mathbb{Z}^n$, or equivalently, arises as the top row of some element of $\mathrm{GL}_n(\mathbb{Z})$. The proof of the latter fact is a good exercise.)

First, the standard facts, which hold more generally with $\mathbb{Z}$ replaced by any principal ideal domain:

Theorem 2.1. *A submodule M of a free $\mathbb{Z}$-module L is free, with the rank of M at most that of L.*

Theorem 2.2. *A finitely-generated torsion-free $\mathbb{Z}$-module is free.*

Next, the consequences. Let L be a finitely-generated free $\mathbb{Z}$-module.

Definition 2.3. A submodule M of L is called *primitive* if $\mathbb{Q}M \cap L = \mathbb{Z}M$; here $\mathbb{Q}M$ and $\mathbb{Z}M$ denote the respective spans under $\mathbb{Q}$ and $\mathbb{Z}$ of M inside $L \otimes_{\mathbb{Z}} \mathbb{Q}$. A vector $v \in L$ is called *primitive* if it generates a primitive submodule, or equivalently, if $\mathbb{Q}v \cap L = \mathbb{Z}v$.

Lemma 2.4. *The following are equivalent for a submodule $M \subseteq L$:*

(i) M is primitive.
(ii) L/M is torsion-free.
(iii) L/M is free.
(iv) M admits a complement, i.e., a submodule M' of L for which $L = M \oplus M'$.

Proof. The equivalence of (i) and (ii) follows readily from the definition, while that of (ii) and (iii) follows from Theorem 2.2. (iv) implies (iii) because M' is free, while (iii) implies (iv) because the lift to L of any basis for L/M generates the required complement. □

In particular, we may construct a basis for L by iterated application of the lemma, as follows. We inductively construct $v_1, \dots, v_n$ so that for each $k \in \{1, \dots, n\}$, the module $\langle v_1, \dots, v_k \rangle$ is primitive. First, we choose any primitive vector $v_1 \in L$. Clearly $\langle v_1 \rangle$ is primitive. For $k \geqslant 2$, we choose any $v_k \in L$ whose image in the free module $L/\langle v_1, \dots, v_{k-1} \rangle$ is primitive. By induction on k, we see that $\langle v_1, \dots, v_k \rangle$ is primitive. The construction stops at $k = n$ with $\langle v_1, \dots, v_n \rangle = L$.

Exercise 2.1. Let M be a submodule of L. Write m and n for the respective ranks of M and L. Choose a basis of L, so that we may identify L with $\mathbb{Z}^n$ and speak of coordinates. Let $v_1, \dots, v_m \in \mathbb{Z}^n$ be a basis of M. Let A denote the $m \times n$ matrix with rows $v_1, \dots, v_m$. Show that M is primitive if and only if the greatest common divisors of the maximal minors of A is one.

2.2. The two-dimensional case. Let $\mathcal{L} \subseteq \mathbb{R}^2$ be a lattice. Choose a nonzero lattice vector v_2 of minimal length. By rotating and scaling the lattice, we may assume that $v_2 = (1, 0)$ and that all other vectors have length $\geqslant 1$. Note that v_2 is primitive, i.e., $\mathcal{L} \cap \mathbb{R}v_2 = \mathbb{Z}v_2$, as we would otherwise have $v_2/N \in \mathcal{L}$, contrary to the minimality of the length of v_2. Choose a vector $v_1 \in \mathcal{L} - \mathbb{Z}v_1$ of shortest length. Write $v_1 = (x, y)$. By replacing v_1 with its negative, we may assume that $y > 0$. Then $\|v_1\| \geqslant 1$. Also, $-1/2 \leqslant x \leqslant 1/2$, as otherwise we could translate v_1 by a multiple of v_2 to produce a shorter vector.

Thus we may classify lattices in $\mathbb{R}^2$ up to scaling and rotation by points (x, y) with $y > 0$, $-1/2 \leqslant x \leqslant 1/2$ and $x^2 + y^2 \geqslant 1$. (There is some mild ambiguity in this description when (x, y) lies on the "boundary.") We obtain the picture of the standard fundamental domain of the modular group $\mathrm{SL}_2(\mathbb{Z})$.

2.3. Iwasawa decomposition. Let $v_1, \dots, v_n$ be a basis for $\mathbb{R}^n$. There is a unique *orthogonal* basis $v_1', \dots, v_n'$ so that

$$v_i = v_i' + \sum_{j<i} x_{ij} v_j'$$

for some scalars x_{ij}. Indeed, we are forced to take for v_m' the orthogonal projection of v_m onto $\langle v_{m+1}, \dots, v_n \rangle^{\perp}$, which may be constructed explicitly using the Gram–Schmidt procedure. Said another way, one has the matrix identity

$$\begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_{n-1} \\ v_n \end{pmatrix} = \begin{pmatrix} 1 & x_{12} & x_{13} & \cdots & x_{1,n} \\ & 1 & x_{23} & \ddots & \vdots \\ & & \ddots & x_{n-2,n-1} & x_{n-2,n} \\ & & & 1 & x_{n-1,n} \\ & & & & 1 \end{pmatrix} \begin{pmatrix} v_1' \\ v_2' \\ \cdots \\ v_{n-1}' \\ v_n' \end{pmatrix}$$

where each entry x_{ij} is uniquely determined.

Set $a_j := |v_j'| > 0$, so that $v_j' = a_j v_j''$ where v_j'' is a unit vector. Then we have the matrix identity

$$\begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_{n-1} \\ v_n \end{pmatrix} = \begin{pmatrix} 1 & x_{12} & x_{13} & \cdots & x_{1,n} \\ & 1 & x_{23} & \ddots & \vdots \\ & & \ddots & x_{n-2,n-1} & x_{n-2,n} \\ & & & 1 & x_{n-1,n} \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_1 & & & & \\ & a_2 & & & \\ & & \ddots & & \\ & & & a_{n-1} & \\ & & & & a_n \end{pmatrix} \begin{pmatrix} v_1'' \\ v_2'' \\ \cdots \\ v_{n-1}'' \\ v_n'' \end{pmatrix}.$$

The rows of the rightmost matrix are orthogonal, and so that matrix lies in the orthogonal group $\mathrm{O}(n)$.

This way of describing a basis may be interpreted group-theoretically. Ordered bases as above correspond to elements of $G := \mathrm{GL}_n(\mathbb{R})$, with $(v_1, \dots, v_n)$ corresponding to the element g having jth row v_j, or equivalently, for which $e_j g = v_j$. Denote by N the group of strictly upper-triangular matrices, A^+ the group of diagonal matrices with positive entries, and $K = \mathrm{O}(n)$. Then every $g \in G$ may be written uniquely as $g = xak$ with $x \in N, a \in A^+, k \in K$. Symbolically, $G = NA^+K$, and the map $N \times A^+ \times K \to G$ is a diffeomorphism. This is known as the *Iwasawa decomposition.*

2.4. Reduced bases and Siegel domains. Let $L \subseteq \mathbb{R}^n$ be a lattice. Then L admits many different ordered bases.

Definition 2.5. We call an ordered basis $v_1, \dots, v_n$ of $\mathbb{R}^n$ *reduced* if, in Iwasawa coordinates as above,

- $|x_{ij}| \leqslant 1/2$ for $1 \leqslant i < j \leqslant n$, and
- $a_j/a_{j+1} \geqslant \sqrt{3}/2$ for $1 \leqslant j < n$.

What matters in this definition is that $1/2 < \infty$ and $\sqrt{3}/2 > 0$ are fixed; the precise numerical values will not be of much significance for us.

We summarize some of the basic facts concerning reduced bases of lattices, due to Minkowski and Siegel. In what follows the notation $A \ll B$ signifies that $|A| \leqslant C|B|$ where C depends only upon n, while $A \asymp B$ means that $A \ll B \ll A$.

Theorem 2.6. *Let $L \subseteq \mathbb{R}^n$ be a lattice.*

(i) There is at least one reduced basis.

(ii) Let $v_1, \dots, v_n$ be any reduced basis, with accompanying notation as above. Let T be the linear transformation sending v_i' to v_i. Then T and its inverse have operator norm $\ll 1$; equivalently, for all $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, we have

$$\| \sum c_i v_i \| \asymp \| \sum c_i v_i' \| \asymp \max_{1 \leqslant i \leqslant n} a_i |c_i|. \tag{2.1}$$

In particular, $|v_i| \asymp |v_i'| = a_i$.

(iii) If $v_1, \dots, v_n$ and $w_1, \dots, w_n$ are two reduced bases, then $|v_i'| \asymp |w_i'|$. Thus the Iwasawa coordinates a_i of a lattice are "well-defined up to constants depending only upon n."

(iv) Let $C \geqslant 1$ be large enough in terms of n. Suppose that $v_1, \dots, v_n$ is a reduced basis. Let $i \in \{1, \dots, n-1\}$. Suppose that

$$a_i/a_{i+1} > C.$$

Then the $\mathbb{Z}$-module

$$\langle v_{i+1}, \dots, v_n \rangle$$

depends only upon L, not upon the choice of reduced basis.

(v) Let $v_1, \dots, v_n$ and $w_1, \dots, w_n$ be reduced bases of the same lattice. Let $\gamma \in \mathrm{GL}_n(\mathbb{Z})$ be the change of basis matrix relating the two bases, thus

$$w_i = \sum_j \gamma_{ij} v_j.$$

Then each entry γ_{ij} is $\ll 1$.

Proof. **We start with part (i).**

Let $v_n \in \mathcal{L}$ be a shortest vector (i.e., a nonzero vector of minimal length). Such a vector exists in view of the discreteness of $\mathcal{L}$, and is necessarily primitive. By definition, $v_n' = v_n$.

Let $\mathcal{L}_{n-1}$ denote the image of $\mathcal{L}$ under orthogonal projection to $v_n^{\perp}$; it is a lattice of rank $n-1$. Let $v_{n-1}' \in \mathcal{L}_{n-1}$ be a shortest vector (which exists, as $\mathcal{L}_{n-1}$ is again discrete). Lift it to a vector $v_{n-1} \in \mathcal{L}$ of shortest length. This lift is primitive. (The superscripted "prime" notation is consistent: by construction, v_{n-1}' is the orthogonal projection of v_{n-1} to $\langle v_n \rangle^{\perp}$.) By the $n = 2$ discussion, we have $|v_{n-1}'| \geqslant \frac{\sqrt{3}}{2} |v_n'|$.

Next, let $\mathcal{L}_{n-2}$ denote the orthogonal projection of $\mathcal{L}$ to $\langle v_{n-1}, v_n \rangle^{\perp}$, let $v'_{n-2} \in \mathcal{L}_{n-2}$ be a shortest vector, and let $v_{n-2} \in \mathcal{L}$ be any lift (necessarily primitive). Then

$$v_{n-2} = v'_{n-2} + a v'_{n-1} + b v'_n$$

for some real numbers a and b. (We have used here that $\langle v_{n-1}, v_n \rangle_{\mathbb{R}} = \langle v'_{n-1}, v'_n \rangle_{\mathbb{R}}$.) We are free to translate v_{n-2} by any integral multiple of v_{n-1} or v_n. Adjusting first by v_{n-1}, we may arrange that $|a| \leqslant 1/2$. Adjusting further by v_n, we may arrange further that $|b| \leqslant 1/2$. We observe now that $u := v'_{n-2} + a v'_{n-1}$ is the orthogonal projection of v_{n-2} to $\mathcal{L}_{n-1}$. By the minimality of v'_{n-1}, we have

$$|v'_{n-1}|^2 \leqslant |u|^2 = |v'_{n-2}|^2 + a^2 |v'_{n-1}|^2.$$

Using that $|a| \leqslant 1/2$, it follows readily that $|v'_{n-2}| \geqslant \frac{\sqrt{3}}{2} |v'_{n-1}|$.

Continuing in this way, we obtain the desired reduced basis.

We turn to part (ii). By definition, $v'_i T = v_i = v'_i + \sum_{j>i} x_{ij} v'_j$. Let us express this relation in terms of the orthonormal basis $\{v''_i\}$:

$$v''_i T = v''_i + \sum_{j>i} x_{ij} \frac{a_j}{a_i} v''_j.$$

Thus in this basis, the matrix of T is strictly upper-triangular, with each entry $\ll 1$. The same is then true for the inverse of T. By computing $\|v\|$ and $\|Tv\|$ with respect to $\{v''_i\}$, we deduce the desired estimate.

We turn to part (iii). Let $v_1, \dots, v_n$ and $w_1, \dots, w_n$ be reduced bases. By (ii) and symmetry, it is enough to show that for each $k \in \{1, \dots, n\}$,

$$|w_k| \gg |v_k|.$$

To that end, write $w_i = \sum_j c_{ij} v_j$ for some integers c_{ij}. Since the vectors $w_k, \dots, w_n$ are linearly independent, they cannot all be contained in the span of the vectors $v_{k+1}, \dots, v_n$. We may thus find $\ell \in \{k, \dots, n\}$ and $m \in \{1, \dots, k\}$ so that $c_{\ell m} \neq 0$. By (2.1) and the integrality of $c_{\ell m}$, we then have

$$|w_k| \geqslant \left(\frac{\sqrt{3}}{2} \right)^{\ell - k} |w_\ell| \gg |w_\ell| \asymp \max_j |c_{\ell j} v'_j| \gg |v'_m| \geqslant \left(\frac{\sqrt{3}}{2} \right)^{k-j} |v_k| \gg |v_k|,$$

as required.

We turn to part (iv). Let $v_1, \dots, v_n$ and $w_1, \dots, w_n$ be reduced bases. Suppose that

$$\langle v_{i+1}, \dots, v_n \rangle \neq \langle w_{i+1}, \dots, w_n \rangle. \tag{2.2}$$

Without loss of generality, suppose more precisely that $w_j \notin \langle v_{i+1}, \dots, v_n \rangle$ for some $j \in \{i+1, \dots, m\}$. We may write $w_j = \sum_k c_k v_k$ with $c_k \in \mathbb{Z}$. By hypothesis, there exists $k \in \{1, \dots, i\}$ with $c_k \neq 0$. By (2.1) and part (iii), we have

$$a_{i+1} \asymp |w_{i+1}| \gg |w_j| \gg |v_k| \gg |v_i| \asymp |w_i| \asymp a_i.$$

Thus there exists $C \geqslant 1$, depending only upon n, so that $a_{i+1}/a_i \leqslant C$ whenever (2.2) holds. Taking contrapositives, we obtain the required implication.

We turn, finally, to part (v).

Let $a_i = |v'_i|$ denote the Iwasawa coordinates of the first reduced basis. By part (iii), they are approximately those of the second basis: $a_i \asymp |w'_i|$. By part (ii), we have $a_i \asymp |w_i| \asymp \sum_j a_j|\gamma_{ij}|$. In particular, there is a constant C, depending only upon n, so that

$$|\gamma_{ij}| \leqslant C a_i/a_j. \tag{2.3}$$

A key observation (implicit already in the proof of (iv)) is that if $a_j/a_i > C$, then $Ca_i/a_j < 1$. Since γ has integral entries, this forces $\gamma_{ij} = 0$. This feature motivates that our argument will depend upon how far apart the neighboring a_i's are from one another.

Recall that $a_1 \ggg \dots \ggg a_n$.

Consider first the case that all a_i are roughly of the same size:

$$a_1 \asymp \dots \asymp a_n.$$

Then (2.3) implies that each $\gamma_{ij} \ll 1$, which was the desired conclusion.

Suppose next that, for some $m \in \{1, \dots, n\}$, the numbers $a_1, \dots, a_m$ are all of the same size, as are the numbers $a_{m+1}, \dots, a_n$, but those in the first group are significantly larger than those in the second:

$$a_1 \asymp \dots \asymp a_m, \quad a_{m+1} \asymp \dots \asymp a_n, \quad a_i/a_j > C \text{ for } i \leqslant m, j > m. \tag{2.4}$$

Then (2.3) implies that

- $\gamma_{ij} \ll 1$ whenever i, j are both $\leqslant m$ or both $> m$, and
- $\gamma_{ij} = 0$ whenever $i > m$ and $j \leqslant m$,

but does not immediately usefully constrain the remaining variables γ_{ij} for $i \leqslant m, j > m$. We may in any event write γ in block-diagonal form as

$$\gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ 0 & \Gamma_{22} \end{pmatrix}$$

where $\Gamma_{11} \in \mathrm{GL}_m(\mathbb{Z})$ and $\Gamma_{22} \in \mathrm{GL}_{n-m}(\mathbb{Z})$ have entries $\ll 1$. Note that the inverses of these last matrices are also of size $\ll 1$, in view of Cramer's rule and the fact that $\det(\Gamma_{ii}) = \pm 1$. Our remaining task is to estimate the entries of the $m \times (n-m)$ matrix Γ_{12}; we wish to show that those entries are $\ll 1$.

To that end, let $g \in \mathrm{GL}_n(\mathbb{R})$ denote the matrix with rows $e_j g = v_j$. Then γg has rows $e_j \gamma g = w_j$. The Iwasawa decomposition of g reads $g = xak$ where $|x_{ij}| \leqslant 1/2$, $a = \operatorname{diag}(a_1, \dots, a_n)$ and $k \in \mathrm{O}(n)$. The Iwasawa decomposition of γg satisfies similar bounds. We wish to combine the known bounds on g and γg to constrain Γ_{12}. To that end, it is convenient to write the Iwasawa components in block-diagonal form, like for γ. First, write

$$x = \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix},$$

where U_{11}, U_{22} are strictly upper-triangular. Next, taking into account that the numbers $a_1, \dots, a_m$ and also $a_{m+1}, \dots, a_n$ are of similar size, we write

$$\operatorname{diag}(a_1, \dots, a_m) = A_1 M_1,$$

$$A_1 := \operatorname{diag}(a_1, \dots, a_1), \quad M_1 := \operatorname{diag}(a_1/a_1, a_2/a_1, \dots, a_m/a_1)$$

and

$$\operatorname{diag}(a_{m+1}, \dots, a_n) = A_2 M_2,$$

$$A_2 = \operatorname{diag}(a_n, \dots, a_n), \quad M_1 = \operatorname{diag}(a_{m+1}/a_n, \dots, a_{n-1}/a_n, a_n/a_n),$$

say. The precise choice of the scalar matrices A_1, A_2 is unimportant; what matters is just that each entry of M_1 or M_2 has size $\asymp 1$. Thus

$$g = \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix} \begin{pmatrix} A_1 M_1 & 0 \\ 0 & A_2 M_2 \end{pmatrix} k.$$

After a bit of matrix multiplication, we find that

$$\gamma g = \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Gamma_{11} U_{11} M_1 A_1 & 0 \\ 0 & \Gamma_{22} U_{22} M_2 A_2 \end{pmatrix} k,$$

where

$$B := \Gamma_{12} \Gamma_{22}^{-1} + \Gamma_{11} U_{12} U_{22}^{-1} \Gamma_{22}^{-1}.$$

Since the matrices $\Gamma_{11}, \Gamma_{22}, U_{12}, U_{22}$ and their inverses are of size $\ll 1$, our goal of showing that Γ_{12} is of size $\ll 1$ will be achieved if we can verify that B is of size $\ll 1$.

To that end, we compute the Iwasawa decomposition of γg. Note that the matrices $\Gamma_{ii} U_i M_i$ and their inverses are each of size $\ll 1$, so the same holds for their Iwasawa components, say $\Gamma_{ii} U_i M_i = u^{(i)} a^{(i)} k^{(i)}$. The Iwasawa decomposition of γg thus reads

$$\gamma g = u' \begin{pmatrix} a^{(1)} A_1 & 0 \\ 0 & a^{(2)} A_2 \end{pmatrix} \begin{pmatrix} k^{(1)} & \\ & k^{(2)} \end{pmatrix} k,$$

where

$$u' = \begin{pmatrix} u^{(1)} & B u^{(2)} \\ 0 & u^{(2)} \end{pmatrix}.$$

Since the basis $w_1, \dots, w_n$ is reduced, we know that the individual above-diagonal entries of u' are of size at most $1/2$. Since $u^{(1)}, u^{(2)}$ and their inverses are of size $\ll 1$, it follows that B is likewise of size $\ll 1$, as required.

This completes the proof in the special case (2.4). The general case is similar. We may write $n = \alpha_1 + \dots + \alpha_r$ in such a way that, setting $n_0 := 0$ and $n_k := \alpha_1 + \dots + \alpha_k$, we have

- $a_i \asymp a_j$ whenever i, j lie in the same segment, i.e., $n_{k-1} < i, j \leqslant n_k$ for some k, and
- $a_i / a_j > C$ whenever $i > j$ and i, j lie in distinct segments.

We then write our matrices in terms of $r \times r$ blocks and argue as above. □

Definition 2.7. Let $\mathfrak{S}$ denote the set of all $g \in \mathrm{GL}_n(\mathbb{R})$ for which the corresponding basis $e_1 g, \dots, e_n g$ of $\mathbb{R}^n$ is reduced. Thus $\mathfrak{S}$ consists of those g whose Iwasawa coordinates $g = xak \in N \times A^+ \times K$ satisfy $|x_{ij}| \leqslant 1/2$ $(i < j)$ and $a_i / a_{i+1} \geqslant \sqrt{3}/2$.

Theorem 2.8. *We have*

$$\mathrm{GL}_n(\mathbb{Z}) \mathfrak{S} = \mathrm{GL}_n(\mathbb{R}). \tag{2.5}$$

Moreover,

$$\#\{\gamma \in \mathrm{GL}_n(\mathbb{Z}) : \gamma \mathfrak{S} \cap \mathfrak{S} \neq \emptyset\} < \infty. \tag{2.6}$$

Proof. We apply Theorem 2.6.

Part (i) implies (2.5). Indeed, let $g \in \mathrm{GL}_n(\mathbb{R})$. Let $\mathbb{Z}^n g$ denote the corresponding lattice, with basis $e_1 g, \dots, e_n g$ given by the rows of g. Choose a reduced basis $v_1, \dots, v_n$ of g. Let $\gamma \in \mathrm{GL}_n(\mathbb{Z})$ denote the corresponding change of basis matrix, characterized by $v_i = \sum_j \gamma_{ij} e_j g = e_i \gamma g$. Then $\gamma g \in \mathfrak{S}$, as required.

Part (v) implies (2.6). Indeed, suppose $\gamma \in \Gamma$ satisfies $\gamma g \in \mathfrak{S}$ for some $g \in \mathfrak{S}$. Let $v_1, \dots, v_n$ and $w_1, \dots, w_n$ denote the rows of γ and γg, respectively. These define reduced bases of the same lattice $\mathbb{Z}^n g$. The hypotheses for part (v) are thus satisfied, and we obtain $|\gamma_{ij}| \leqslant C$ where C depends only upon n. The set of integral matrices satisfying this last condition is finite. □

We call $\mathfrak{S}$ the standard *Siegel domain* for $\mathrm{GL}_n(\mathbb{Z})$. It contains a fundamental domain for $\mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$, but is "not much bigger" than a fundamental domain in the sense quantified by the above results. The same results apply to the larger domains defined like $\mathfrak{S}$, but with $1/2$ replaced by something larger and $\sqrt{3}/2$ by something smaller.

There are obvious analogues of Definition 2.7 and Theorem 2.8 for $\mathrm{SL}_n(\mathbb{R})$.

For future reference, we introduce the following terminology.

Definition 2.9. Let $G = \mathrm{GL}_n(\mathbb{R})$ or $\mathrm{SL}_n(\mathbb{R})$. Write $G = NAK$ for its Iwasawa decomposition, where now A consists of diagonal elements with positive entries. A *Siegel domain* for G is a subset of the form $\mathfrak{S} = \omega A_t K$, where $\omega \subseteq N$ is compact and $A_T := \{a \in A : a_i/a_{i+1} \geqslant t\}$.

The examples above arise by taking $\omega = \{u \in N : |u_{ij}| \leqslant B \text{ for all } i < j\}$.

Exercise 2.2. Adapt the proof of Theorem 2.6 to show that any discrete subgroup $\Lambda \subseteq \mathbb{R}^n$ is of the form $\Lambda = \oplus_{i=1}^m \mathbb{Z} v_i$ for some linearly independent subset $\{v_1, \dots, v_m\} \subseteq \mathbb{R}^n$. Deduce that the following conditions on a subgroup $\Lambda \subseteq \mathbb{R}^n$ are equivalent.

(i) Λ is a lattice in the sense defined earlier (i.e., it is the $\mathbb{Z}$-module generated by some $\mathbb{R}$-basis of $\mathbb{R}^n$).
(ii) Λ is discrete and cocompact.

Exercise 2.3 (Minkowski's theorem). Let L be a lattice with reduced basis $v_1, \dots, v_n$. As usual, set $a_i = |v_i'|$.

(i) Show that $\mathrm{vol}(\mathbb{R}^n/L) \asymp a_1 \cdots a_n$,
(ii) Show that L contains a nonzero vector of length $\ll \mathrm{vol}(\mathbb{R}^n/L)^{1/n}$.

Exercise 2.4. Give an explicit example of a lattice $\mathcal{L} \subseteq \mathbb{R}^2$ with reduced bases v_1, v_2 and w_1, w_2 so that v_1, v_2 is not a permutation of $\pm w_1, \pm w_2$ for any combination of signs.

Exercise 2.5 (Mahler's theorem). Let L be a lattice. The *successive minima* of L are the numbers $M_1 \leqslant \dots \leqslant M_n$ described as follows: M_k is the smallest positive real for which L contains a k-element linearly independent subset $\{w_1, \dots, w_k\}$ with each $|w_i| \leqslant M_k$. Show that for any reduced basis $v_1, \dots, v_n$ of L, we have $M_k \asymp a_{n-k+1}$, where as usual $a_k = |v_k'|$.

Exercise 2.6 (Mahler compactness criterion). Let S be a subset of the space of unimodular lattices. Show that the following are equivalent.

(i) S is precompact.
(ii) There exists $\delta > 0$ so $|v| \geqslant \delta$ for all $0 \neq v \in L \in S$.

2.5. Haar measure.

For a summary with references, see [1, §2.9].

On any locally compact group G, there is a unique (up to scaling) measure $d^L g$ (resp. $d^R g$) on G invariant under left (resp. right) translation by G. These are called left (resp. right) Haar measures. If G is a Lie group and ω^L is a nonzero volume form invariant under left translation, then one may construct such a Haar measure explicitly by $\int f(g)\,d^L g = \int f \cdot \omega^L$ (and similarly with "left" replaced by "right"). The left and right Haar measures are related (up to scaling) by $d^R g = \delta(g)\,d^L g$, where $\delta : G \to \mathbb{R}_+^\times$ is the continuous homomorphism given by the magnitude of the determinant of the adjoint representation:

$$\delta(g) := |\det \mathrm{Ad}(g)|, \quad \mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}).$$

We call δ the *modulus character* of G.

We call G *unimodular* if δ is trivial, i.e., $\delta(g) = 1$ for all $g \in G$; in that case, we may take $d^L g = d^R g$ and speak unambiguously of a (left and right) Haar measure dg.

Example 2.10. Any abelian group is unimodular. In particular,

- the additive group $G = \mathbb{R}$ is unimodular, with a Haar measure given by the Lebesgue measure dx, and
- the multiplicative group $G = \mathbb{R}^\times$ is unimodular, with a Haar measure is given by $d^\times x := \frac{dx}{|x|}$.

Example 2.11. Any nilpotent Lie group is unimodular: the adjoint representation admits a basis with respect to which it is strictly upper-triangular, hence has determinant one.

Example 2.12. Any compact group G is unimodular. Indeed, since $\delta : G \to \mathbb{R}_+^\times$ defines a continuous homomorphism and G is compact, the image of δ is a compact subgroup of $\mathbb{R}_+^\times$, but the only such subgroup is $\{1\}$. Compact groups are moreover of finite volume, so they admit a unique probability Haar measure.

Example 2.13. $\mathrm{SL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{R})$ and, more generally, all reductive algebraic groups G are unimodular. Indeed, the adjoint determinant is identically one. One can see this directly in these examples using the Iwasawa decomposition and the preceding examples. Alternatively, it suffices to check this at the level of complex points (e.g., for $\mathrm{GL}_n(\mathbb{C})$), in which case it follows from the fact that G possesses a Zariski dense compact subgroup K (e.g., $\mathrm{U}(n)$); Example 2.12 implies that δ is trivial on K, hence also on G.

Example 2.14. The two-dimensional affine group $B = \left\{ b = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R}, y \in \mathbb{R}^\times \right\}$ is not unimodular; we have $\delta(b) = |y|$ and may take

$$d^L b = \frac{dx\,dy}{|y|^2}, \quad d^R b = \frac{dx\,dy}{|y|}.$$

Example 2.15. Consider the group $B = NA$ of upper-triangular matrices in $\mathrm{GL}_n(\mathbb{R})$. A Haar measure on N is given by the product of Lebesgue measures in coordinates: $dx = \prod_{i<j} dx_{ij}$. A Haar measure on A is given by $da = \prod_i d^\times a_i$, where as before $d^\times x := \frac{dx}{|x|}$ denotes the Haar measure on the multiplicative group. The modulus character is given by $\delta(xa) = \prod_{i<j} a_i/a_j$, since for $z = (z_{ij})_{i \leqslant j} \in \mathrm{Lie}(B)$ we have $\mathrm{Ad}(xa)z = (a_i a_j^{-1} z_{ij})_{i \leqslant j}$. As left and right Haar measures on B, we may take

$$d^R(xa) = dx\,da, \quad d^L(xa) = dx\,\frac{da}{\delta(a)}.$$

Example 2.16. Let $G = \mathrm{GL}_n(\mathbb{R})$, let $B = NA$ be as above, and let $K = \mathrm{O}(n)$. Then a Haar measure dg on G is given by

$$\int_G f(g)\,dg = \int_{b\in B}\int_{k\in K} f(bk)\,d^L b\,dk,$$

where $d^L b$ is a left Haar measure on B (as described above) and dk is a Haar measure on K.

To see this, observe first that the RHS defines a measure μ on G. Using that the map $B \times K \to G$ has surjective differential (as observed in our discussion of the Iwasawa decomposition), we see that μ is absolutely continuous with respect to any Haar measure dg on G, say $\mu = \alpha\,dg$ for some function α on G. The invariance properties of $d^L b$ and dk imply that α is left B-invariant and right K-invariant, hence $\alpha(bk) = \alpha(1)$ for all $(b,k) \in B \times K$. Since $BK = G$, it follows that α is constant.

We might as well take $d^L b$ as in Example 2.15 and dk as in Example 2.12.

Example 2.17. A Haar measure on $G = \mathrm{SL}_n(\mathbb{R})$ may be constructed similarly; we take $k \in \mathrm{SO}(n)$, restrict to a with $a_1 \cdots a_n = 1$, and omit the factor $d^\times a_n$. Explicitly, we may take

$$\int_G f = \int_{x\in N}\int_{\substack{a_1,\dots,a_n>0:\\ a_1\cdots a_n=1}}\int_{k\in K} f(xak)\,dx\,\frac{da}{\delta(a)}\,dk, \quad da = \frac{da_1}{a_1}\cdots\frac{da_{n-1}}{a_{n-1}}. \tag{2.7}$$

Given a locally group G and a closed subgroup H, both unimodular and equipped with Haar measures dg and dh, the quotient space G/H admits a unique G-invariant quotient measure $d\bar{g}$ for which $dg = d\,\bar{g}\cdot dh$; this last identity means that for any $f \in C_c(G)$,

$$\int_G f(g)\,dg = \int_{G/H}\left(f(\bar{g}h)\,dh\right)\,d\bar{g}.$$

Similar considerations apply to $H\backslash G$. We will refer to such quotient measures simply as Haar measures.

In particular, for any choices of Haar measures on $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{R})$, we obtain quotient measures on $\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{Z})\backslash\mathrm{GL}_n(\mathbb{R})$; here the discrete subgroups $\mathrm{SL}_n(\mathbb{Z})$ and $\mathrm{GL}_n(\mathbb{Z})$ are equipped with counting measure.

Exercise 2.7. Let μ be a Haar measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$. Show that there exists $c > 0$ with the following property. Let $f \in C_c(\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R}))$. Suppose that f is right-invariant under $\mathrm{SO}(2)$, so that it may be identified with a function $\varphi : \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \to \mathbb{C}$. Then

$$\int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})} f\,d\mu = c\int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} \varphi(x+iy)\,\frac{dx\,dy}{y^2}. \tag{2.8}$$

2.6. Finiteness of the volume.

Theorem 2.18. *The quotient* $\mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$ *has finite volume with respect to any Haar measure.*

Proof. Let $\mathfrak{S} \subseteq \mathrm{SL}_n(\mathbb{R})$ denote the standard Siegel domain, defined as in Definition 2.7. By Theorem 2.8, it is enough to verify that $\mathfrak{S}$ has finite volume with respect to the given Haar measure. Since all Haar measures are multiples of one another, it doesn't matter which we work with, so let's use the Haar measure given by (2.7). The volume of $\mathfrak{S}$ with respect to this measure is

$$\int_{\substack{a_1\cdots a_n=1,\\ a_i/a_{i+1}\geqslant\sqrt{3}/2}} \frac{da}{\delta(a)}, \tag{2.9}$$

so our task is to verify that this last integral converges absolutely.

We introduce the change of variables

$$y_1 = a_1/a_2, \quad y_2 = a_2/a_3, \ldots, y_{n-1} = a_{n-1}/a_n.$$

We can recover the a_i's from the y_j's and a_n: for $1 \leqslant i \leqslant n-1$,

$$a_i = y_i y_{i+1} \cdots y_{n-1} a_n.$$

We can determine a_n by calculating

$$1 = a_1 a_2 \cdots a_n = (y_1 \cdots y_{n-1} a_n)(y_2 \cdots y_{n-1} a_n) \cdots a_n = y_1 y_2^2 y_3^3 \cdots y_{n-1}^{n-1} a_n^n,$$

from which we deduce that

$$a_n = \prod_i y_i^{i/n}. \tag{2.10}$$

Up to the constant factor coming from the Jacobian of the change of variables $(a_i) \mapsto (y_j)$, we may rewrite the integral (2.9) as

$$\int_{y_1,\ldots,y_{n-1}=\sqrt{3}/2}^{\infty} \frac{1}{\delta(a(y))}\, d^\times y_1 \cdots d^\times y_{n-1}, \tag{2.11}$$

where

$$\delta(a(y)) := \begin{pmatrix} y_1 \cdots y_{n-1} a_n & & & & \\ & y_2 \cdots y_{n-1} a_n & & & \\ & & \cdots & & \\ & & & y_{n-1} a_n & \\ & & & & a_n, \end{pmatrix}$$

where a_n is defined as in (2.10), so that $\det(a(y)) = 1$. The point is now just that we may write

$$\delta(a(y)) = y_1^{\beta_1} \cdots y_{n-1}^{\beta_{n-1}}$$

for some *positive* real numbers $\beta_1, \ldots, \beta_{n-1}$, given explicitly by

$$\beta_j = j(n-j). \tag{2.12}$$

To see his, note that for $x = (x_{ij})_{i \leqslant j} \in \mathfrak{g} = \operatorname{Lie}(B)$, we have $\operatorname{Ad}(a(y))x = (y_i \cdots y_{j-1} x_{ij})_{i \leqslant j}$. For example, when $n = 3$,

$$\operatorname{Ad}(a(y))x = \begin{pmatrix} x_{11} & y_1 x_{12} & y_1 y_2 x_{13} & y_1 y_2 y_3 x_{14} \\ 0 & x_{22} & y_2 x_{23} & y_2 y_3 x_{24} \\ 0 & 0 & x_{33} & y_3 x_{34} \\ 0 & 0 & 0 & x_{44} \end{pmatrix}.$$

The matrix of eigenvalues of $\operatorname{Ad}(a(y))|_{\operatorname{Lie}(B)}$ is thus given by

$$\begin{pmatrix} 1 & y_1 & y_1 y_2 & y_1 y_2 y_3 \\ & 1 & y_2 & y_2 y_3 \\ & & 1 & y_3 \\ & & & 1 \end{pmatrix}.$$

The appearances of y_j form a square of dimensions $j \times (n-j)$, whence the claim (2.12).

Since $\int_{\sqrt{3}/2} y^{-\beta} \frac{dy}{y} < \infty$ for each $\beta > 0$, it follows that the integral (2.11) converges absolutely, as required. □

2.7. How many vectors does a random lattice have in a given region? Let $\mathbb{E}_L$ denote an integral over unimodular lattices L respect to the probability Haar measure on $X_n^{(1)} = \mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R})$ ($\mathcal{E}$ stands for "expectation").

As further illustration of the content of this section, we sketch a proof of the following theorem of Siegel.

Theorem 2.19. *Let Ω be a bounded open subset of $\mathbb{R}^n$ with $0 \notin \Omega$. Then*

$$\mathbb{E}_L |L \cap \Omega| = \mathrm{vol}(\Omega),$$

where the volume of Ω is computing using the usual Lebesgue measure. In words, for a random unimodular lattice, the expected number of lattice points in Ω is given by its volume.

It is not *a priori* obvious that the expectation in question is finite, in view of the following simple exercise.

Exercise 2.8. Assume $n \geqslant 2$. Show that for each nonempty open $\Omega \subseteq \mathbb{R}^n$ and $N \geqslant 0$, there is a unimodular lattice L with $|L \cap \Omega| \geqslant N$.

The proof requires a couple lemmas. First, we give some basic estimates and check that the integral converges. For $R > 0$, let $B_R \subseteq \mathbb{R}^n$ denote the ball of radius R.

Lemma 2.20. *Let $L \subseteq \mathbb{R}^n$. Choose a reduced basis, hence Iwasawa coordinates*

$$a_1 \gg \cdots \gg a_n \tag{2.13}$$

as usual. Then

$$|L \cap B_R| - 1 \ll \frac{R}{a_n} + \frac{R^2}{a_n a_{n-1}} + \cdots + \frac{R^n}{a_n \cdots a_1}. \tag{2.14}$$

Proof. The LHS counts the number of nonzero lattice vectors of size at most R. Let $v_1, \ldots, v_n$ be a reduced basis. Any nonzero element of L may be written $x = \sum x_i v_i$ where the x_i are integers, not all zero. By the estimate (2.1), we have $|x| \asymp \max_{1 \leqslant i \leqslant n} a_i |x_i|$. Thus, after enlarging R by a constant multiple, our task is to estimate

$$|\{0 \neq x \in \mathbb{Z}^n : \max_{1 \leqslant i \leqslant n} a_i |x_i| \leqslant R\}|.$$

The number of possibilities for x_i is at most $1 + R/a_i$, so the number of possibilities for x (excluding $x = 0$) is

$$\prod_{i=1}^{n} \left(1 + \frac{R}{a_i}\right) - 1.$$

We expand out the product and note that, thanks to (2.13), each term in the expansion is majorized by a corresponding term in (2.14). We are done. □

Lemma 2.21. *We have*

$$\mathbb{E}_L(|L \cap B_R| - 1) \ll R + R^n.$$

In particular, the indicated expectation is finite.

Proof. It suffices to bound the corresponding integral taken over the standard Siegel domain. By lemma 2.20, we reduce to verifying for each $0 \leqslant m \leqslant n-1$ that

$$\int_{\substack{a_1 \cdots a_n = 1 \\ a_i/a_{i+1} \geqslant \sqrt{3}/2}} \frac{1}{a_n \cdots a_{m+1}} \frac{da}{\delta(a)} < \infty.$$

We introduce the change of variables $a = a(y)$ as in the proof of Theorem 2.18. The key point is once again that

$$\frac{1}{a_n \cdots a_{m+1}} \frac{1}{\delta(a)} = y_1^{-\beta_1} \cdots y_{n-1}^{-\beta_n - 1} \tag{2.15}$$

where each $\beta_j > 0$.

To compute the β_j, it is convenient to rewrite

$$\frac{1}{a_n \cdots a_{m+1}} = a_1 \cdots a_m,$$

using here that $a_1 \cdots a_n = 1$. We compute as before that

$$a_n = \prod_{j=1}^{n-1} y_j^{-j/n},$$

so that

$$a_{m+1} = \prod_{j \leqslant m} y_j^{-j/n} \prod_{j > m} y_j^{(n-j)/n}.$$

Using the definition $y_i = a_i / a_{i+1}$, we compute moreover that

$$a_1 \cdots a_m = y_1 y_2^2 \cdots y_m^m a_{m+1}^m = \prod_{j \leqslant m} y_j^{j(1-m/n)} \prod_{j > m} y_j^{(n-j)/n}.$$

Recall, finally, that

$$\delta(a(y)) = \prod_j y_j^{j(n-j)}.$$

We obtain in summary that (2.15) holds with

$$j \leqslant m \implies \beta_j = j(n-j) - j(1-m/n) = j(n-j-1+m/n),$$
$$j > m \implies \beta_j = j(n-j) - (n-j)/n = (n-j)(j-1/n).$$

In either case, since $1 \leqslant j \leqslant n-1$, we have $\beta_j > 0$, as required. □

Exercise 2.9. Show that

$$\mathbb{E}_L(|L \cap B_R| - 1)^2 = \infty.$$

Remark 2.22. In the case $n = 2$, the calculations of Lemma 2.21 and Exercise 2.9 boil down to the following (cf. Exercise 2.7):

$$\int_{y=\sqrt{3}/2}^{\infty} y^{1/2} \frac{dy}{y^2} < \infty, \qquad \int_{y=\sqrt{3}/2}^{\infty} y \frac{dy}{y^2} = \infty.$$

Next, for a measurable function f on $\mathbb{R}^n$, define (if convergent) $\operatorname{Eis}[f] : X_n^{(1)} \to \mathbb{C}$ by $\operatorname{Eis}[f](L) := \sum_{0 \neq v \in L} f(v)$. Taking for f the characteristic function 1_Ω of a set Ω as above, we then have $\operatorname{Eis}[1_\Omega](L) = |L \cap \Omega| - 1$.

Suppose now that $f \in C_c(\mathbb{R}^n)$. The sums defining $\operatorname{Eis}[f]$ are then finite. Lemma 2.21 (applied with R large enough that B_R contains the support of f) implies that $\operatorname{Eis}[f]$ defines an integrable function on $X_n^{(1)}$. The map

$$C_c(\mathbb{R}^n) \ni f \mapsto \mathbb{E}_L \operatorname{Eis}[f](L),$$

being finite on its domain and nonnegative whenever f is, defines a measure μ on $\mathbb{R}^n$.

For $g \in \mathrm{SL}_n(\mathbb{R})$, set $R(g)f(v) := f(vg)$. Then $\mathrm{Eis}[R(g)f](L) = \mathrm{Eis}(f)(Lg)$. Thus the $\mathrm{SL}_n(\mathbb{R})$-invariance of the Haar measure on $X_n^{(1)}$ implies the $\mathrm{SL}_n(\mathbb{R})$-invariance of μ.

By arguments similar to those used to establish the uniqueness properties of Haar measure, one can verify that any $\mathrm{SL}_n(\mathbb{R})$-invairant measure μ on $\mathbb{R}^n$ is of the form

$$\mu = c_0 \delta_0 + c_1 \lambda, \tag{2.16}$$

where c_0, c_1 are scalars, δ is the δ-mass at the origin, and λ denotes Lebesgue measure. The main point in completing the proof is to verify that $c_0 = 0$ and $c_1 = 1$.

Lemma 2.23. $c_1 = 1$.

Proof. We consider the $R \to \infty$ asymptotics of $\mu(f_R)$, where f_R denotes the normalized characteristic function of B_R, i.e., $f_R = 1_{B_R} / \operatorname{vol}(B_R)$. We see first from Lemma 2.21 that $\limsup_{R \to \infty} \mu(f_R) < \infty$.

On the other hand, for each unimodular lattice L, it follows from Riemann integration in $\mathbb{R}^n$ that $\mathrm{Eis}[f_R](L) \to 1$ as $R \to \infty$. (Indeed, $\mathrm{Eis}[f_R](L)$ is one less than the number of points of L in the ball B_R, divided by the volume of B_R. Equivalently, it is one less than the number of points of the rescaled lattice $R^{-1}L$ in the ball B_1, divided by $R^n \operatorname{vol}(B_1)$. Since L is unimodular, the count $R^{-n}|R^{-1}L \cap B_1|$ is a Riemann sum approximation to the volume of B_1. Thus $\mathrm{Eis}[f_R](L) \to 1$.)

We see moreover from the proof of lemma 2.21 that the functions $\mathrm{Eis}[f_R]$ are all majorized on the standard Siegel domain by the same integrable function, a constant multiple of $\max_{1 \leqslant m \leqslant n} a_1 \cdots a_m$. By the dominated convergence theorem, we deduce that $\mu(f_R) = \mathbb{E}_L \mathrm{Eis}[f_R](L) \to 1$ as $R \to \infty$.

On the other hand, it is clear from (2.16) that $\mu(f_R) \to c_1$ as $R \to \infty$. Therefore $c_1 = 1$. □

Lemma 2.24. $c_0 = 0$.

Proof. The proof is similar to that of lemma 2.23, but we instead consider the behavior as $R \to 0$ of $\mu(f_R)$ with $f_R := 1_{B_R}$ the unnormalized characteristic function. We see as before that the $\mathrm{Eis}[f_R]$ are majorized by a common integrable function. On the other hand, for each unimodular lattice L, we have $\mathrm{Eis}[f_R](L) \to 0$ as $R \to 0$. Therefore $\mu(f_R) = \mathbb{E}_L \mathrm{Eis}[f_R](L)$ converges to 0 as $R \to \infty$. On the other hand, it also converges to c_0. □

2.8. Convergence of Eisenstein series. Here I largely follow the exposition of Borel [2, §11–12].

2.8.1. *Statement of result.* Let (G, Γ) be either

$$(\mathrm{GL}_n(\mathbb{R}), \mathrm{GL}_n(\mathbb{Z})) \text{ or } (\mathrm{SL}_n(\mathbb{R}), \mathrm{SL}_n(\mathbb{Z})).$$

We write as usual $G = NAK$ for the Iwasawa decomposition, here with A the group of diagonal matrices having positive entries.

Each n-tuple $s = (s_1, \dots, s_n) \in \mathbb{C}^n$ of complex numbers defines a character of A, i.e., a continuous homomorphism $A \to \mathbb{C}^\times$, given by

$$a \mapsto a^s := a_1^{s_1} \cdots a_n^{s_n}.$$

This character extends to a character of the group $B = NA$, given by $ua \mapsto a^s$, having trivial restriction to the subgroup $\Gamma_B := \Gamma \cap B$ of elements with integral entries.

Definition 2.25. We call $s \in \mathbb{C}^n$ *dominant* if

$$\mathrm{Re}(s_1) \geqslant \cdots \geqslant \mathrm{Re}(s_n)$$

and *strictly dominant* if

$$\mathrm{Re}(s_1) > \cdots > \mathrm{Re}(s_n).$$

We recall from §2.5 that the modulus character δ_B of B is given by

$$\delta(ua) = \prod_{i<j} a_i/a_j = a_1^{n-1} a_2^{n-3} \cdots a_n^{1-n}.$$

With the notation introduced above, we may write $\delta(ua) = a^{2\rho}$, where $\rho \in \mathbb{R}^n$ is the dominant element defined by

$$\rho = \left(\frac{n-1}{2}, \frac{n-3}{2}, \dots, \frac{1-n}{2} \right).$$

For $g \in G$, let us write $a(g) \in A$ for its Iwasawa component, so that $g \in Na(g)K$.

Theorem 2.26. *For $s \in \mathbb{C}^n$ with $s - 2\rho$ strictly dominant and $g \in G$, the series*

$$\sum_{\gamma \in \Gamma_B \backslash \Gamma} a(\gamma g)^s$$

converges absolutely, locally uniformly.

We note that, since the character $B \ni ua \mapsto a^s$ has trivial restriction to Γ_B, the function $\Gamma \ni \gamma \mapsto a(\gamma g)^s$ is left-invariant under Γ_B. Thus the terms in the above series are well-defined.

Corollary 2.27. *For given g, the above series defines a holomorphic of s in the indicated domain.*

The idea of the proof, developed leisurely over the next few subsections, is to compare the sum to an integral.

Exercise 2.10. Assuming that the theorem holds for $G = \mathrm{SL}_n(\mathbb{R})$, deduce it for $G = \mathrm{GL}_n(\mathbb{R})$.

We henceforth focus on the case $G = \mathrm{SL}_n(\mathbb{R})$.

2.8.2. *Recap on the finite-dimensional representation theory of* SL_n. Here we recall the classification of finite-dimensional representations of $G = \mathrm{SL}_n(\mathbb{R})$. The full contents of this section are not necessary for our immediate aims, but help place some of our arguments in context.

Definition 2.28. By a *weight* we mean an equivalence class of elements ω of $\mathbb{Z}^n$, with two such elements declared equivalent if their difference is a multiple of $(1, 1, \dots, 1)$.

We observe that if two elements of $\mathbb{Z}^n$ are equivalent, then they define the same character of $A \leqslant \mathrm{SL}_n(\mathbb{R})$, since any element $a \in A$ satisfies $a_1 \cdots a_n = 1$. The map $\omega \mapsto [a \mapsto a^{\omega}]$ identifies weights with polynomial characters of A.

Example 2.29. For $1 \leqslant m \leqslant n-1$, the element

$$\beta_m := (\underbrace{1, \dots, 1}_{m}, 0, \dots, 0) : a \mapsto a_1 \cdots a_m$$

is a nontrivial dominant weight. These are called the *fundamental weights*.

Exercise 2.11. The set of weights is a free $\mathbb{Z}$-module with basis given by the set $\{\beta_1, \dots, \beta_{m-1}\}$ of fundamental weights. A weight ω is dominant iff $\omega \in \sum_m \mathbb{Z}_{\geqslant 0} \beta_m$.

Definition 2.30. We say that $a \in A$ is *dominant* if $a_1 \geqslant \cdots \geqslant a_n$. Given weights ω_1 and ω_2, we say that ω_1 is *higher than* ω_2, and write $\omega_1 \geqslant \omega_2$, if $a^{\omega_1} \geqslant a^{\omega_2}$ for all dominant $a \in A$. We say that a weight ω is *positive* if $\omega \geqslant 0$.

The relation $\omega_1 \geqslant \omega_2$ defines a partial order on the set of weights. We note that if ω is dominant, then it is positive, but not conversely; for instance, $\omega = (1, -1, 0)$ is positive but not dominant.

Definition 2.31. By a *representation* of $G = \mathrm{SL}_n(\mathbb{R})$, we mean a pair (σ, V), where V is a finite-dimensional complex vector space and $\sigma : G \to \mathrm{GL}(V)$ is a continuous homomorphism. Such a representation is *irreducible* if it is nonzero and if it has, other than $\{0\}$ and V, no invariant subspaces (i.e., subspaces $W \leqslant V$ with $\sigma(G)W \subseteq W$). As a matter of notation, one often abbreviates $gv := \sigma(g)v$ for $g \in G$ and $v \in V$ when the homomorphism σ is understood from context.

Definition 2.32. Let (σ, V) be a representation of G. A nonzero vector $v \in V$ is said to be a *weight vector of weight* ω if $\sigma(a)v = a^{\omega} v$ for all $a \in A$.

Example 2.33. The *standard representation* (σ, V) is $V = \mathbb{C}^n$, with σ the "identity" map. The standard basis vectors $e_1, e_2, \dots, e_n$ satisfy

$$a e_j = a_j e_j$$

and are thus weight vectors, with weights $\beta_1 = (1, 0, \dots, 0), (0, 1, 0, \dots, 0), 0, \dots, 0, 1)$. Note that the weight of e_1 is higher than the weights of all the other basis elements.

Example 2.34. More generally, for $1 \leqslant m \leqslant n-1$, the mth exterior power $\Lambda^m \mathbb{C}^n$ of the standard representation admits the basis of weight vectors

$$e_{i_1} \wedge \cdots \wedge e_{i_m} \quad (1 \leqslant i_1 < \cdots < i_m \leqslant n)$$

with weights given by elements of $\mathbb{Z}^n$ consisting of m ones and $n-m$ zeros, corresponding to the characters $a \mapsto a_{i_1} \cdots a_{i_m}$. The vector $e_1 \wedge \cdots \wedge e_m$ has weight β_m, which is higher than all the other weights.

The previous examples are special cases of a more general result, most of which is known as the theorem of the highest weight.

Theorem 2.35. *For each dominant weight ω, there is an irreducible representation $(\sigma_\omega, V_\omega)$, unique up to isomorphism, with the following properties.*

1. *V_ω contains a unique (up to scaling) weight vector e_ω of weight ω that is moreover N-invariant.*
2. *V_ω admits a basis of weight vectors of weight $\leqslant \omega$.*

This association defines a bijection

$$\{\textit{dominant weights}\} \leftrightarrow \{\textit{irreducible representations up to isomorphism}\}.$$

Each V_ω admits an "integral structure," i.e., a $\mathbb{Z}$-submodule $E_\omega \subseteq V_\omega$ which is invariant by $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ and for which $V_\omega = E_\omega \otimes_{\mathbb{Z}} \mathbb{C}$. We can choose e_ω to lie in E_ω.

The weight ω is referred to as the highest weight of the representation, and e_ω as a highest weight vector.

Example 2.36. Suppose that $\omega = \beta_m$ is a fundamental weight. Then, following Example 2.34, we may take $V_\omega = \Lambda^m \mathbb{C}^n$. An integral structure E_ω is given by $\mathbb{Z}$-span of the basis elements $e_{i_1} \wedge \cdots \wedge e_{i_m}$. The highest weight vector $e_\omega := e_1 \wedge \cdots \wedge e_m$ then lies in E_ω.

We equip each such V_ω with a K-invariant inner product, hence a K-invariant Euclidean norm. By the discreteness of the submodule E_ω, we see that the numbers $\|v\|$ for $0 \neq v \in E_\omega$ are bounded away from zero. For instance, in the setting of Example 2.36, we may take the inner product with respect to which the indicated basis is orthonormal; we then have $\|v\| \geqslant 1$ for all $0 \neq v \in E_\omega$. We assume that the highest weight vector e_ω is a unit vector. We then have the following useful formula:

Lemma 2.37. *$a(g)^\omega = \|g^{-1} e_\omega\|^{-1}$ for all $g \in G$.*

Proof. Write $g = uak$. Then $g^{-1} e_\omega = k^{-1} a^{-1} u^{-1} e_\omega$. Recalling that

- $u^{-1} e_\omega = e_\omega$,
- $a^{-1} e_\omega = a^{-\omega} e_\omega$,
- $\|.\|$ is K-invariant, and
- e_ω is a unit vector,

we obtain

$$\|g^{-1} e_\omega\| = \|k^{-1} a^{-1} e_\omega\| = a^{-\omega} \|k^{-1} e_\omega\| = a^{-\omega},$$

as required. □

2.8.3. *Geometric lemmas.*

Lemma 2.38. *For each dominant weight ω and each $\gamma \in \Gamma$, we have $a(\gamma)^\omega \leqslant 1$.*

Proof. Since the dominant weights are nonnegative integral combinations of the fundamental weights, it suffice to consider the case that ω is a fundamental weight β_m. We've seen then that $\|v\| \geqslant 1$ for all $0 \neq v \in E_\omega$. Since E_ω contains e_ω and is Γ-invariant, we have in particular $\|\gamma^{-1} e_\omega\| \geqslant 1$ for all $\gamma \in \Gamma$. The required inequality follows now from Lemma 2.37. □

Example 2.39. Take $n = 2$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma = \mathrm{SL}_2(\mathbb{Z})$. One can check directly that

$$a(\gamma) = \begin{pmatrix} 1/\sqrt{c^2+d^2} & 0 \\ 0 & \sqrt{c^2+d^2} \end{pmatrix}, \quad a(\gamma)^{\beta_1} = 1/\sqrt{c^2+d^2}.$$

In the setting of the above proof, this amounts to noting that

$$\gamma^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \gamma^{-1} e_1 = \begin{pmatrix} d \\ -c \end{pmatrix}, \|\gamma^{-1} e_1\| = \sqrt{c^2+d^2}.$$

The inequality $a(\gamma)^{\beta_1} \leqslant 1$ thus amounts to that $c^2 + d^2 \geqslant 1$ for integers c and d, not both zero.

Definition 2.40. The *negative cone* ${}^{-}A \subseteq A$ is defined by

$$ {}^{-}A := \{a \in A : a^{\omega} \leqslant 1 \text{ for all dominant weights } \omega\}. $$

Using the relation between dominant and fundamental weights, we see that ${}^{-}A$ may be described by the smaller set of conditions

$$ a^{\beta_m} = a_1 \cdots a_m \leqslant 1 \quad \text{ for } 1 \leqslant m \leqslant n-1. $$

Corollary 2.41 (of Lemma 2.38). *For each $\gamma \in \Gamma$, we have $a(\gamma) \in {}^{-}A$.*

Exercise 2.12. Let $\mathfrak{S} \subseteq G$ be a Siegel domain. Let ω be a dominant weight. Verify that

$$ a(gx)^{\omega} \ll a(g)^{\omega} a(x)^{\omega} \quad \text{ for all } g \in G, x \in \mathfrak{S}, $$

where the implied constant may depend upon $(n, \mathfrak{S}, \omega)$ but not (g, x).

[Hint: using the Iwasawa decomposition, reduce to the case $g \in K$, then further to the case that g lies in a fixed compact set and x is a dominant element of A. Then apply Lemma 2.37.]

Exercise 2.13. Interpret the result of the previous exercise as follows. Let L be a unimodular lattice. Let $1 \leqslant m \leqslant n-1$.

(i) With notation as in Theorem 2.6, let $v_n, \dots, v_1$ be a reduced basis of L (it will be convenient to reverse the usual numbering) and $a_j := |v_j'|$ the corresponding Iwasawa coordinates, so that $a_1 \ll \dots \ll a_n$. Verify that $a_1 \cdots a_m$ is the covolume of the lattice $\oplus_{j=1}^m \mathbb{Z} v_j$ (inside its real span, equipped with the volume form induced by the restriction of the standard Euclidean metric on $\mathbb{R}^n$).

(ii) Let $w_1, \dots, w_n$ be any basis of L. Show that the covolume of the lattice $\oplus_{j=1}^m \mathbb{Z} w_j$ is $\gg a_1 \cdots a_m$.

Thus a reduced basis $v_n, \dots, v_1$ simultaneously minimizes (up to constants) the covolumes of each of the lattices $\oplus_{j=1}^m \mathbb{Z} v_j$ for $1 \leqslant m \leqslant n-1$.

It may be instructive at this point to draw a picture depicting the various conditions in play here.

One more easy lemma:

Lemma 2.42. *Let Ω be a compact subset of G. Then there is a compact subset Ω_A of A so that $a(xg) \in a(x)\Omega_A$ for all $x \in G$ and $g \in \Omega$.*

Proof. By enlarging Ω if necessary, we may assume that it is closed under left and right multiplication by the compact subgroup K of G. Write $x = uak$ (Iwasawa decomposition). Then $xg = uakg$ with $kg \in \Omega$. By the Iwasawa decomposition, we may find a compact subset $\Omega_A \subseteq A$ so that $\Omega \subseteq N \Omega_A K$. Then kg admits the Iwasawa decomposition $u'a'k'$ with $a' \in \Omega_A$. We have $xg = u(au'a^{-1})aa'k'$. Since A normalizes N, this shows that $a(xg) = aa' \in a\Omega_A$, as required. □

2.8.4. *Proof of Theorem 2.18.* We must show that if g belongs to a fixed compact subset of G and s belongs to a fixed compact subset of the region $\{s : s - 2\rho \text{ strictly dominant}\}$, then $\sum_{\gamma \in \Gamma_B \backslash \Gamma} a(\gamma g)^s \ll 1$.

By Lemma 2.42, we have for g and s as above and all $\gamma \in \Gamma$ the estimate

$$ a(\gamma g)^s \asymp a(\gamma)^s $$

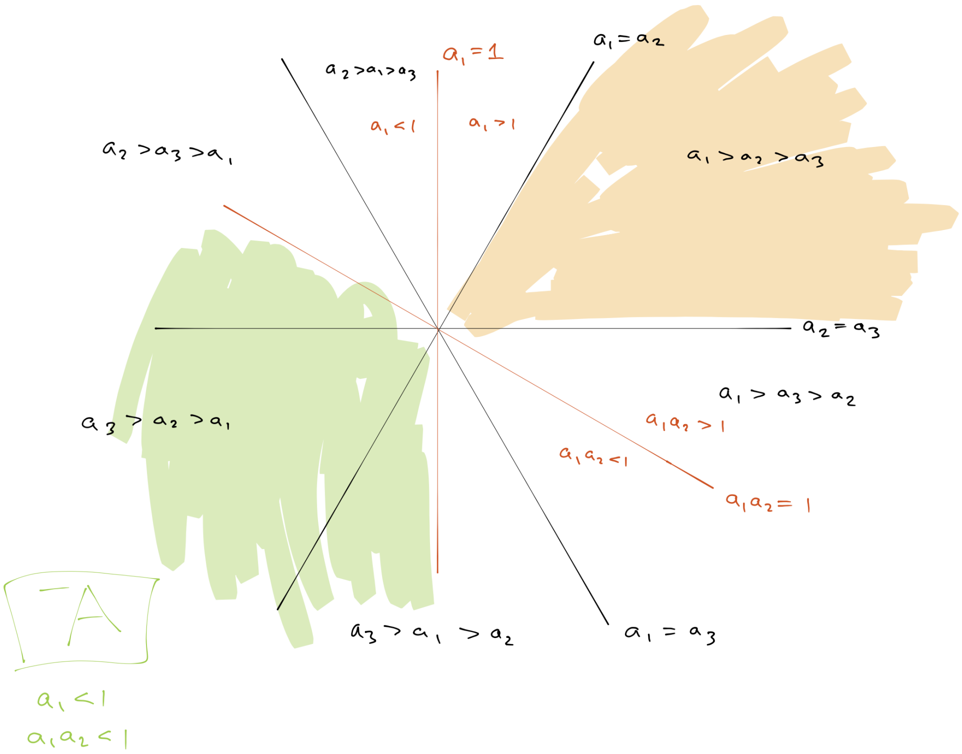

FIGURE 1. A "map" of the diagonal subgroup $A \leqslant \mathrm{SL}_3(\mathbb{R})$.

Moreover, if s lies in $\mathbb{R}^n$, then each of the above quantities is positive. The theorem will thus follow if we can show that for each $s \in \mathbb{R}^n$ with $s - 2\rho$ strictly dominant,

$$\sum_{\gamma \in \Gamma_B \backslash \Gamma} a(\gamma)^s < \infty.$$

Since the subgroup Γ of G is discrete, we may find a nonempty bounded neighborhood $\Omega \subseteq G$ of the identity element such that the translates $\gamma\Omega$ for $\gamma \in \Gamma$ are disjoint from one another. By another application of Lemma 2.42, we have

$$a(\gamma)^s \asymp a(\gamma g)^s$$

for all $\gamma \in \Gamma$ and $g \in \Omega$. We thereby reduce to verifying that

$$\int_{g \in \Omega} \sum_{\gamma \in \Gamma_B \backslash \Gamma} a(\gamma g)^s \, dg < \infty.$$

The sets $\Gamma_B \gamma \Omega$ for coset representatives $\gamma \in \Gamma_B \backslash \Gamma$ are disjoint from one another, so it is enough to verify that

$$\int_{g \in \Gamma_B \backslash \Gamma \Omega} a(g)^s \, dg < \infty,$$

where now dg denotes the quotient Haar measure on $\Gamma_B \backslash G$.

Let $\gamma \in \Gamma$ and $g \in \Omega$. Let ω be a dominant weight. By Lemma 2.42, we have $a(\gamma g)^\omega \ll a(\gamma)^\omega a(g)^\omega$. Recalling from Lemma 2.38 that $a(\gamma)^\omega \leqslant 1$ and from, e.g., Lemma 2.42 that $a(\gamma)^{\omega} \ll 1$, we obtain

$$a(\gamma g)^{\omega} \ll 1.$$

This last estimate tells us that each $g \in \Gamma\Omega$ satisfies $a(g) \in c \cdot {}^{-}A$ for some fixed $c \in A$, depending at most upon Ω. In other words,

$$\Gamma\Omega \subseteq N \cdot c \cdot {}^{-}A \cdot K.$$

We thereby reduce to verifying that

$$\int_{g \in \Gamma_B \backslash N \cdot c \cdot {}^{-}A \cdot K} a(g)^s \, dg < \infty.$$

We now recall from Example 2.17 that a Haar measure dg may be given in Iwasawa coordinates $g = uak$ by $du \, \frac{da}{a^{2\rho}} \, dk$. The quotient $\Gamma_N \backslash N$ is compact because it is the image of the compact set $\{u \in N : |u_{ij}| \leqslant 1/2 \forall i < j\}$, as one verifies readily either directly or by adopting the proof of Theorem 2.6. The group K is compact. We thereby reduce to verifying the finiteness of

$$\int_{a \in {}^{-}A} a^{s-2\rho} \, da \tag{2.17}$$

whenever $s - 2\rho$ is strictly dominant, where as usual da denotes a Haar measure on A, e.g., that given by

$$da = \frac{da_1 \cdots da_{n-1}}{a_1 \cdots a_{n-1}}.$$

At this point, we might as well substitute $s \mapsto s + 2\rho$ to reduce to verifying that

$$\int_{a \in {}^{-}A} a^s \, da \tag{2.18}$$

is finite whenever s is strictly dominant.

To understand the last integral (2.18), we change coordinates to

$$t_m = a_1 \cdots a_m \quad (1 \leqslant m \leqslant n-1).$$

In these coordinates, the domain ${}^{-}A$ is described by the conditions $t_m \leqslant 1$ for all m. Moreover, if define $\nu_1, \dots, \nu_{n-1}$ by

$$\nu_m := s_m - s_{m+1} > 0,$$

then we have

$$s_m = \nu_m + \nu_{m+1} + \cdots + \nu_{n-1} + s_n$$

and

$$a^s = \prod_{m=1}^{n-1} t_m^{\nu_m}.$$

The integral (2.18) of interest thus factors as the product of one-dimensional integrals

$$\int_{t_m=0}^{1} t_m^{\nu_m} \frac{dt_m}{t_m},$$

each of which converge. The proof is thus complete.

Exercise 2.14. For $\gamma \in \Gamma = \mathrm{SL}_n(\mathbb{Z})$, write $a_1(\gamma), \dots, a_n(\gamma) > 0$ for the Iwasawa coordinates. Note that these depend only upon the coset $\Gamma_B \gamma$. For $1 \leqslant m \leqslant n-1$, define $t_m(\gamma) := a_1(\gamma) a_2(\gamma) \cdots a_m(\gamma)$.

Show that for $X_1, \dots, X_{n-1} \geqslant 1$,

$$\# \{\gamma \in \Gamma_B \backslash \Gamma : t_m(\gamma) \geqslant 1/X_m \text{ for all } 1 \leqslant m \leqslant n-1\} \ll (X_1 \cdots X_{n-1})^2.$$

Show also that for $X \geqslant 1$,

$$\# \{\gamma \in \Gamma_B \backslash \Gamma : a_i(\gamma)/a_{i+1}(\gamma) \geqslant 1/X \text{ for all } 1 \leqslant i \leqslant n-1\} \ll X^p$$

for some $p = p(n) > 0$.

Exercise 2.15. Let $\mathfrak{S} \leqslant G$ be a Siegel domain, and $\Omega \subseteq G$ a compact subset.

(1) Show that there exists $c \geqslant 1$ with the following property. Let $x \in \mathfrak{S}$ and $\gamma \in \Gamma$. Suppose that $\gamma x \in N x \Omega$. Let $1 \leqslant i < j \leqslant n$ be indices such that the Iwasawa coordinates $a_1, \dots, a_n$ of x satisfy $a_i / a_j \geqslant c$. Then $\gamma_{ji} = 0$.

(2) Show that for each $g \in \mathfrak{S}$,

$$\#\{\gamma \in \Gamma : g^{-1} \gamma g \in \Omega\} \ll a(g)^{2\rho}.$$

(3) Show that there is a bounded subset of $\mathfrak{S}$ containing all $g \in \mathfrak{S}$ for which $\gamma g \in N \Omega$ for some $\gamma \in \Gamma$.

3. Basics on automorphic forms

At least initially, we'll follow in large part the exposition of Borel [2, §2].

3.1. Motivation. Sometimes it's useful to break problems up into analytic and algebraic parts. For instance, as a first step towards the analysis of spaces of functions on the circle group $\mathbb{R}/\mathbb{Z}$, one can study the algebraically-flavored space of trigonometric polynomials, i.e., finite linear combinations of the functions $x \mapsto e(nx)$ ($n \in \mathbb{Z}$). It's straightforward to verify that the Fourier inversion formula and the Parseval relation hold for such functions. One can then deduce the corresponding assertions for more analytically-flavored spaces (e.g., the space $L^2(\mathbb{R}/\mathbb{Z})$ of square-integrable functions the space $C^\infty(\mathbb{R}/\mathbb{Z})$ of smooth functions, etc.) via limiting arguments. Similarly, on the real line, one can approach analytic questions concerning function spaces ($L^2(\mathbb{R})$, the Schwartz space $\mathcal{S}(\mathbb{R})$, etc.) by first studying algebraic questions concerning finite linear combinations of Fourier modes $x \mapsto e(\xi x)$ ($\xi \in \mathbb{R}$), or finite linear combinations of hermite functions, i.e., functions of the form $x \mapsto P(x) e^{-\pi x^2}$ for a polynomial P.

Automorphic forms are algebraically-flavored spaces of functions on $\Gamma \backslash G$.

3.2. Definition without explanation. We continue to focus on the examples $\Gamma \backslash G = \mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$. We accordingly let K denote $\mathrm{SO}(n)$ or $\mathrm{O}(n)$.

Definition 3.1. An *automorphic form* $\varphi : G \to \mathbb{C}$ is a continuous function such that

(A_1) $\varphi(\gamma x) = \varphi(x)$ for all $\gamma \in \Gamma$, $x \in G$

(A_2) φ is *right K-finite*, i.e., the span of $\{\varphi(\cdot k) : k \in K\}$ is finite-dimensional.

(A_3) φ is $\mathfrak{Z}(\mathfrak{g})$-*finite*, where $\mathfrak{Z}(\mathfrak{g})$ denotes the center of the universal enveloping algebra.

(A_4) φ is of *moderate growth*.

The conditions in this definition are elaborated below.

3.3. Norms on G. For $x \in G$, we define

$$\|x\|^2 := \sum_{i,j} \left(|x_{ij}|^2 + |(x^{-1})_{ij}|^2\right).$$

We observe that

$$\|xy\| \leqslant \|x\|.\|y\|, \quad \|x\| = \|x^{-1}\|,$$

that $\{x \in G : \|x\| \leqslant c\}$ is compact for each $c > 0$, and that for each compact subset Ω of G,

$$\|uxu'\| \asymp \|x\| \quad (u, u' \in \Omega, x \in G). \tag{3.1}$$

Exercise 3.1. Let $\mathfrak{S} \subseteq G = \mathrm{SL}_n(\mathbb{R})$ be a Siegel domain (Definition 2.9). Let ω be a dominant weight (§2.8.2). Show that there exists $m > 0$ such that for each compact subset $\Omega \subseteq G$ there exists C so that for all $\gamma \in \Gamma$, $x \in \mathfrak{S}$ and $g \in \Omega$,

$$a(x)^\omega \leqslant C\|\gamma x g\|^m.$$

[Hint: use that $\|h\| \asymp \|hg\|$ for $g \in \Omega$ and that if $x \in B$, then $a(x)^\omega = \|xe_\omega\|$.]

3.4. Functions of moderate growth. We say that a function $f : G \to \mathbb{C}$ is of *moderate growth* if there exists $m \in \mathbb{Z}$ so that

$$f(x) \ll \|x\|^m \quad (x \in G). \tag{3.2}$$

We say that f is of *rapid decay* if the estimate (3.2) holds for all $m \in \mathbb{Z}$.

It is clear that a sum or product of functions of moderate growth or rapid decay has the same property, while the product of one function of moderate growth with another function of rapid decay is of rapid decay. Thus the functions of moderate growth form a ring which contains the functions of rapid decay as an ideal.

3.5. Finite functions.

3.5.1. *Generalities.*

Lemma 3.2. *Let G be a locally compact group. Let $f : G \to \mathbb{C}$ be a continuous function (measurability would also suffice). The following conditions are equivalent:*

(i) The span of the right G-translates of f is finite-dimensional.
(ii) The span of the left G-translates of f is finite-dimensional.
(iii) The span of the $G \times G$-translates of f is finite-dimensional.
(iv) f arises as a matrix coefficient for some finite-dimensional representation (π, V) of G: there exists $\alpha \in \mathrm{End}(V)^*$ so that $f(g) = \alpha(\pi(g))$.

Definition 3.3. A *finite function* on G is one satisfying the equivalent conditions of the lemma.

Example 3.4. For $G = \mathbb{R} = \mathbb{R}/\mathbb{Z}$, the finite functions are the trigonometric polynomials, i.e., the finite linear combinations of $x \mapsto e(nx)$.

Example 3.5. For a compact group G, any finite-dimensional representation decomposes as a direct sum of irreducible representations, so the finite functions are the finite linear combinations of matrix coefficients of irreducible finite-dimensional representations. The Peter–Weyl theorem says that these are dense in $L^2(G)$.

Example 3.6. For $G = \mathbb{R}$, the finite functions are the finite linear combinations of the *exponential polynomials*

$$x \mapsto e^{\alpha x} x^{\beta}$$

defined for $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{Z}_{\geqslant 0}$.

These are also the generalized eigenfunctions of the differentiation operator $\partial_x = \frac{d}{dx}$, i.e., the function f for which $(\partial_x - \alpha)^{\beta+1} f = 0$.

Example 3.7. For $G = \mathbb{R}_+^\times \xrightarrow{\log} \mathbb{R}$, the finite functions are those of the form

$$y \mapsto y^{\alpha} (\log y)^{\beta},$$

with α and β as above. These are also called exponential polynomials.

3.5.2. *Isotypic decomposition.* Let K and G be as in Definition 3.1. Let $\varphi : G \to \mathbb{C}$ be a right K-finite continuous function. Write V for the (finite-dimensional) span of the right translates of φ. Then K acts on V via right translation. By Maschke's theorem, we may decompose V as a direct sum of irreducible representations of K.

Let σ be an irreducible representation of K. We say that φ has *K-type* σ if V is isomorphic to a direct sum of copies of σ. Define $e_\sigma \in C^\infty(K)$ by the formula

$$e_\sigma(k) := (\dim \sigma) \chi_\sigma(k),$$

where $\chi_\sigma(k) := \operatorname{trace}(\sigma(k))$ denotes the character of σ. Using the Schur orthogonality relations, we may check that φ has K-type σ if and only if the convolution

$$\varphi * e_\sigma : g \mapsto \int_{k \in K} \varphi(gk^{-1}) e_\sigma(k) \, dk,$$

defined using the probability Haar measure dk on K, coincides with φ, i.e., $\varphi * e_\sigma = \varphi$.

We say that V is *K-isotypic* if it has K-type σ for some σ. In general, we may write any K-finite φ as a finite sum of K-isotypic vectors, indexed by σ in some finite collection Ξ of (isomorphism classes of) irreducible representations of K. For any such Ξ, set

$$e_\Xi := \sum_{\sigma \in \Xi} e_\sigma.$$

Then a continuous function φ on G is K-finite if and only if it satisfies $\varphi * e_\Xi = \varphi$ for some such Ξ.

3.6. Universal enveloping algebra.

3.6.1. *General definition and construction.* Let $\mathfrak{g}$ be a Lie algebra.

Definition 3.8. A *universal enveloping algebra* $\mathfrak{U}(\mathfrak{g})$ is a unital associative algebra equipped with a morphism of Lie algebras $\mathfrak{g} \to \mathfrak{U}(\mathfrak{g})$ that is universal among such morphisms in the following sense: for each morphism of Lie algebras $\mathfrak{g} \to A$ whose target is unital associative algebra A, there is a unique morphism of associative algebras $\mathfrak{U}(\mathfrak{g}) \to A$ such that the composition $\mathfrak{g} \to \mathfrak{U}(\mathfrak{g}) \to A$ is the given map $\mathfrak{g} \to A$.

By the usual arguments, if such an algebra exists, then it is unique. For this reason we typically speak of "the" universal enveloping algebra. We may construct $\mathfrak{U}(\mathfrak{g})$ explicitly as the quotient $T(\mathfrak{g})/I$, where

- $T(\mathfrak{g})$ denotes the tensor algebra of $\mathfrak{g}$, i.e., the graded algebra $\oplus_{k \in \mathbb{Z}_{\geqslant 0}} \mathfrak{g}^{\otimes k}$ equipped with the multiplication
$$(x_1 \otimes \cdots \otimes x_{k_1}) \cdot (y_1 \otimes \cdots \otimes y_{k_2}) = x_1 \otimes \cdots \otimes x_{k_1} \otimes y_1 \otimes \cdots \otimes y_{k_2},$$
and
- I denotes the ideal of $T(\mathfrak{g})$ generated by the elements $x \otimes y - y \otimes x - [x, y]$ for $x, y \in \mathfrak{g}$.

In other words, $\mathfrak{U}(\mathfrak{g})$ may be regarded as the non-commutative polynomial ring generated by elements $X \in \mathfrak{g}$ subject to the relations $XY - YX = [X, Y]$.

The proof that this construction defines a universal enveloping algebra is the same as the proof that, e.g., the usual construction of the tensor product of a pair of vector spaces satisfies its universal property.

3.6.2. *Construction via Lie theory.* Let G be a Lie group. We'll assume that it's a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Then $\mathfrak{g} := \mathrm{Lie}(\mathfrak{g})$ may be identified with a Lie subalgebra of the algebra $M_n(\mathbb{R})$ of $n \times n$ matrices.

The group G acts on itself by left and right translation, hence also on functions on G. For $g \in G$ and $f : G \to \mathbb{C}$, we set
$$r_g f(x) := f(xg), \quad \ell_g f(x) := f(g^{-1}x).$$
These define the right and left regular representations of G.

Exercise 3.2. Verify that $r_{g_1} r_{g_2} = r_{g_1 g_2}$ and $\ell_{g_1} \ell_{g_2} = \ell_{g_1 g_2}$.

For $X \in \mathfrak{g}$, we write $e^X = \exp(X) \in G$ for its exponential, which may be defined by the usual series. Thus
$$e^{tX} = \sum_{n \geqslant 0} \frac{t^n X^n}{n!}.$$

The group G acts on itself by conjugation. This induces the adjoint representation of G on $\mathfrak{g}$, denoted $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$. We have
$$g e^{tX} g^{-1} = e^{t \,\mathrm{Ad}(g) X}.$$

Each $X \in \mathfrak{g}$ defines a left-invariant differential operator on G, i.e., an element $[f \mapsto Xf]$ of $\mathrm{End}(C^\infty(G))$ given by
$$Xf(g) := \partial_{t=0} f(g e^{tX}).$$
In other words, $Xf = \partial_{t=0} r_{e^{tX}} f$ is obtained by differentiating the right regular representation of G. The Lie bracket $[,]$ on $\mathfrak{g}$ is characterized by the identity
$$X(Yf) - Y(Xf) = [X, Y]f \quad (f \in C^\infty(G)). \tag{3.3}$$
The map $\mathfrak{g} \to \mathrm{End}(C^\infty(G))$ is thus a Lie algebra homomorphism. This map is readily seen to be injective (consider the first order Taylor coefficients of f at the identity), so we may identify $\mathfrak{g}$ with its image.

The universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ may be defined as the subalgebra of $\mathrm{End}(C^\infty(G))$ generated by $\mathfrak{g}$. (It is clear from the universal propetry that $\mathfrak{U}(\mathfrak{g})$, as constructed in §3.6.1, admits a map to $\mathrm{End}(C^\infty(G))$. The point is just that this map is injective.) We may also identify $\mathcal{U}(\mathfrak{g})$ with the algebra of left-invariant differential operators on G, or with the convolution algebra of distributions on G supported at the identity element.

3.6.3. *The center.*

Definition 3.9. We denote by $\mathfrak{Z}(\mathfrak{g})$ the center of $\mathfrak{U}(\mathfrak{g})$.

Definition 3.10. In the setting of §3.2, we say that $\varphi : G \to \mathbb{C}$ is $\mathfrak{Z}(\mathfrak{g})$-*finite* if the image of its span under the action of $\mathfrak{Z}(\mathfrak{g})$ described in §3.6.2 is finite-dimensional.

Example 3.11. For $G = \mathrm{GL}_n(\mathbb{R})$, $\mathfrak{Z}(\mathfrak{g})$ is isomorphic to a polynomial ring in n variables, described explicitly via the Harish–Chandra isomorphism [8, p.220].

Example 3.12. For $G = \mathrm{SL}_2(\mathbb{R})$, $\mathfrak{Z}(\mathfrak{g})$ is isomorphic to a polynomial ring in one variable. It is freely generated by the element

$$\Omega = ef + fe + \frac{1}{2}h^2, \tag{3.4}$$

where

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

This is the simplest example of a "Casimir element." Note that in (3.4), the products ef, fe, h^2 take place inside $\mathfrak{U}(\mathfrak{g})$, not inside the space of 2×2 matrices.

3.7. Finiteness. The algebra $\mathfrak{Z}(\mathfrak{g})$ acts on the space of smooth functions φ on G (and also on, e.g., the space of distributions). The space $\mathfrak{Z}(\mathfrak{g})\varphi = \{D\varphi : D \in \mathfrak{Z}(\mathfrak{g})\}$ is isomorphic to the quotient algebra $\mathfrak{Z}(\mathfrak{g})/\operatorname{Ann}(\varphi)$, where $\operatorname{Ann}(\varphi) = \{D : D\varphi = 0\}$ denotes the annihilator ideal of φ. Thus that space is finite-dimensional if and only if that ideal is of finite codimension. In particular, φ is $\mathfrak{Z}(\mathfrak{g})$-finite if and only if it is annihilated by some ideal J in $\mathfrak{Z}(\mathfrak{g})$ of finite codimension. We say in the latter case that φ has $\mathfrak{Z}(\mathfrak{g})$-*type* J.

3.8. Relation to classical modular forms. Take $\Gamma\backslash G = \mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$, $K = \mathrm{SO}(2)$. Then G/K identifies with the upper half-plane $\mathbb{H} = \{x + iy : y > 0\}$ via the map $gK \mapsto g \cdot i$, where G acts on $\mathbb{H}$ via fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

Definition 3.13. A function $f : \mathbb{H} \to \mathbb{C}$ is a *modular form of weight* m if

(M_1) $f(\gamma z) = (cz+d)^m f(z)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ and $z \in \mathbb{H}$.
(M_2) f is holomorphic.
(M_3) f is "regular at the cusp."

Explanation of (M_3): from (M_1) we see that $f(z+1) = f(z)$, then from (M_2) that $f(z) = \sum_{n \in \mathbb{Z}} a_n e(nz)$, where as usual $e(z) := e^{2\pi i z}$. The factor $e(nz)$ blows up as $y \to \infty$ if $n < 0$. Condition (M_3) says that $a_n = 0$ unless $n \geqslant 0$, so that such blow-up does not occur.

The relation between Definition 3.13 and Definition 3.1 is as follows. Given a modular form f of weight m, define a function $\varphi : G \to \mathbb{C}$ by

$$\varphi(g) := (ci+d)^{-m} f(g \cdot i) \quad \text{if } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Exercise 3.3. Show that (M_1) implies (A_1).

Exercise 3.4. Show that (M_1) implies (A_2); more precisely, verify that

$$\varphi\left(g\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\right) = e(m\theta)\varphi(\gamma).$$

Exercise 3.5. Show that (M_2) implies (A_3); more precisely, with Ω as in Example 3.12, verify that $\Omega\varphi = c\varphi$ for some real number c depending only upon m.

Exercise 3.6. Show that (M_3) implies (A_4).

Exercise 3.7. Let $f : \mathbb{H} \to \mathbb{C}$ be a smooth function on the upper half-plane that is an eigenfunction for the hyperbolic Laplacian $\Delta = y^2(\partial_{xx} + \partial_{yy})$. Define φ as above, taking $m = 0$. Verify that (A_3) holds for φ; more precisely, show that φ is an eigenfunction for Ω, with eigenvalue deterimined by that for f with respect to Λ.

3.9. Some basic properties. Recall that a Lie group G comes with natural analytic charts, given by local exponential coordinates. We state here, without proof, some basic consequences of Definition 3.1.

Let $G = \mathrm{SL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{R})$.

Theorem 3.14. *Any automorphic form is analytic.*

More generally, any function $G \to \mathbb{C}$ that is right K-finite and $\mathfrak{Z}(\mathfrak{g})$-finite is analytic. The idea of the proof is that the K-finiteness and $\mathfrak{Z}(\mathfrak{g})$-finiteness conditions imply that φ satisfies an elliptic PDE. See for instance [1, Theorem 2.13] or [2, §3.1].

Theorem 3.15 (Harish–Chandra)**.** *For any automorphic form φ we may find $f \in C_c^\infty(G)$, supported arbitrarily close to the identity element, so that φ equals the convolution $\varphi * f$ defined as usual by*

$$\varphi * f(x) := \int_{g\in G} \varphi(xg^{-1})f(g)\,dg.$$

We may arrange moreover that $f(k^{-1}gk) = f(g)$ for all $k \in K, g \in G$.

This result again applies more generally to any right K-finite and $\mathfrak{Z}(\mathfrak{g})$-finite function on G. See for instance [1, Theorem 2.14], which develops a detailed proof in the special case $G = \mathrm{SL}_2(\mathbb{R})$, and [8, Corollary 8.41] or [2, §3.2] concerning the general case.

Exercise 3.8. Assuming the statement of [7, Theorem 1.14], deduce the conclusion of Theorem 3.15 in the setting of Exercise 3.7.

Definition 3.16. We say that a function $\varphi : G \to \mathbb{C}$ has *uniform moderate growth* if there exists m so that for each $D \in \mathcal{U}(\mathfrak{g})$, the function Df, defined as in §3.6.2, satisfies (3.2).

Proposition 3.17. *Any automorphic form φ has uniform moderate growth.*

Proof. Write $\varphi = \varphi * f$ with $f \in C_c^\infty(G)$. By elementary calculations with the definitions, we verify that

$$D\varphi = D(\varphi * f) = \varphi * Df.$$

On the other hand, using (3.1), we see that

$$\varphi(g) \ll \|g\|^m \implies (\varphi * Df) \ll \|g\|^m.$$

Therefore $D\varphi(g) \ll \|g\|^m$, as required. □

4. Constant terms

The main purpose of this section is to generalize the basic estimate

$$\sum_{n \geqslant 0} a_n e^{2\pi i n z} \sim a_0 \tag{4.1}$$

describing the asymptotic behavior as $y \to \infty$ of classical modular forms. More precisely, we initiate the study of the asymptotics "near infinity" in $\Gamma \backslash G$ of an automorphic form. In this section $\Gamma \backslash G$ is either $\mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$. We loosely follow the exposition of [2, §5–6].

4.1. Reduction from GL_n to SL_n. Recall from Example 3.7 that the finite functions on $\mathbb{R}_+^\times$ are the exponential polynomials. The finite functions on $\mathbb{R}^\times$ will also be referred to as exponential polynomials (they are given on each of the two connected components by an exponential polynomial).

Take $G := \mathrm{GL}_n(\mathbb{R})$. We may identify the center Z of G with $\mathbb{R}^\times$ via the map

$$\begin{pmatrix} z & & & & \\ & z & & & \\ & & \ddots & & \\ & & & z & \\ & & & & z \end{pmatrix} \mapsto z.$$

Using that G commutes with Z, we deduce that the Lie algebra $\mathfrak{z}$ of Z, regarded as a subspace of the universal enveloping algebra $\mathfrak{U}(\mathfrak{g})$, lies in the center $\mathfrak{Z}(\mathfrak{g})$. Thus any automorphic form φ on $\Gamma \backslash G$ is in particular $\mathfrak{z}$-finite. By analyticity, it is also Z-finite. Therefore for each $g \in G$, the function $Z \ni z \mapsto \varphi(zg)$ is an exponential polynomial. We get a slightly stronger conclusion by taking into account also the $\mathfrak{Z}(\mathfrak{g})$-finiteness of φ (see [2, §5.6] and references):

Lemma 4.1. *For each smooth $\mathfrak{Z}(\mathfrak{g})$-finite function $\varphi : G \to \mathbb{C}$, there are exponential polynomials Q_i on Z and elements $P_i \in \mathfrak{Z}(\mathfrak{g})$ such that for all $z \in Z$ and $g \in G$, we have*

$$\varphi(zg) = \sum_i Q_i(z) \cdot P_i f(g), \tag{4.2}$$

with $P_i f$ defined as in §3.6.2.

Since the multiplication map $Z \times \mathrm{SL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})$ has image of finite index and since exponential polynomials may be regarded as well-understood, we see that the study of the behavior near infinity of automorphic forms on $\mathrm{GL}_n(\mathbb{Z}) \backslash \mathrm{GL}_n(\mathbb{R})$ largely reduces to that of automorphic forms on $\mathrm{SL}_n(\mathbb{Z}) \backslash \mathrm{SL}_n(\mathbb{R})$.

4.2. Standard parabolic subgroups. Fix a Siegel domain $\mathfrak{S}$ for $\Gamma \backslash G$ (§2.4).

Question 4.2. *How can a sequence $x^{(\ell)} \in \mathfrak{S}$ ($\ell = 1, 2, 3, \dots$) tend off to ∞?*

Recall that we may write $x^{(\ell)} = u^{(\ell)} a^{(\ell)} k^{(\ell)}$, where $u^{(\ell)}$ lies in some fixed compact subset of N, $k^{(\ell)}$ lies in K, and $a^{(\ell)}$ is a diagonal matrix with positive entries for which each ratio $a_i^{(\ell)} / a_{i+1}^{(\ell)}$ is bounded from below by some fixed positive quantity. Thus after passing to a subsequence, we may assume that each such ratio either remains bounded or tends off to infinity as $\ell \to \infty$. Let $I \subseteq \{1, \dots, n-1\}$ denote the set of indices for which the ratio in question remains bounded. Then I describes the "directions" in which $x^{(\ell)}$ remains bounded in $\mathfrak{S}$, while the complement of I describes those in which it tends off to infinity.

It is often convenient to think of the set I in terms of a related equivalence relation $\sim$ on $\{1, \dots, n\}$. We declare $i \sim j$ when the ratio $a_i^{(\ell)}/a_j^{(\ell)}$ is bounded from above and below by fixed positive quantities, independent of ℓ. Thus for $i \leqslant j$, we have $i \sim j$ precisely when $\{i, i+1, \dots, j-1\} \subseteq I$. Such equivalence relations have the property that their equivalence classes consist of intervals of integers. Let us call such relations *continguous*. We obtain a natural bijection

$$\{\text{subsets } I \subseteq \{1, \dots, n-1\}\} \leftrightarrow \{\text{contiguous equivalence relations } \sim \text{ on } \{1, \dots, n\}\}.$$

We may, in turn, identify a continguous equivalence relation with the *composition* of n describing the cardinalities of each equivalence class. For instance, if $n = 8$ and $I = \{1, 3, 4, 7\}$, then the equivalence classes of $\sim$ are $\{1, 2\}$, $\{3, 4, 5\}$, $\{6\}$ and $\{7, 8\}$, and the corresponding composition is $n = 2 + 3 + 1 + 2$.

(In the lecture, the following definition was given a bit later, when it seemed better motivated.)

Definition 4.3. Given an equivalence relation $\sim$ on $\{1, \dots, n\}$ (or equivalently, a subset of $\{1, \dots, n-1\}$, or a composition of n), the corresponding *standard parabolic subgroup* P of G is defined by

$$P := \{g \in G : g_{ij} = 0 \text{ if } i > j \text{ and } i \not\sim j\}.$$

The corresponding *standard Levi subgroup* M of P is defined by

$$M := \{g \in G : g_{ij} = 0 \text{ if } i \not\sim j\}.$$

The *unipotent radical* U of P is given by

$$U := \{g \in N : g_{ij} = 0 \text{ if } i \sim j\}.$$

Such subgroups have already appeared implicitly: in our discussion of the shape of Siegel domains (see handwritten notes from Lecture 2) and in our proof of Theorem 2.6, part (v).

Exercise 4.1. With notation as above, we have $P = MU$. More precisely, P is the semidirect product of M and U, with U normal.

Example 4.4. Take $G = \mathrm{SL}_4$. Given a sequence $x^{(\ell)}$, we get an equivalence relation $\sim$ (equivalently, a subset of $\{0, 1, 2, 3\}$, or a composition of 4), hence a standard parabolic subgroup $P = MU$.

(i) If $x^{(\ell)} = \mathrm{diag}(\ell^2, \ell, \ell^{-1}, \ell^{-2})$, then

$$P = \begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{pmatrix}, \quad M = \begin{pmatrix} * & & & \\ & * & & \\ & & * & \\ & & & * \end{pmatrix}, \quad U = \begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{pmatrix}.$$

The corresponding subset and composition are

$$\emptyset \subseteq \{0, 1, 2, 3\}, \quad 4 = 1 + 1 + 1 + 1.$$

(ii) If $x^{(\ell)} := \mathrm{diag}(\ell^2, \ell^2, \ell^{-1}, \ell^{-3})$, then

$$P = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ & & * & * \\ & & & * \end{pmatrix}, \quad M = \begin{pmatrix} * & * & & \\ * & * & & \\ & & * & \\ & & & * \end{pmatrix}, \quad U = \begin{pmatrix} 1 & & * & * \\ & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{pmatrix}.$$

The corresponding subset and composition are

$$\{0\} \subseteq \{1,2,3\}, \quad 4 = 2+1+1.$$

(iii) If $x^{(\ell)} := \operatorname{diag}(\ell, \ell, \ell^{-1}, \ell^{-1})$, then

$$P = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ & & * & * \\ & & * & * \end{pmatrix}, \quad M = \begin{pmatrix} * & * & & \\ * & * & & \\ & & * & * \\ & & * & * \end{pmatrix}, \quad U = \begin{pmatrix} 1 & & * & * \\ & 1 & * & * \\ & & 1 & \\ & & & 1 \end{pmatrix}.$$

The corresponding subset and composition are

$$\{1,3\} \subseteq \{0,1,2,3\}, \quad 4 = 2+2.$$

(iv) If $x^{(\ell)} := \operatorname{diag}(\ell, \ell, \ell, \ell^{-3})$, then

$$P = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ & & & * \end{pmatrix}, \quad M = \begin{pmatrix} * & * & * & \\ * & * & * & \\ * & * & * & \\ & & & * \end{pmatrix}, \quad U = \begin{pmatrix} 1 & & & * \\ & 1 & & * \\ & & 1 & * \\ & & & 1 \end{pmatrix}.$$

The corresponding subset and composition are

$$\{1,2\} \subseteq \{1,2,3\}, \quad 4 = 3+1.$$

(v) If $x^{(\ell)} = 1$ for all ℓ, then $P = M = G$ and $U = \{1\}$. The corresponding subset and composition are

$$\{1,2,3\} \subseteq \{1,2,3\}, \quad 4 = 4.$$

Lemma 4.5. *Let $\mathfrak{S}$ be a Siegel domain for $\Gamma \backslash G$. Let $x^{(\ell)} \in \mathfrak{S}$ be a sequence. After passing to a subsequence, let $\sim$ denote the corresponding equivalence relation describing the directions along which the A-coordinates of this sequence escape to infinity. Let $P = MU$ denote the corresponding standard parabolic subgroup. Let $v \in U$. Define the sequence $g^{(\ell)} \in G$ by writing*

$$v x^{(\ell)} = x^{(\ell)} g^{(\ell)}.$$

Then $g^{(\ell)}$ tends to the identity element 1 of G as $\ell \to \infty$. The conclusion remains valid if we replace v by a sequence $v^{(\ell)}$ that traverses some fixed compact subset of U.

Proof. Consider the Iwasawa decomposition $x^{(\ell)} = u^{(\ell)} a^{(\ell)} k^{(\ell)}$. For notational simplicity in what follows, we drop the superscripted ℓ's and write simply $x = uak$, keeping in mind that all matrices considered may depend upon ℓ. We write

$$vx = uv'ak = uav''k = uakv''',$$

where

$$vu = v'u,$$
$$v'a = av'',$$
$$v''k = kv'''.$$

Thus $g^{(\ell)} = v'''$. Our aim is to show that v''' tends to the identity element of G as $\ell \to \infty$.

Since x lies in a fixed Siegel domain, we know that u lies in a fixed compact subset of N. Since N normalizes U (check this), it follows that v' lies in a fixed compact subset of U.

Next, we deduce by matrix multiplication that

$$v''_{ij} = \frac{a_j}{a_i} v'_{ij}.$$

Since $v' \in U$, we have for $i > j$ that $v'_{ij} \neq 0$ only if $i \not\sim j$; in that case, by the construction of the equivalence relation $\sim$, we know that the ratio a_j/a_i tends to zero as $\ell \to \infty$. Thus v'' tends to the identity element 1 of U as $\ell \to \infty$.

Since k lies in the fixed compact subset K of G, we conclude as required that v''' tends to the identity element of G. □

4.3. Approximation by constant terms.

4.3.1. *Informal discussion.* Let $\varphi : \Gamma\backslash G \to \mathbb{C}$ be an automorphic form. We have noted that we may then write $\varphi = \varphi * f$ for some $f \in C_c^\infty(G)$. Let us encode this relationship informally as the assertion that

$$\varphi(xg) \approx \varphi(x) \tag{4.3}$$

for all $x \in \Gamma\backslash G$ and all *small* $g \in G$, or in words, that φ enjoys some form of uniform continuity under right translation by G. The precise meaning is that this approximation becomes an identity when integrated against $f(g)\,dg$. We then have the following informal lemma.

Lemma 4.6 (Informal). *Let $\mathfrak{S}$ be a Siegel domain. Let $x^{(\ell)} \in \mathfrak{S}$ be a sequence. Let $P = MU$ denote the standard parabolic subgroup of G attached to $x^{(\ell)}$, as in Lemma 4.5, after passing to a suitable subsequence. Then for large ℓ and all $u \in U$, we have*

$$\varphi(x^{(\ell)}) \approx \varphi(ux^{(\ell)}).$$

Proof. Since φ is left-invariant under Γ_U and the quotient $\Gamma_U\backslash U$ is compact, we may take u to lie in a fixed compact fundamental domain for that quotient (e.g., that obtained by requiring that each above-diagonal entry have magnitude bounded by 1/2). Then, writing $ux^{(\ell)} = x^{(\ell)}g^{(\ell)}$, we have by Lemma 4.5 that $g^{(\ell)} \to 1$ as $\ell \to \infty$. The claim then follows from (4.3). □

By averaging the above estimate over u in the compact group $\Gamma_U\backslash U$ with respect to a probability Haar measure du, we obtain:

Corollary 4.7 (Informal). *With notation and assumptions as above, we have*

$$\varphi(x^{(\ell)}) \approx \int_{u \in \Gamma_U\backslash U} \varphi(ux^{(\ell)})\,du.$$

In the following subsections, we formulate a rigorous result in the spirit of the above informal discussion.

4.3.2. *Rigorous formulation.* Let $\varphi : \Gamma\backslash G \to \mathbb{C}$ be an automorphic form, and $\mathfrak{S}$ a Siegel domain. Let m be such that φ satisfies the "uniform moderate growth" condition (Definition 3.16) with exponent m.

Let $\mathfrak{S}$ be a Siegel domain, and suppose now that $\Gamma\backslash G = \mathrm{SL}_n(\mathbb{Z})\backslash\mathrm{SL}_n(\mathbb{R})$ (cf. §4.1). Let $x \in \mathfrak{S}$, and write $x = uak$. Then $a_1 \gg \cdots \gg a_n$ and $a_1 \cdots a_n = 1$. On the other hand, since u and k lie in fixed compact sets, we have (by Exercise 3.1)

$$\|x\| \asymp \|a\| \asymp \max(a_1, \dots, a_n, 1/a_1, \dots, 1/a_n) \asymp \max(a_1, 1/a_n) \ll a_1/a_n,$$

say. In particular, $\|x\|^m \ll a(x)^\lambda$ with λ the dominant weight defined by $a^\lambda := a_1/a_n$. The uniform moderate growth condition on φ thus implies that

$$D\varphi(x) \ll a(x)^\lambda \tag{4.4}$$

for fixed $D \in \mathfrak{U}(\mathfrak{g})$ and all $x \in \mathfrak{S}$.

Let $P = MU$ be a standard parabolic subgroup of G.

Definition 4.8. The *constant term* of φ with respect to P is the function $\varphi_P : G \to \mathbb{C}$ defined by

$$\varphi_P(g) := \int_{u \in \Gamma_U \backslash U} \varphi(ug)\, du,$$

where du denotes the probability Haar measure on $\Gamma_U \backslash U$.

Example 4.9. If $P = G$, then $M = G$ and $U = \{1\}$, so $\varphi_P = \varphi_G = \varphi$. In the remaining cases $P \subsetneq G$, this definition is more interesting.

The group A acts on U and its Lie algebra $\mathfrak{u}$ by conjugation. We denote below by β one of the (dominant) weights describing an eigenvalue for the latter action. Explicitly, β is uniquely of the form $a^\beta = a_i/a_j$ for some $i < j$ with $i \not\sim j$ in the equivalence relation attached to P.

Proposition 4.10. *Let φ be an automorphic form. Then for each fixed $N \geqslant 0$ and all $x \in \mathfrak{S}$,*

$$\varphi(x) - \varphi_P(x) \ll \max_\beta a(x)^{\lambda - N\beta}, \tag{4.5}$$

where β runs over eigenvalues for $A \circlearrowright \mathfrak{u}$ as above.

Example 4.11. Suppose P is *maximal* among proper standard parabolic subgroups. Thus P corresponds to a composition of the form $n = n_1 + n_2$ with $n_1, n_2 \geqslant 1$ (as in cases (iii) and (iv) of Example 4.4, but not the other cases). Then every β as in Proposition 4.10 is an integer multiple of the element α given by $a^\alpha = a_{n_1}/a_{n_1+1}$, and so the conclusion may be formulated equivalently as the estimate

$$\varphi(x) - \varphi_P(x) \ll a(x)^{\lambda - N\alpha}. \tag{4.6}$$

Before giving the proof of Proposition 4.10, let us indicate how one might apply this lemma in practice. Given $x \in \mathfrak{S}$, we may define (e.g., as in the above discussion concerning a sequence $x^{(\ell)}$) an equivalence relation $\sim$ on $\{1, \dots, n\}$ describing which ratios a_i/a_j are considered "large" and which are considered "small." We then form the corresponding standard parabolic subgroup $P = MU$. By construction, the quantities $a(x)^\beta$ are all "large." Since N can be taken arbitrarily large, it follows that the RHS of (4.5) is "small" – that is to say, that φ is well-approximated by its constant term φ_P on such elements x. Such conclusions may be understood as rigorous versions of Corollary 4.7.

4.3.3. *Proof when $n = 2$.* Here we record the proof of Proposition 4.10 in the simplest nontrivial case, in which $n = 2$ and P is the group of upper-triangular matrices. It will be convenient to introduce, for $x \in \mathbb{R}$ and $y \in \mathbb{R}_+^\times$, the notation

$$u(x) := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \quad a(y) := \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix}.$$

There is only one relevant value of β, for which $a(y)^\beta = y$. Our Siegel domain consists of elements $g = u(x)a(y)k$ with $|x| \leqslant 1/2$ and $y \geqslant t > 0$, say. Our hypothesis is that φ and its derivatives are majorized on such elements by y^{λ} for some $\lambda \geqslant 0$. Our task is thus to show, for $y \gg 1$, that

$$\varphi(g) - \varphi_P(g) \ll y^{\lambda - N\beta}. \tag{4.7}$$

for each fixed N. Note that this translates, via the recipe of §3.8, to weaker forms of estimates like (4.1).

Define the function $\phi : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ by

$$\phi(t) := \varphi(n(t)g).$$

Then $\varphi_P(g) = \int \phi := \int_0^1 \phi(t)\, dt$, so the quantity to be estimated in (4.7) is

$$\phi(0) - \int \phi = \int_0^1 (\phi(0) - \phi(t))\, dt = -\int_0^1 \int_0^t \phi'(s)\, ds\, dt = -\int_0^1 (1-s)\phi'(s)\, ds. \tag{4.8}$$

It will thus suffice to show that

$$\phi'(t) \ll y^{\lambda - N\beta}$$

for all t.

Our hypotheses (4.4) on ϕ concerns its derivatives with respect to left-invariant differential operators, induced by the action of G on $\Gamma\backslash G$ via right translation. To apply these here, we need to express the "ordinary" derivative $\phi'(t)$ in terms of such differential operators. This is achieved as follows. First, the Lie algebra element

$$X := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

of $\mathfrak{g} = \mathrm{Lie}(G)$ is an infinitesimal generator of the subgroup N. We indeed have

$$\phi'(t) = \frac{d}{dt}\varphi(e^{tX}g).$$

The RHS of this last identity may be understood as the evaluation of the derivative of φ at the tangent vector to G at g given by

$$Xu(x)a(y)k.$$

We can rewrite this tangent vector as

$$y^{-1}u(x)a(y)Xk,$$

using here the matrix conjugation identities $u(x)X = Xu(x)$ and $a(y)^{-1}Xa(y) = y^{-1}X$, and then further as

$$y^{-1}g(k^{-1}Xk).$$

In other words, we have shown that

$$\phi'(t) = y^{-1}\frac{d}{dt}\varphi(ge^{tX^k}), \quad X^k := k^{-1}Xk.$$

Now since k lies in the fixed compact subgroup $\mathrm{SO}(2)$, we know that X^k lies in some fixed compact subset of $\mathfrak{g}$. Therefore the uniform moderate growth condition (4.4) implies that

$$\phi'(t) \ll y^{\lambda - 1}.$$

This gives the required estimate in the special case $N = 1$. In general, we iterate the argument. We verify first, exactly as above, that

$$\phi^{(N)}(t) \ll y^{\lambda - N}.$$

We then use the integral formulas, for $j = 1, \dots, N-1$,

$$\phi^{(j)}(t) = -\int_0^1 (1-s)\phi^{(j+1)}(s+t)\,ds,$$

proved in the same way as (4.8), to show by descending induction on $j = N, N-1, \dots, 2, 1$ that

$$\phi^{(j)}(t) \ll y^{\lambda - N}.$$

The same estimate then holds for $\phi(t)$ thanks again to (4.8). The proof is now complete.

Remark 4.12. Another useful way to analyze differences as in (4.8) is via Fourier inversion, which gives for a smooth function $\phi : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ the identity

$$\phi(0) = \sum_{n \in \mathbb{Z}} a_n, \quad a_n := \int_{\mathbb{R}/\mathbb{Z}} \phi(x) e(nx)\,dx. \tag{4.9}$$

We have $a_0 = \int \phi$. On the other hand, integration by parts gives for any $j \in \mathbb{Z}_{\geqslant 0}$ that

$$a_n = (-2\pi i n)^{-j} \int_{\mathbb{R}/\mathbb{Z}} \phi^{(j)} e(nx)\,dx, \tag{4.10}$$

so that in particular

$$|a_n| \leqslant |2\pi n|^{-j} \|\phi^{(j)}\|_\infty. \tag{4.11}$$

Thus for j large enough that $\sum_{n \neq 0} |n|^{-j} < \infty$ (i.e., $j \geqslant 2$), we have

$$\phi(0) - \int \phi \ll \|\phi^{(j)}\|_\infty. \tag{4.12}$$

This last estimate may be applied in the above argument and in many other situations.

4.3.4. *The general argument.* In general, we find a filtration of U by subgroups

$$U = U_1 > U_2 > \cdots > U_q > U_{q+1} = \{1\},$$

with dimensions decreasing by one, in such a way that

- each Lie-algebraic quotient $\operatorname{Lie}(U_j)/\operatorname{Lie}(U_{j+1})$ corresponds to some eigenvalue β_j for $A \circlearrowright \mathfrak{u}$ as above, and
- each group-theoretic quotient U_j/U_{j+1} is central in U/U_{j+1}.

Example 4.13. In the case

$$G = \mathrm{SL}_3(\mathbb{R}), \quad P = \begin{pmatrix} * & * & * \\ & * & * \\ & & * \end{pmatrix}, \quad U = \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix},$$

we may take $q = 3$ and

$$U_2 = \begin{pmatrix} 1 & 0 & * \\ & 1 & * \\ & & 1 \end{pmatrix}, \quad U_3 = \begin{pmatrix} 1 & 0 & * \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Exercise 4.2. Verify in general that such a filtration exists.

We then define

$$\varphi_{q+1} := \varphi$$

and, by descending induction on $j = q, q-1, \dots, 1$,

$$\varphi_j(g) := \int_{u \in \Gamma_{U_j} \backslash U_j} \varphi_{j+1}(ug)\, du.$$

Exercise 4.3. Verify that $\varphi_P = \varphi_1$.

Since $\varphi = \varphi_{q+1}$, we have

$$\varphi - \varphi_P = \sum_{j=1}^{q} (\varphi_{j+1} - \varphi_j),$$

and our task reduces to verifying for each fixed $N \geqslant 0$ and all $x \in \mathfrak{S}$ that

$$\varphi_{j+1}(x) - \varphi_j(x) \ll a(x)^{\lambda - N\beta_j}. \tag{4.13}$$

The groups U_j/U_{j+1} are isomorphic to $\mathbb{R}$, while the further quotients $\Gamma_j \backslash U_j / U_{j+1}$ are isomorphic to $\mathbb{R}/\mathbb{Z}$. Indeed, if we let $X_j \in \mathfrak{u}$ denote the elementary matrix with eigenvalue β_j under the adjoint action of A, then the map

$$\mathbb{R}/\mathbb{Z} \ni t \mapsto u_j(t) := e^{tX_j} U_j$$

induces an isomorphism onto $\Gamma_j \backslash U_j / U_{j+1}$.

Example 4.14. In the setting of Example 4.13,

$$u_3(t) = \begin{pmatrix} 1 & 0 & t \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad u_2(t) = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & t \\ & & 1 \end{pmatrix}, \quad u_1(t) = \begin{pmatrix} 1 & t & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Now, given $j \in \{1, \dots, q\}$ and $x \in \mathfrak{S}$, define $\phi : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ by

$$\phi(t) := \varphi_{j+1}(u_j(t)x).$$

Clearly $\phi(0) = \varphi_{j+1}(x)$.

Exercise 4.4. Verify that $\int \phi = \varphi_j(x)$.

Our task (4.13) thereby reduces to showing that

$$\phi(0) - \int \phi \ll a(x)^{\lambda - N\beta_j}. \tag{4.14}$$

This last estimate may be proved by the same argument as in §4.3.3 (see [2, §6.7] or [9, §I.2.10] for detailed expositions).

4.4. Properties of the constant term. Above, we defined for each reasonable function $\varphi : \Gamma \backslash G \to \mathbb{C}$ and each standard parabolic subgroup $P = MU$ of G the constant term φ_P. Here we summarize some of its basic properties.

Observe that U is a normal subgroup of P, so $\Gamma_P U$ is a group, and we may form the quotient $\Gamma_P U \backslash G$. This quotient fits into the fiber bundle

$$\begin{tikzcd} \Gamma_M \backslash M \arrow[r] & \Gamma_P U \backslash G \arrow[d] \\ & P \backslash G. \end{tikzcd}$$

The quotient space $P \backslash G$ is compact, by the Iwasawa decomposition $G = PK$.

Suppose for instance that $G = \mathrm{GL}_n(\mathbb{R})$. Then $M \cong \prod_{j=1}^{k} \mathrm{GL}_{n_k}(\mathbb{R})$ (see the examples of §4.2), $\Gamma_M \backslash M \cong \prod_{j=1}^{k} \mathrm{GL}_{n_j}(\mathbb{Z}) \backslash \mathrm{GL}_{n_j}(\mathbb{R})$. If P is a *proper* parabolic (i.e., $P \neq G$), then each $n_j < n$. One has a similar but slightly more complicated description when $G = \mathrm{SL}_n(\mathbb{R})$. In either case, we can understand $\Gamma_P U \backslash G$ as the total space of a fiber bundle with compact base and fibers given by products of quotients much like $\Gamma \backslash G$, but with n replaced by something smaller. This feature often permits us to establish properties of $\Gamma \backslash G$ (or spaces of functions on it, etc.) by induction on n; we will see a first example later today.

Lemma 4.15. *φ_P is left-invariant under $\Gamma_P U$.*

Proof. Left-invariance under U follows from the right-invariance of the Haar measure du on $\Gamma_U \backslash U$: for $x \in G$ and $g \in U$,

$$\varphi_P(gx) = \int_{u \in \Gamma_U \backslash U} \varphi(ugx)\, du = \int_{u \in \Gamma_U \backslash U} \varphi(ux)\, du = \varphi_P(x).$$

To see the left invariance under $\gamma \in \Gamma_P$, use the left invariance of φ under Γ to write

$$\varphi_P(\gamma x) = \int_{u \in \Gamma_U \backslash U} \varphi(u\gamma x)\, du = \int_{u \in \Gamma_U \backslash U} \varphi(\gamma^{-1} u \gamma x)\, du = |c| \int_{u \in \Gamma_U \backslash U} \varphi(ux)\, du,$$

where $c \in \mathbb{R}^\times$ denotes the determinant of the adjoint action of γ on the Lie algebra $\mathfrak{u}$ of U. We may compute this determinant explicitly: writing $M = \mathrm{GL}_{n_1}(\mathbb{R}) \times \cdots \times \mathrm{GL}_{n_k}(\mathbb{R})$ with $n_1 + \cdots + n_k = n$ and writing $\gamma_j \in \mathrm{GL}_{n_j}(\mathbb{R})$ for the image of the projection of γ to M, we see that $c = \prod_{j=1}^{k} \det(\gamma_j)^{b_i}$ for some integers b_i. On the other hand, since γ lies in Γ_P, we have $\gamma_j \in \mathrm{GL}_{n_j}(\mathbb{Z})$, so $\det(\gamma_j) \in \{\pm 1\}$. Thus $|c| = 1$, as required. □

Definition 4.16. We say that a function $f : \Gamma_P U \backslash G \to \mathbb{C}$ is an *automorphic form* if it satisfies the following conditions, analogous to those in Definition 3.1.

(A_2) f is right K-finite.
(A_3) f is $\mathfrak{Z}(\mathfrak{m})$-finite, where $\mathfrak{Z}(\mathfrak{m})$ denotes the center of the universal enveloping algebra of $\mathfrak{m} = \mathrm{Lie}(M)$.
(A_4) f is of moderate growth.

Lemma 4.17. *Let φ be an automorphic form on $\Gamma \backslash G$. Then φ_P is an automorphic form on $\Gamma_P U \backslash G$.*

Proof. Right K-finiteness follows immediately from that of φ (note that φ_P is defined as an average of left translates of φ and that left translation commutes with right translation). The moderate growth property follows from that of f thanks to the observation (3.1) and the compactness of $\Gamma_U \backslash U$. It remains to verify the $\mathfrak{Z}(\mathfrak{m})$-finiteness property.

We observe first what this amounts to in the case that $G = \mathrm{SL}_2(\mathbb{R})$ and $\varphi = \varphi_f$ for some classical holomorphic modular form $f : \mathbb{H} \to \mathbb{C}$ of weight m (see §3.8) and $P = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$. Then φ_P is described in Iwasawa coordinates by the classical constant term $a_0(z) := \int_{u \in \mathbb{R}/\mathbb{Z}} f(z+u)\, du$. Of course $a_0(x+iy) = a_0(iy)$, so a_0 is really a function of one variable. On the other hand, it is well-known that the holomorphy of f implies that $a_0(iy)$ is independent of y; it is the constant term in the q-expansion $f(z) = \sum_{n \geqslant 0} a_n e(nz)$. To obtain a direct proof of this independence, one writes down the Cauchy–Riemann equations describing the holomorphy of f in Cartesian coordinates, and separates variables; one deduces in this way that the real and imaginary parts of $a_0(iy)$ have vanishing derivative with respect to y, hence that a_0 is indeed constant.

The general proof is similar, but requires a more explicit description of $\mathfrak{Z}(\mathfrak{g})$ than we have given thus far (see Example 3.11 and references). □

4.5. Eisenstein series defined via series.

Definition 4.18. Let $P = MU$ be a parabolic subgroup of G. Let $f : \Gamma_P U\backslash G \to \mathbb{C}$ and let $g \in G$. We define

$$\operatorname{Eis}(f)(g) := \sum_{\gamma \in \Gamma_P \backslash \Gamma} f(\gamma g), \tag{4.15}$$

provided that the series converges absolutely; in that case, we obtain a function

$$\operatorname{Eis}(f) : \Gamma\backslash G \to \mathbb{C}.$$

Example 4.19. Verify that if f is supported in a compact subset of $\Gamma_P U\backslash G$, then

(i) each sum (4.15) has only finitely many nonzero terms, and
(ii) $\operatorname{Eis}(f)$ is supported in a compact subset of $\Gamma\backslash G$.

[A simpler version of the arguments of §2.8 should suffice.]

Lemma 4.20. *For $f \in C_c(\Gamma_P U\backslash G)$ and $\varphi : \Gamma\backslash G \to \mathbb{C}$ locally integrable, we have*

$$\int_{\Gamma\backslash G} \operatorname{Eis}(f)\varphi = \int_{\Gamma_P U\backslash G} f\varphi_P, \tag{4.16}$$

where the measures are defined as follows:

- *The quotient $\Gamma\backslash G$ is equipped with the Haar measure coming from a given Haar measure dg on G and the counting measure on Γ.*
- *The quotient $\Gamma_P\backslash G$ is equipped with the Haar measure defined in the same way, so that Fubini's theorem holds with respect to the fiber bundle*

$$\begin{tikzcd} \Gamma_P\backslash\Gamma \arrow[r] & \Gamma_P\backslash G \arrow[d] \\ & \Gamma\backslash G \end{tikzcd} \tag{4.17}$$

with the counting measure on the fiber, i.e., for $\Phi \in C_c(\Gamma_P\backslash G)$,

$$\int_{\Gamma_P\backslash G} \Phi = \int_{g \in \Gamma\backslash G} \Big(\sum_{\gamma \in \Gamma_P\backslash\Gamma} \Phi(\gamma g)\Big)\, dg.$$

- *The Haar measure on $\Gamma_P U\backslash G$ is characterized by requiring that Fubini's theorem holds for the fiber bundle*

$$\begin{tikzcd} \Gamma_U\backslash U \arrow[r] & \Gamma_P\backslash G \arrow[d] \\ & \Gamma_P U\backslash G, \end{tikzcd} \tag{4.18}$$

where, as in Definition 4.8, the compact quotient $\Gamma_U\backslash U$ is equipped with the probability Haar measure.

Proof. We first use the left invariance of φ under Γ to write the LHS of (4.16) as

$$\int_{g\in\Gamma\backslash G}\sum_{\varphi\in\Gamma_P\backslash\Gamma}\varphi(\gamma g)f(\gamma g)\,dg.$$

We then apply Fubin's theorem twice, first for the fiber bundle (4.17) and then for (4.18), giving that the above is

$$\int_{\Gamma_P\backslash G}\varphi(g)f(g)\,dg=\int_{g\in\Gamma_P U\backslash G}\left(\int_{u\in\Gamma_U\backslash U}\varphi(ug)\,du\right)f(g)\,dg,$$

which in turn evaluates to the RHS of (4.16). □

Lemma 4.20 may be understood as asserting the adjointness of the two maps

$$[\varphi\mapsto\varphi_P]:\{\text{functions on }\Gamma\backslash G\}\to\{\text{functions on }\Gamma_P U\backslash G\}$$

and

$$[f\mapsto\operatorname{Eis}(f)]:\{\text{functions on }\Gamma_P U\backslash G\}\to\{\text{functions on }\Gamma\backslash G\},$$

when they are defined. This observation gives a sense in which the definition (4.15) of the Eisenstein series is inevitable: it describes the adjoint of the constant term map, which arose naturally in our study (e.g., §4.3.1) of the asymptotics of automorphic forms near the boundary of the quotient $\Gamma\backslash P$.

4.6. Cuspidal functions.

Definition 4.21. We say that a measurable, locally integrable function $\varphi:\Gamma\backslash G\to\mathbb{C}$ is *cuspidal* if for all standard parabolic subgroups $P\subsetneq G$, we have $\varphi_P=0$ almost everywhere. We write $L^2_0(\Gamma\backslash G)$ for the space of cuspidal elements of $L^2(\Gamma\backslash G)$. A *cusp form* is a cuspidal automorphic form.

For now, let us abbreviate these last two spaces simply by L^2_0 and L^2, respectively.

Lemma 4.22. *L^2_0 is a closed subspace of L^2.*

Proof. This is similar to the following simple consequence of Fubini's theorem: the subspace V of $L^2(\mathbb{R}^2)$, consisting of all φ for which $\int\varphi(x,y)\,dx=0$ for almost all y, is closed. This fact is contained in Fubini's theorem (or its proof). One way to understand this observation is as follows: V is the orthogonal complement in $L^2(\mathbb{R}^2)$ of the space of all functions of the form $\Phi_f(x,y):=f(y)$ for some $f\in C_c(\mathbb{R})$.

A similar argument applies in the setting of the lemma. Using Lemma 4.20, we see that L^2_0 is the orthogonal complement of the set of Eisenstein series $\operatorname{Eis}(f)$ attached to $f\in C_c(\Gamma_P U\backslash G)$ for some standard parabolic $P\subsetneq G$. Being an orthogonal complement, it is thus closed. □

Lemma 4.23. *Any cusp form is of rapid decay (Definition 3.4).*

Proof. We use that the constant terms φ_P of a cusp form φ vanish for each $P\subsetneq G$, and appeal to the approximation of automorphic forms by their constant terms. See [1, Cor 7.9] or [2, Thm 6.9] or [9, I.2.12, I.2.18] for details. □

5. Finiteness theorems

Here we record a generalization of the fact that the spaces $M_k(\mathrm{SL}_2(\mathbb{Z}))$ of holomorphic modular forms of given weight k are finite-dimensional.

Let $\Gamma\backslash G$ be either $\mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{Z})\backslash \mathrm{GL}_n(\mathbb{R})$. (Like almost all results presented in this course, these results are valid for much more general quotients; see the references below.) Given a finite collection Ξ of (isomorphism classes of) irreducible representations of K and a finite codimension ideal J of $\mathfrak{Z}(\mathfrak{g})$, we denote by

$$\mathcal{A}(\Gamma\backslash G, \Xi, J)$$

the space of automorphic forms on $\Gamma\backslash G$ having K-type Ξ and $\mathfrak{Z}(\mathfrak{g})$-type J (§3.5.2, §3.6). The space of automorphic forms is itself the union of these subspaces.

Theorem 5.1 (Harish–Chandra)**.** *Each of the above subspaces is finite-dimensional.*

Proof. We refer for now to [1, §8] for the special case $n = 2$ and to [2, Thm 7.4] for the general case. □

6. Convolution operators and decomposition of the space of cusp forms

We closely follow [1, §9] and [2, §8]. We refer to those references for summaries, with references, of relevant terminology from functional analysis, as well as proofs.

Theorem 6.1. *Let $\Gamma\backslash G = \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$, $f \in C_c^\infty(G)$. Then the operator T_f on $L^2_0(\Gamma\backslash G)$ is trace class (and in particular, compact).*

The key ingredient in the proof is the following estimate, proved by arguments similar to those in §4.3.

Theorem 6.2. *For $\varphi \in L^2_0(\Gamma\backslash G)$ and $f \in C_c^\infty(G)$, we have $\|\varphi * f\|_\infty \leqslant C(f)\|\varphi\|_2$, where $C(f)$ depends at most upon f.*

For the precise statement and proof of the following result, we refer for now to [1, Thm 16.2], [2, §9], or [6, §1.6].

Theorem 6.3. *The Hilbert space $L^2_0(\Gamma\backslash G)$, as a unitary representation of G, decomposes as a Hilbert direct sum of irreducible subrepresentations, each occurring with finite multiplicity.*

The main point of the proof was to show that there exist irreducible subrepresentations. By contrast, the regular representation of $\mathbb{R}$ on $L^2(\mathbb{R})$ has no irreducible subrepresentations. The result may be understood as an analogue for $\Gamma\backslash G$ of the density of the space of trigonometric polynomials inside $L^2(\mathbb{R}/\mathbb{Z})$.

References

[1] Armand Borel. *Automorphic forms on* $\mathrm{SL}_2(\mathbf{R})$, volume 130 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.

[2] Armand Borel. Automorphic forms on reductive groups. In *Automorphic forms and applications*, volume 12 of *IAS/Park City Math. Ser.*, pages 7–39. Amer. Math. Soc., Providence, RI, 2007.

[3] Armand Borel. *Introduction to arithmetic groups*, volume 73 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2019. Translated from the 1969 French original [MR0244260] by Lam Laurent Pham, Edited and with a preface by Dave Witte Morris.

[4] Armand Borel and Harish-Chandra. Arithmetic subgroups of algebraic groups. *Bull. Amer. Math. Soc.*, 67:579–583, 1961.

[5] Armand Borel and Harish-Chandra. Arithmetic subgroups of algebraic groups. *Ann. of Math. (2)*, 75:485–535, 1962.

[6] I. M. Gel′ fand, M. I. Graev, and I. I. Pyatetskii-Shapiro. *Generalized functions. Vol. 6.* AMS Chelsea Publishing, Providence, RI, 2016. Representation theory and automorphic functions, Translated from the 1966 Russian original [MR0220673] by K. A. Hirsch, Reprint of the 1969 English translation [MR0233772].

[7] Henryk Iwaniec. *Spectral methods of automorphic forms*, volume 53 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2002.

[8] Anthony W. Knapp. *Representation theory of semisimple groups*, volume 36 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1986. An overview based on examples.

[9] C. Mœ glin and J.-L. Waldspurger. *Spectral decomposition and Eisenstein series*, volume 113 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1995. Une paraphrase de l'Écriture [A paraphrase of Scripture].

[10] Dave Witte Morris. *Introduction to arithmetic groups*. Deductive Press, [place of publication not identified], 2015.

[11] Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

[12] Carl Ludwig Siegel. Einheiten quadratischer Formen. *Abh. Math. Sem. Hansischen Univ.*, 13:209–239, 1940.

ETH Zürich, Department of Mathematics, Rämistrasse 101, CH-8092, Zürich, Switzerland