# STABILITY OF LATTICES AND THE PARTITION
# OF ARITHMETIC QUOTIENTS*

BILL CASSELMAN[†]

**1. Introduction.** Elements of the group $G = \mathrm{SL}_2(\mathbb{R})$ act on the upper half plane

$$\mathcal{H} = \{z = x + iy \mid y > 0\}$$

by linear fractional transformations

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} : \quad z \longmapsto \frac{az + b}{cz + d} \, .$$

The arithmetic subgroup $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ acts discretely on $\mathcal{H}$, and as is well known it has as fundamental domain the region

$$\mathcal{D} = \{|z| \geq 1, |x| \leq 1/2\} \, .$$

On the other hand, let $P$ be the group of upper triangular matrices in $G$, containing the group $N$ of upper unipotent matrices. Thus elements of

$$\Gamma \cap P = \left\{ \begin{bmatrix} \pm 1 & n \\ 0 & \pm 1 \end{bmatrix} \, \middle| \, n \in \mathbb{Z} \right\}$$

act on $\mathcal{H}$ by horizontal integral translations $z \mapsto z + n$, and a fundamental domain for $\Gamma \cap P$ is therefore the region $\left\{ z \in \mathcal{H} \, \middle| \, x = |\mathrm{RE}(z)| \leq 1/2 \right\}$.

If for $Y > 0$ we define the region

$$\mathcal{H}_Y = \{z = x + iy \mid y > Y\}$$

then there are a number of properties it possesses that play an important role in analysis on the quotient $\Gamma \backslash \mathcal{H}$, for example in the construction of Eisenstein series and the proof of the Selberg trace formula:

- *The region $\mathcal{H}_Y$ is invariant under the group $N$ as well as the discrete subgroup $\Gamma \cap P$;*
- *The quotient by $N(\Gamma \cap P)$ is isomorphic to the subset $[Y, \infty)$;*
- *for $Y \geq 1$, the canonical projection from $\Gamma \cap P \backslash \mathcal{H}$ to $\Gamma \backslash \mathcal{H}$ when restricted to the quotient $\Gamma \cap P \backslash \mathcal{H}_Y$ embeds it as a neighbourhood of the cusp at infinity;*
- *the complement of the image of $\mathcal{H}_Y$ in $\Gamma \backslash \mathcal{H}$ is compact.*

[†]Department of Mathematics, University of British Columbia, Vancouver, Canada (cass@math.ubc.ca).

$y = Y$
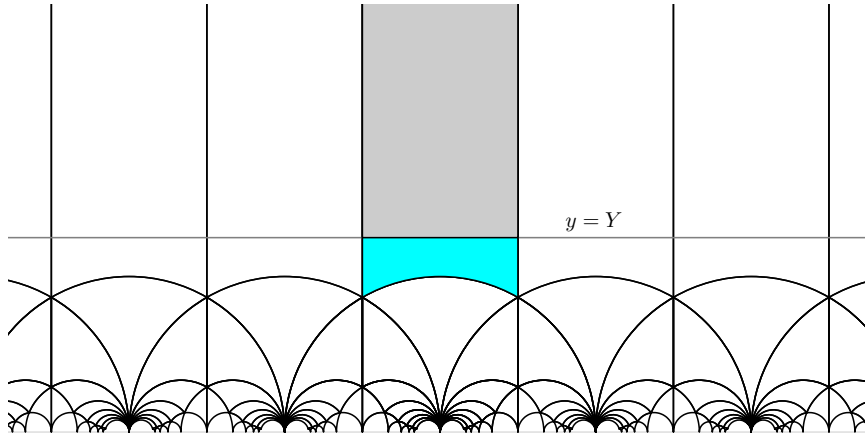
The second property may also be formulated as saying that *if $z$ and $\gamma(z)$ both lie in $\mathcal{H}_Y$ for some $\gamma$ in $\Gamma$ then $\gamma$ lies in $P$.*

In effect, we have a partition of $\Gamma \backslash \mathcal{H}$ into two parts, one a neighbourhood of infinity which is relatively simple, and the other a compact piece of the interior.

There is another way to formulate this result. Let $\mathcal{H}^*$ be the union of $\mathcal{H}$ and the rational cusps, the $\Gamma$-translates of $\infty$, which may be identified with the points of $\mathbb{P}^1(\mathbb{Q})$. If

$$\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then it takes $\infty$ to $a/c$ and $\mathcal{H}_Y$ to the disc centred at $(a/c, 1/2c^2Y)$ tangent to $\mathbb{R}$ at $a/c$. The stabilizers in $G$ of the cusps are the $\Gamma$-conjugates of $P$, and the $\Gamma$-transforms of the regions $\mathcal{H}_Y$, which are discs unless $\gamma$ lies in $P$, are the neighbourhoods of the cusps in the topology of $\mathcal{H}^*$ defined by Satake. The sets $\gamma \mathcal{H}_Y$, with $Y$ fixed, as $\gamma$ ranges over $\Gamma$ are disjoint (when not identical), and their union is stable under $\Gamma$ as is its complement in $\mathcal{H}$. The quotient of this complement by $\Gamma$ is compact.
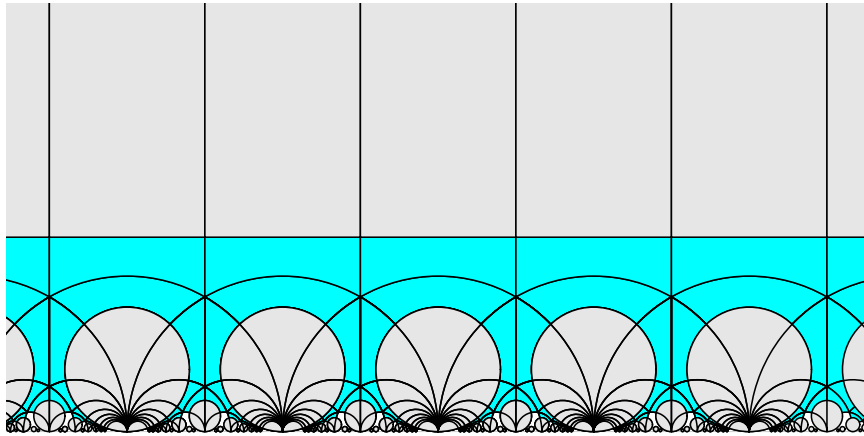


As far as I know, it was Jim Arthur who first generalized this result explicitly to arbitrary arithmetical quotients (in 1977), although I think it's fair to say that this generalization was already implicit in Satake's work on compactifications of arithmetic quotients. In Arthur's generalization the subsets of the partition are parametrized

by $\Gamma$-conjugacy classes of rational parabolic subgroups, which is also how Satake's rational boundary components are parametrized. Of course Arthur did this work with the intention of using it in dealing with his extension of the Selberg trace formula, but subsequently it has also been useful in other contexts.

In this note, which is largely expository, I will explain Arthur's partition for $\mathrm{GL}_n(\mathbb{Z})$, applying ideas almost entirely due to Harder, Stuhler, and Grayson, and including a self-contained account of their work.

A point $z = x + iy$ in $\mathcal{H}$ gives rise to the lattice generated by $z$ and 1. If we choose for this the basis 1 and $-z$, we obtain the positive definite symmetric form

$$Q_z(m, n) = \frac{1}{y}(m - nz)(m - n\overline{z}) = \frac{1}{y}\left(m^2 - 2xmn + n^2|z|^2\right)$$

(normalized so as to have discriminant equal to 1). Its matrix is

$$Q_z = \frac{1}{y}\begin{bmatrix} 1 & 0 \\ -x & -y \end{bmatrix}\begin{bmatrix} 1 & -x \\ 0 & -y \end{bmatrix} = \begin{bmatrix} 1/y & -x/y \\ -x/y & x^2 + y^2/y \end{bmatrix}.$$

On the other hand, the group $\mathrm{SL}_2(\mathbb{R})$ acts on the space of positive definite $2 \times 2$ symmetric matrices $Z$ by the transformations

$$Z \longmapsto {}^t g^{-1} Z g^{-1}.$$

I leave it as an exercise to verify that the two actions are the same—that for any $z$ in $\mathcal{H}$ we have $gQ_z = Q_{g(z)}$ for all $g$ in $\mathrm{SL}_2(\mathbb{R})$. The important part of the verification is that

$$\begin{aligned}
\begin{bmatrix} 1 & -z \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} &= \begin{bmatrix} 1 & -z \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\
&= \begin{bmatrix} (cz + d) & -(az + b) \end{bmatrix} \\
&= (cz + d)\begin{bmatrix} 1 & -(az + b)/(cz + d) \end{bmatrix}.
\end{aligned}$$

In higher dimensions we therefore have the following generalization of the classical theory. For any real vector space $V$, let $X = X_V$ be the space of all positive definite quadratic forms on $V$. For $V = \mathbb{R}^n$ this may be identified with $X_n$, the space of all positive definite symmetric $n \times n$ matrices, if we define

$$x(v) = {}^t v\, x\, v,$$

identifying $\mathbb{R}^n$ with column matrices. The space $X$ is a homogeneous space for $G = \mathrm{GL}(V)$, where an element $g$ in $G$ acts according to the rule $gQ(v) = Q(g^{-1}v)$. The subgroup acting trivially is $\pm I$. On $X_n$, this is equivalent to $x \mapsto {}^t g^{-1} x\, g^{-1}$. There is one peculiar point to mention. Although the classical action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ and that on $X_2$ agree, the corresponding actions of $\mathrm{GL}_2(\mathbb{Z})$ do not—the fractional linear transformations in $\mathrm{GL}_2(\mathbb{Z})$ take $\mathcal{H}$ to its conjugate, but the natural action of $\mathrm{GL}_2$ takes the connected space $X_2$ to itself.

If $x$ is a positive definite symmetric matrix, Gauss elimination applied to $x$ requires no row swapping and hence gives a factorization $x = \ell\, d\, u$ where $\ell$ is lower

triangular unipotent, $d$ diagonal, and $u$ upper triangular unipotent. Since $x$ is symmetric, $\ell = {}^t u$ and hence

$$x = {}^t u \, d \, u \, .$$

The action of $GL_n$ on $X$ is therefore transitive. The isotropy subgroup of the identity matrix (the sum of $n$ squares) is $K = \mathrm{O}_n(\mathbb{R})$, and therefore $X = X_n$ may be identified with $G/K$. In fact, it follows equally from Gauss elimination that the subgroup of upper triangular matrices acts transitively on $X_n$, and hence also any of its conjugates in $G$, or any group that contains one of its conjugates.

Let $\Gamma$ be the subgroup $GL_n(\mathbb{Z})$. The object of these notes is to show how ideas of [Grayson 1984] (which follows [Stuhler 1976], itself depending heavily on [Harder-Narasimhan 1975]) can be used to describe a parabolic decomposition of $\Gamma \backslash X$ used by Arthur and others in the theory of automorphic forms. A **flag** in the vector space $\mathbb{R}^n$ is an increasing sequence of vector subspaces. It is a **rational** flag if the subspaces are rational (defined by linear equations with coefficients in $\mathbb{Q}$). A **parabolic subgroup** of $G$ is the stabilizer of a flag, and a **rational parabolic subgroup** is the stabilizer of a rational flag. Any partition $n = n_1 + n_2 + \cdots + n_k$ of $n$ into positive numbers determines the rational flag

$$0 \subset \mathbb{R}^{n_1} \subset \mathbb{R}^{n_1+n_2} \subset \ldots \subset \mathbb{R}^n$$

and the **standard** parabolic subgroup associated to this partition is the stabilizer of this flag. If $P$ is any rational parabolic subgroup of $G$ with unipotent radical $N$ then there is a canonical surjection from $\Gamma \cap P \backslash X$ to $\Gamma \backslash X$. Arthur's result describes a simple $N$-invariant subset of $\Gamma \cap P \backslash X$ for which this map is an embedding, and partitions $\Gamma \backslash X$ into a disjoint union of the images of such embeddings as $P$ ranges over a set of representatives of $\Gamma$-conjugacy classes of rational parabolic subgroups. As remarked above, this is necessary in the theory of Eisenstein series, where functions in the continuous spectrum of $\Gamma \backslash X$ are constructed in terms of functions on the parabolic quotients $(\Gamma \cap P)N \backslash X$. Since $P$ contains a conjugate of the group of upper triangular matrices, $P$ acts transitively on $X$, which may be identified with the quotient $P/K \cap P$.

Let $\mathcal{C}$ be the acute cone $\{s_i \leq s_{i+1}\}$ in $\mathbb{R}^n$. The principal result of this paper, stated roughly, is that

> There exists a canonical map associating to each $x$ in $X_n$ a parabolic subgroup $P_x$ and a point $s_x$ in $\overline{\mathcal{C}}$ lying in the face of $\mathcal{C}$ naturally associated to $P$. The point $\gamma x$ maps to $\gamma P_x \gamma^{-1}$ and $s_x$, and the structure of the fibres of this map may be described recursively in terms of analogous maps on lower dimension symmetric spaces.

The space $\Gamma \backslash X$ is therefore partitioned by $\Gamma$-conjugacy classes of rational parabolic subgroups.

Existing discussions of these matters for arbitrary arithmetic groups can be found in [Arthur 1978], [Osborne-Warner 1983], [Saper 1994], and [Leuzinger 1995]. Another recent treatment, more arithmetical in flavour, can be found in the Trieste lectures [Harder-Stuhler 1997]. But techniques explained in the two papers [Grayson 1984] and [Grayson 1986] seem to me to be close to ideal, and provide as well an elegant derivation of classical reduction theory. Incidentally the authors of many of these papers often seem to be largely unaware of each other and particularly not to have known about the much older result stated in [Arthur 1978] (Lemma 6.4). Arthur works

with adèle groups, but his results are easily reformulated and proven for arithmetic ones (as has been done by Osborne and Warner).

One of the virtues of this approach is that it strengthens known analogies between symmetric varieties and the buildings of Bruhat-Tits associated to $p$-adic fields, for example those pointed out so strikingly in [Manin 1994].

The main reference here is [Grayson 1984], which considers symmetric spaces associated to $\mathrm{GL}_{n,F}$ for number fields $F$, as well as various orthogonal groups with respect to symmetric or anti-symmetric forms. In the second paper [Grayson 1986] he extends his techniques to an arbitrary semi-simple group defined over $\mathbb{Q}$. These papers of Grayson are just part of a large literature dealing with related material, perhaps originating with [Harder 1969]. In these notes the only new contribution is to explain the link between Grayson's ideas and those of Arthur, and I shall discuss in detail only $\mathrm{SL}_n$ and $\mathrm{GL}_n$. Incidentally, it seems to me that the theory explained in Grayson's papers for these groups and the orthogonal groups is just about perfect, whereas for other groups there are some loose ends to be tied up. As Grayson himself points out, for example, it would be interesting to handle arbitrary reductive groups in a similar spirit, whereas his current theory applies only to semi-simple ones. In this respect Grayson's theory again has points in common with the Bruhat-Tits theory. Another loose end in Grayson's papers is the role of relative discriminants. The recent paper [Harder-Stuhler 1997] deals with this question a little more precisely by discussing the reduction theory for Chevalley groups over number fields.

This paper was written mostly during a visit to the Université de Lyon I. Thanks are due to Fokko du Cloux for arranging the visit. Armand Borel spent much of his professional energy on the reduction theory of arithmetic groups, so it is appropriate that I dedicate this paper to his memory.

**2. The basic definitions.**   I follow Grayson in defining a **lattice** of rank $n$ to be a pair $\Lambda = (L_\Lambda, Q_\Lambda)$ where $L_\Lambda$ is an abelian group isomorphic to $\mathbb{Z}^n$ and $Q_\Lambda$ a Euclidean metric on it. Usually I'll just refer to the group $L$, with $Q$ implicit. The metric $Q$ also induces a Euclidean metric on the real vector space $V = L_\mathbb{R} = L \otimes \mathbb{R}$, and a uniform Riemannian metric on the torus quotient $V/L$. If $(\ell_i)$ is a basis of $L$ and $(e_j)$ is an orthonormal basis of $V$, then the volume of the parallelogram spanned by the $\ell_i$ is the absolute value of the determinant of the matrix $E$ with entries $\ell_i \bullet e_j$. The matrix $Q$ of the quadratic form with respect to the basis $\ell$, on the other hand, is that with entries $\ell_i \bullet \ell_j$. But the matrix $Q$ is also the matrix product ${}^t\!E\,E$, so that $\det(Q) = \det(E)^2$. A **unit** lattice is a lattice whose fundamental parallelograms in $L_\mathbb{R}$ have unit area, or equivalently $|\det(E)| = \det(Q) = 1$.

Two lattices are **isomorphic** to each other if there is an isomorphism of the groups inducing an isomorphism of metrics. Two lattices are **similar** if their metrics differ by a positive scalar. Our basic problem, here and more generally, is to describe as explicitly as possible the isomorphism classes of these structures.

If $L$ is a free subgroup of $V$ of maximal rank, then two quadratic forms $x_1$ and $x_2$ on $V$ give rise to isomorphic lattices based on $L$ if and only if $x_2 = \gamma x_1$ with $\gamma$ in $\mathrm{GL}(L)$. Thus the set $\mathcal{L}_L$ of isomorphism classes of lattices with free group $L$ may be identified with $\mathrm{GL}(L)\backslash X_V$. On the one hand, the group $L$ may be assumed to be $\mathbb{Z}^n$. If $Q$ is a positive definite metric on $\mathbb{Z}^n$ with associated inner product $\langle \bullet, \bullet \rangle$, then the matrix $(\langle e_i, e_j \rangle)$ is positive definite and symmetric. This leads to an identification

of the isomorphism classes of lattices associated of dimension $n$ with the arithmetic quotient $\mathrm{GL}_n(\mathbb{Z})\backslash X_n$, those of unit lattices with $\mathrm{GL}_n(\mathbb{Z})\backslash\mathfrak{X}_n$ where $\mathfrak{X}_n$ is the subset of matrices in $X_n$ with determinant 1. On the other hand, $L_{\mathbb{R}}$ may be identified with $\mathbb{R}^n$ and the quadratic form with the sum of squares, in which case the isomorphism classes of lattices may be identified with the set of discrete subgroups of $\mathbb{R}^n$ of rank $n$, modulo rotations. Classically, both of these complementary identifications have been used.

Even if one wants to work only with unit lattices in dimension $n$ it is necessary to work with arbitrary lattices of smaller rank. In terms of the group $\mathrm{SL}_n$ this amounts to using the copies of $\mathrm{GL}_m$ embedded along the diagonals of $n \times n$ matrices. For this reason I generally deal with all lattices.

**3. Dimension two.**    We shall first look more carefully at the case $n = 2$, where things can be easily understood. In identifying the isomorphism classes of unit lattices with $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ a certain number of coincidences play an important role and it is probably best if I recall them.

First of all, any pair $u$ and $v$ in $\mathbb{C}$ which are not real multiples of one another determine a lattice. I'll choose for this the opposite of the usual orientation in $\mathbb{C}$. In particular a pair $z = x + iy$ with $y > 0$ and 1 determine a lattice.

Suppose let $u_*$ and $v_*$ be a basis of an arbitrary two-dimensional lattice. Let $v$ be a complex number with $|v| = \|v_*\|$ and $u$ such that $|u| = \|u_*\|$, $\mathrm{IM}(u) > 0$, with the angle between $u$ and $v$ equal to that between $u_*$ and $v_*$. Then the pair $z = u/v$ and 1 are similar to the pair $u_*$ and $v_*$. Thus

• *The upper half-plane $\mathcal{H}$ classifies similarity classes of two-dimensional lattices.*

It also classifies bases of unit lattices, since there is a unique unit lattice in every oriented similarity class. The lattice spanned by $z$ and 1 has area $y$, so that it corresponds to the unit lattice spanned by $z/\sqrt{y}$ and $1/\sqrt{y}$.

Every point in $\mathcal{H}$ can be transformed by an element of $\mathrm{SL}_2(\mathbb{Z})$ into an essentially unique point in the region

$$\mathfrak{D} = \left\{ z = x + iy \mid -1/2 < x \leq 1/2, \ |z| \geq 1 \right\}.$$
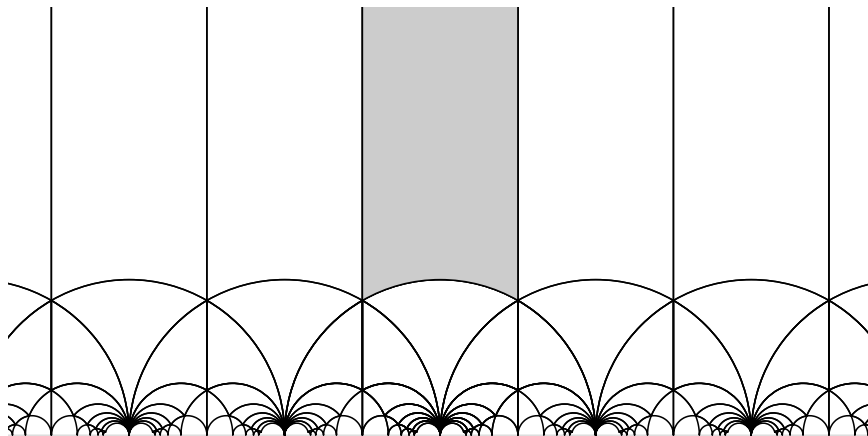
Equivalently, the lattice spanned by 1 and $z$ will be similar to an essentially unique one spanned by 1 and a point of this region.

I think it was Lagrange who first described the algorithm that carries out the necessary reduction, although for a slightly different purpose. The basic reasoning behind the algorithm is contained in this very elementary result, to which I give Lagrange's name for subsequent reference:

LEMMA 3.1. (Lagrange) *If $L$ is any lattice, $u$ a primitive vector in $L$, and $v'$ a vector in $L' = L/\mathbb{Z}u$, then there exists a unique representative $v$ of $v'$ in $L$ with the property that its projection onto $u$ lies in the interval $(-u/2, u/2]$. The inequality*

$$\|v\|^2 \leq \frac{\|u\|^2}{4} + \|v'\|^2 .$$

*holds, where we identify $v'$ with a vector $v^\perp$ in the orthogonal complement of $u$.*
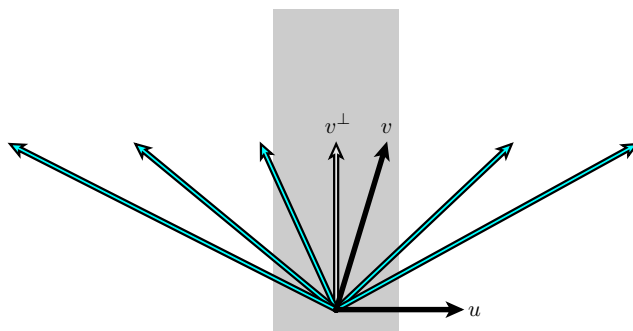


To see from this why every point $z$ in $\mathcal{H}$ may be transformed to an essentially unique point in the region $\mathfrak{D}$, take $u$ to be a vector of least length in the lattice generated by 1 and $z$, and apply the Lemma. The vector $v$ will then have length at least as large as that of $u$, and after we rotate and scale to get $u = 1$ the vector $v$ will lie in the region $\mathcal{D}$. For some points there may be several vectors of least length in this lattice (i.e. more than just one and its negative), and this will cause some ambiguity in the choice of point in $\mathfrak{D}$. What this amounts to is that the points $z$ and $-1/z$ in $\mathfrak{D}$ on the unit circle $|z| = 1$ will be associated to the same lattice in this procedure.

In summary, if $(L, Q)$ is a lattice of rank two there exists an essentially unique positively oriented basis $u$ and $v$ of $L$ where $u$ is a vector in $L$ of shortest length and the projection of $v$ on the line through $u$ lies between $\pm u/2$. For exceptional lattices corresponding to points on the boundary of $\mathfrak{D}$ there will be some harmless ambiguity in the choice of $u$ and $v$. Our knowledge of the domain $\mathfrak{D}$ allows us to classify completely the isomorphism classes of unit lattices. The main result of these notes will be to generalize this classical result in a somewhat weak sense.

Grayson (following Stuhler) associates to every lattice $L$ of rank two its **Newton polygon**. First we make up a set in the plane in the following way: (1) We put $(0, 0)$ in it. (2) Let $L)$ be the common area of any one of the fundamental parallelograms of $L$. We put $(2, \log L))$ in the set. (3) If $v$ is any **primitive** vector in $L$ (i.e. not a multiple of one in $L$) then we put $(1, \log \|v\|)$ in the set. The first coordinate in each of these points is just a dimension. We plot these points in the plane. For example, when the lattice is this:
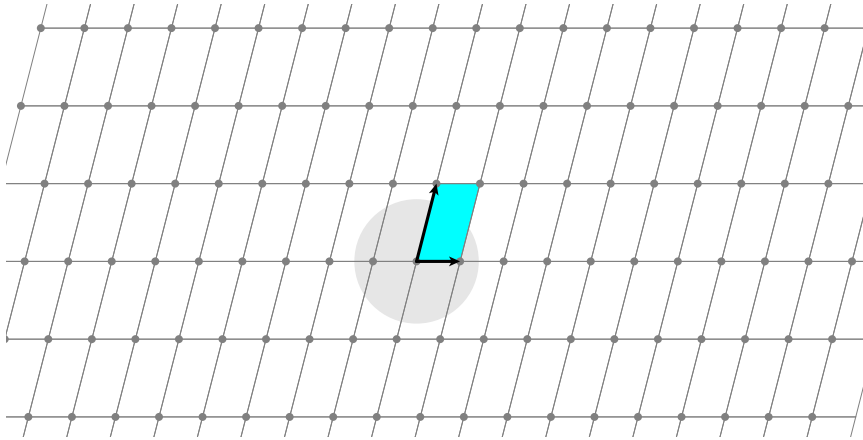
(with $\|u\| < 1$) we get the plot on the left, and if we shrink-wrap it—for reasons I'll explain in a moment—we get the figure on the right:
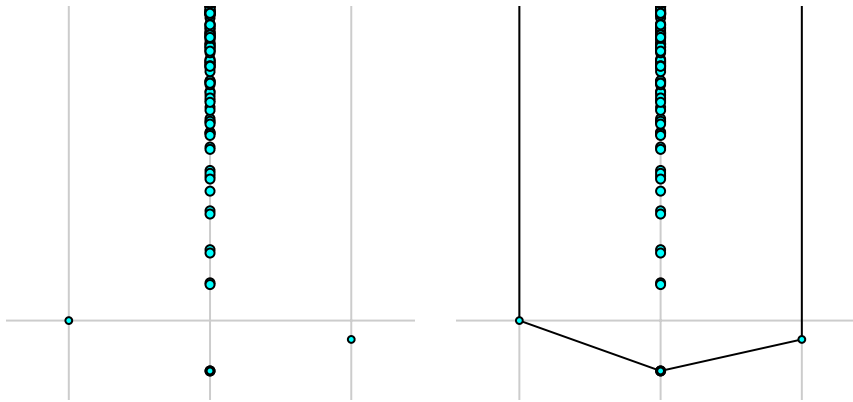


The shorter a vector, the lower its plotted point. Since the vectors in every lattice have length bounded away from 0, the plot points are certainly bounded from below. Therefore the convex hull of the collection of plot points is a polygon bounded from below. Since there are arbitrarily long primitive vectors in the lattice the left and right sides of the hull are vertical lines. Grayson calls the set of points plotted the **canonical plot** of the lattice, and the boundary of the convex hull of the plot its **canonical polygon**. I'll call it the lattice's **profile**.

Let $z = x + iy$ be a point of $\mathfrak{D}$, and let $a = \sqrt{y}$. The lattice of unit area corresponding to $z$ is that spanned by $1/a$ and $(x/a) + ia$. The vector $1/a$ is a vector of least length in this lattice, by definition of $\mathfrak{D}$. The point Grayson attaches to the lattice is thus $(1, -\log a)$. This will lie below the $x$ axis when $a > 1$. Therefore the points of the interesting part of $\mathcal{D}$ where $y \leq 1$ correspond to canonical plots lying entirely on or above the $x$-axis, and the profile of such a lattice has its only vertices at $(0,0)$ and $(2,0)$. It is called a **semi-stable** lattice by Grayson and Stuhler, and if we don't assume the lattice to have area $A = 1$ then a lattice is called semi-stable if the bottom of its profile is a straight line. If $u$ is the shortest vector in the lattice then semi-stability means that $\log \|u\| \geq (1/2) \log A$ or $\|u\| \geq \sqrt{A}$. The terminology is taken from Mumford's geometrical invariant theory—stable lattices are the arithmetic analogues of stable vector bundles on Riemann surfaces, which are

discussed for example in the paper [Harder-Narasimhan 1975].

Points in $\mathfrak{D}$ where $y > 1$ correspond to plots falling below the $x$-axis, and the profile will have an additional vertex below the $x$-axis on the line $x = 1$. More generally, a lattice which is not semi-stable is one in which $\|u\| < \sqrt{A}$. It is said to be **unstable**. Thus *the degree of instability of a rank two lattice is measured by the size of its smallest vectors, compared to its volume.* One important property that unstable lattices possess is that for them the line containing a shortest vector, the one giving rise to the middle vertex, is unique.

Something stronger is true, however, for unstable lattices—something that can be noticed in the figures above. If $u$ is a shortest vector in an unstable lattice then $\ell(u) = (1, \log \|u\|)$ is a vertex on the profile. The lattice's profile will break at this point. *The nature of the break tells something about the second shortest primitive vectors in the lattice.* Let $v$ be a primitive vector such that $u$ and $v$ span the lattice. The area $A$ is equal to $\|u\| \cdot \|v^\perp\|$, where $v^\perp$ is the projection of $v$ orthogonal to $u$. The slope of the profile to $\ell(u)$ is $\log \|u\|$, and that from $\ell(u)$ to $(2, \log A)$ is $\log \|v^\perp\| \leq \|v\|$. The existence of the break for $u$ means that the second slope is greater than the first. Furthermore, any other primitive vector in the lattice will project onto a multiple of $v^\perp$. Therefore the inside of the parallelogram shown in the following picture is empty of plotted points:
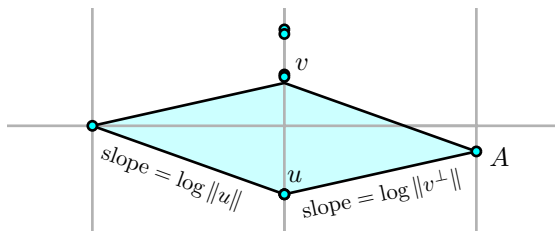


where we have matched the bottom of the canonical polygon with matching sides of a parallelogram. This explains the apparent gap towards the bottom of the canonical plot.

Let $L$ be an unstable lattice with shortest vector $u$, let $V_1$ be the rational line through $u$, and $L_1 = V_1 \cap L$. This determines a **lattice flag** $F$

$$0 \subset L_1 \subset L_2 = L$$

called the **canonical flag** associated here to $L$. This gives rise in turn to a flag of rational subspaces

$$0 \subset V_1 = L_1 \otimes \mathbb{R} \subset V_2 = L \otimes \mathbb{R} .$$

Conversely, if $\mathcal{F}$ is any rational flag in $\mathbb{R}^2$, let $\mathcal{H}_\mathcal{F}$ be the set of all unstable lattices with flag $\mathcal{F}$. It follows from the remarks just above that $\mathcal{H}_F$ is invariant under the unipotent radical $N_\mathcal{F}$ of the parabolic subgroup $P_\mathcal{F}$ stabilizing $\mathcal{F}$. More explicitly, if $\mathcal{F}_\infty$ is the flag fixed by the subgroup $P$ of upper triangular matrices then $\mathcal{H}_{\mathcal{F}_\infty}$ is the region $\{y > 1\}$. If $\gamma$ lies in $\mathrm{SL}_2(\mathbb{Z})$ and $\gamma(\infty) = p/q$ then $\mathcal{H}_{\gamma \mathcal{F}_\infty}$ is the $\gamma \mathcal{H}_{F_\infty}$, the interior of the circle tangent to $\mathbb{R}$ at $p/q$ of radius $1/2q^2$.
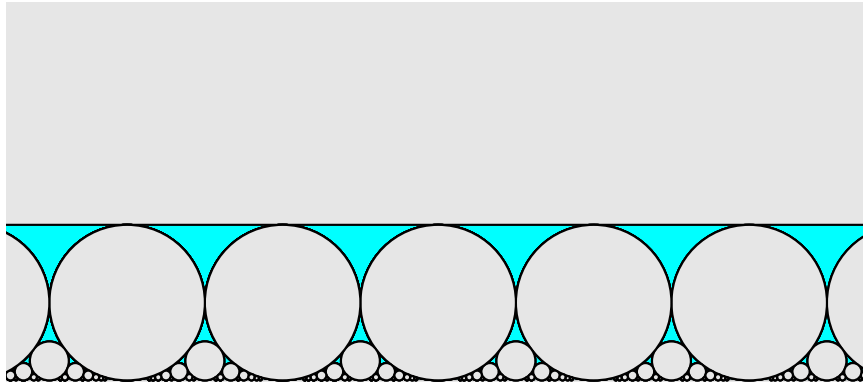
The distinction between stable and unstable partitions the fundamental domain $\mathcal{D}$.
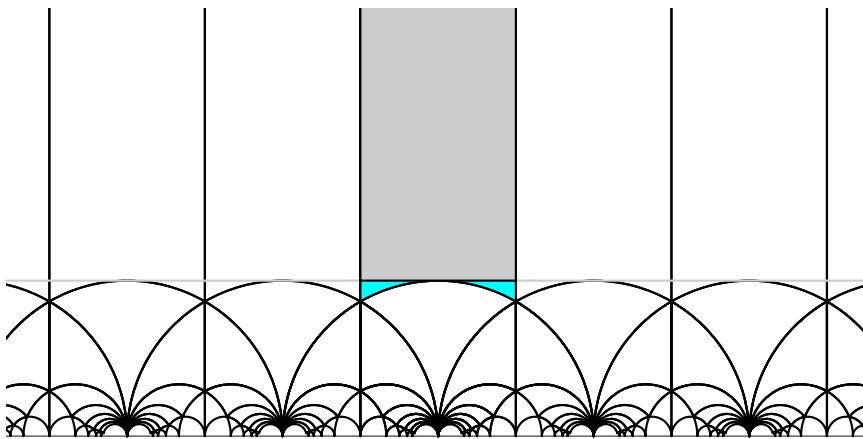


What is the significance of this partition? The group $\Gamma \cap P$ is made up of matrices

$$\pm \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

with $n$ an integer, and elements of $\Gamma \cap P$ act by horizontal integral translation on $\mathcal{H}$. The group $\Gamma \cap P$ is far simpler than $\Gamma$ itself. A fundamental domain for $\Gamma \cap P$ is the band

$$\{x + iy \mid -1/2 < x \leq 1/2\}$$

The region $y > 1$ in $\mathfrak{D}$ may therefore be identified with a very simple subregion of $\Gamma \cap P \backslash \mathcal{H}$. Its structure doesn't mirror any of the complexity of $\mathcal{D}$ itself. The region $y \leq 1$ in $\mathfrak{D}$, on the other hand, is rather more complicated. I call it the **core** of $\mathfrak{D}$. What the partition of Arthur does for $n > 2$ is to divide up similarly the space $\mathrm{GL}_n(\mathbb{Z}) \backslash X_n$, partitioning isomorphism classes of lattices of dimension $n$ into components associated to **parabolic subgroups** of $\mathrm{GL}_n$ (or certain conjugacy classes of them). The component corresponding to the group $P$ may be identified with a subset of $\Gamma \cap P \backslash X_n$ describable in terms of the geometry of $P$ rather than that of $G$.

**4. Lattices of arbitrary rank.** Fundamental domains for the action of $\mathrm{GL}_n(\mathbb{Z})$ on $X_n$ have been completely described for a few low values of $n$. The details

are useful in certain computations, but since their complexity grows rapidly with $n$ it is fortunate that explicit knowledge of this sort is rarely necessary in the theory of automorphic forms. For large $n$, then, each component in Arthur's partition will possess a core of a perhaps unknown (and even unknowable) nature, but the exact description of that core should not be required to elicit interesting and important information. In fact the opposite is in some sense true—analytical techniques should be able to say something about the geometry of the core of an arithmetic quotient that is almost impossible to access directly.

The simplest way to construct the partition uses the **canonical flag** of a lattice of arbitrary dimension. This concept originated perhaps with Gunter Harder, was extended by Ulrich Stuhler, and improved by Dan Grayson. Grayson's ideas might be said merely to add graphic content to those of Stuhler, but the effect on the clarity of arguments is dramatic. He associates to every lattice its canonical plot, its profile, and then finally its canonical flag.

If $L$ is a lattice and $M$ is a discrete subgroup, $M$ is called a **sublattice** if one of these equivalent conditions holds:

(1) *$L/M$ has no torsion*;
(2) *$M$ is a summand of $L$*;
(3) *every basis of $M$ may be extended to a basis of $L$*;
(4) *the group $M$ is the intersection of $L$ with a rational vector subspace of $L_\mathbb{R}$*;
(5) *the quotient $L/M$ is a free $\mathbb{Z}$-module.*

The sublattices of dimension one, for example, are the free subgroups spanned by a single primitive vector, one which is not a multiple of another lattice vector. If $M$ is a sublattice then the vector space $M_\mathbb{R}$ inherits a metric from $L_\mathbb{R}$, so from every sublattice, as indeed from every discrete subgroup, one obtains again a lattice of generally lower rank.

The **volume** of a lattice $L$ is that of the compact torus $L_\mathbb{R}/L$, or equivalently the $n$-dimensional volume of the parallelopiped spanned by any basis of $L$. Suppose $L$ to have rank $n$. If $(\ell_i)$ is a basis of $L$ and $(e_j)$ an orthonormal basis of $L_\mathbb{R}$ and then the volume of $L$ is the absolute value of the determinant of the square matrix $[\langle \ell_i, e_j \rangle]$ whose $i$-th column is made up of the coordinates of $\ell_i$ with respect to the basis $(e_j)$. If $M$ is a sublattice of rank $m$ in $L$ with basis $\ell_1, \ldots, \ell_m$ then the volume of $M$ is the length of the vector $\ell_1 \wedge \ldots \wedge \ell_m$ in $\bigwedge^m L$, or in other words the square root of the sum of the squares of the determinants of the $m \times m$ minor matrices in the $n \times m$ matrix whose columns are the coordinates of the $\ell_i$ with respect to any orthonormal basis of $L_\mathbb{R}$.

If $M$ is a lattice, let $M)$ be the volume of the quotient $M_\mathbb{R}/M$, and let $\dim(M)$ be its rank. We associate to $M \subseteq L$ the point

$$\ell(M) = (\dim(M), \log M))$$

in $\mathbb{R}^2$, and define (following Grayson and Stuhler again) the **canonical plot** of the lattice $L$ to be the set of all points $\ell(M)$ as $M$ ranges over all its sublattices. The origin all by itself is considered to be a lattice of dimension 0 and, by convention, volume 1. It therefore corresponds to the plotted point $(0,0)$. If $M$ has rank one then its volume is the length of a generator. Since the lengths of vectors in a lattice are bounded below so are the plots $(1, \log M))$ as $M$ ranges over all rank one sublattices.

For $M$ of rank $m$ the volume of $M$ is the same as the volume of the rank one sublattice lattice $\bigwedge^m M$ in $\bigwedge^m L$, and again the point $(m, \log M))$ must be bounded from below by a constant depending only on $L$. Define the **profile** of $L$ to be the polygonal boundary of the convex hull of its canonical plot. The plot of a lattice is just about impossible to compute in any sense, but its profile can be computed (in principle) by finding the shortest vectors in each of its exterior products. In practice, this is an infeasible computation for large dimension.

Since there exist arbitrarily long primitive vectors in $L$ and more generally lattices of any rank smaller than $n = \dim(L)$ of arbitrarily large volume, we may as well add to the profile the points $(0, \infty)$ and $(n, \infty)$. The sides of the profile are therefore vertical. Its bottom is a convex polygonal line from $(0, 0)$ to $(n, \log L))$ if $n$ is the rank of $L$.

The profile will contain inside it at least the convex hull of the four points $(0, \infty)$, $(0, 0)$, $(n, \log L))$, $(n, \infty)$, and it may happen that this is all of it. When this is the case, $L$ is said to be **semi-stable**. When this is not the case, the profile of $\Lambda$ will lie strictly below the straight line from $(0, 0)$ to $(\dim(L), \log L))$.

Here, for example, is the plot we get from the three-dimensional lattice with basis $(1, 1, 2)$, $(2, 0, -3)$, $(2, 1, 5)$ scaled suitably to obtain a unit lattice:
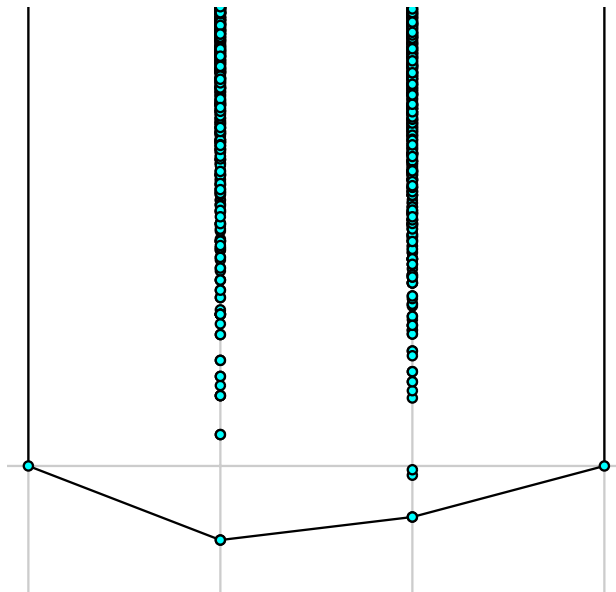


As in the earlier two-dimensional plot, the gaps at the bottom are significant, as we shall see in Proposition 4.3.

If $M$ is a sublattice of $L$, then the projection from $L_\mathbb{R}/M_\mathbb{R}$ onto the orthogonal complement of $M_\mathbb{R}$ in $L_\mathbb{R}$ is an isomorphism, and in this way the quotient space inherits a metric from that on $L_\mathbb{R}$. The quotient group $L/M$ in the quotient space together with this metric defines therefore a lattice, the **quotient lattice**.

Suppose $M$ to be a sublattice of $L$ with basis $(m_i)$, $(e_i)$ to be an orthonormal basis of $M_\mathbb{R}$, $(n_j)$ to be a complement to $M$ in a basis of $L$. Suppose also that the $f_j$ extend the $e_i$ to an orthonormal basis of $L$. Then $\langle m, f_j \rangle = 0$ for $m$ in $M$, and the

volume of $L$ is

$$L) = \left| \det \begin{bmatrix} \langle m, e \rangle & \langle n, e \rangle \\ \langle m, f \rangle & \langle n, f \rangle \end{bmatrix} \right|$$

$$= \left| \det \begin{bmatrix} \langle m, e \rangle & \langle n, e \rangle \\ 0 & \langle n, f \rangle \end{bmatrix} \right|$$

$$= |\det [\, \langle m, e \rangle \,]| \; |\det [\, \langle n, f \rangle \,]| \; .$$

The columns of the matrix $[\, \langle n, f \rangle \,]$ are the coefficients of the projections of the $n_j$ onto the orthogonal complement of $M$, and its determinant is therefore the volume of the quotient lattice $L/M$. It donates one term among several non-negative terms to the volume of the sublattice of $L$ spanned by the $\ell_j$. All in all, as a generalization of the formula $A = b \cdot h$ for the area of a parallelogram spanned by a two-dimensional lattice:

PROPOSITION 4.1. *If $M$ is any sublattice of $L$ then*

$$L) = M)\, L/M)$$

*and if $N$ is any sublattice of $L$ complementary to $M$ then*

$$N) \geq L/M) \; .$$

The second assertion generalizes the simple fact that the length of the orthogonal projection of a vector cannot be larger than the length of the vector. It reduces to that result, in fact, if one considers exterior powers of $L$. As first pointed out by Stuhler, it has a simple useful generalization, when it is applied to the lattices $M/M \cap M_*$ and $M_*/M \cap M_*$ in $M + M_*/M \cap M_*$:

COROLLARY 4.2. *If $M$ and $M_*$ are any two sublattices of $L$ then*

$$\frac{M + M_*)}{M)} \leq \frac{M_*)}{M \cap M_*)}$$

*or equivalently*

$$M + M_*)\, M \cap M_*) \leq M)\, M_*)$$

This result is expressed by Grayson in additive terms:

PROPOSITION 4.3. (Grayson's parallelogram rule) *Suppose that $M$ and $M_*$ are sublattices of $L$. Then*

$$\log M) \geq \log M + M_*) + \log M \cap M_*) - \log M_*) \; .$$

Why is it called the parallelogram rule? We have a short chain of lattices

$$M_* \cap M \subseteq M_* \subseteq M_* + M \; .$$

Let

$$d = \dim(M + M_*) + \dim(M \cap M_*) - \dim(M_*)$$
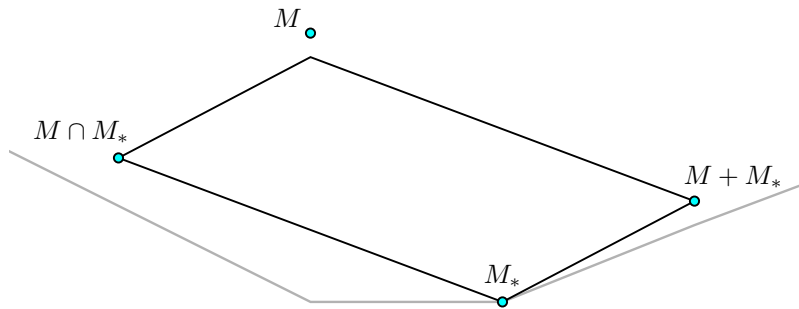$$\ell = \log M + M_*) + \log M \cap M_*) - \log M_*) \ .$$

Then $d$ is the dimension of $M$. The Stuhler-Grayson inequality says neither more nor less than that the point $\ell(M) = (\dim M, \log M))$ lies on or above the point $(d, \ell)$. It will lie exactly at $(d, \ell)$, furthermore, if and only if $M_*$ and $M$ project to orthogonal lattices in $(M + M_*)/(M_* \cap M)$. But the point $(d, \ell)$ is the fourth corner of a parallelogram whose other vertices are $\ell(M \cap M_*)$, $\ell(M_*)$, and $\ell(M + M_*)$. The useful situation is that illustrated below:
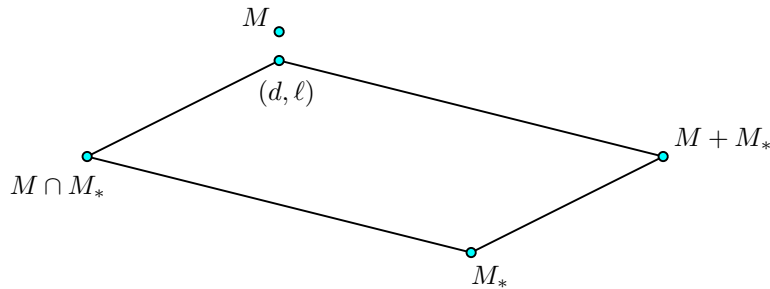


The **vertices** of a profile are its extremal points, where it actually bends. The points $(0, 0)$ and $\ell(L) = (n, \log L))$ are certainly vertices. The first of two main results in this theory concerns other possibilities.

LEMMA 4.4.  *Suppose $M_*$ to be a lattice with $\ell(M_*)$ a vertex on the profile. Whenever $M$ is any other lattice with $\ell(M)$ on the profile we must have either $M \subseteq M_*$ or $M_* \subseteq M$.*

*Proof.* Start off by letting $M$ be arbitrary, $M_*$ a vertex of the profile.



Then $M \cap M_*$ will lie somewhere to the (inclusive) left of both, and $M + M_*$ will lie somewhere to the (inclusive) right of both. The parallelogram whose bottom boundary is $M \cap M_*$, $M$, $M + M_*$ will, by the parallelogram rule, lie underneath $M$. Unless it is one-dimensional, $M$ will be separated from the profile. Therefore if $M$ lies on the profile, the parallelogram must be degenerate, and this means that either $M \cap M_* = M_*$ and $M_* \subseteq M$, or $M + M_* = M_*$, in which case $M \subseteq M_*$.  □

As a consequence:

THEOREM.  *(a) The sublattices of $L$ giving rise to the vertices of the profile of $L$*

*are unique. (b) Any set of sublattices corresponding to extremal points of the profile form a flag.*

This flag is called by Grayson the **canonical filtration** of $L$. I call it the **canonical flag**. A lattice is semi-stable if and only if its canonical flag is trivial.

If $M \subseteq L$ is the sublattice corresponding to a vertex $(i, \ell)$ of the profile of $L$, then that part of the polygon of $L$ running from $x = 0$ to $x = i$ is the profile of $M$, and that part running from $x = i$ to $x = n$ is a translation of that of $L/M$. If $N$ occurs in the canonical flag of $L$ and contains $M$ then $N/M$ occurs in the canonical flag of $L/M$.

Here is another corollary of the Lemma.

THEOREM. (Grayson's criterion) *Suppose*

$$L_0 = \{0\} \subset L_1 \subset L_2 \subset \ldots \subset L_k = L$$

*to be a flag with the property that each quotient $L_i/L_{i-1}$ is semi-stable, and such that the slope of $(L_{i-1}, L_i)$ is less than the slope of $(L_i, L_{i+1})$. Then this flag is the canonical flag.*

*Proof.* Suppose $M$ to be any other sublattice of $L$. We want to know that $\ell(M)$ lies above the plot $P$ of the $\ell(L_i)$. We prove by induction that if $M \subseteq L_i$ then this is so. For $i = 1$ this is immediate.
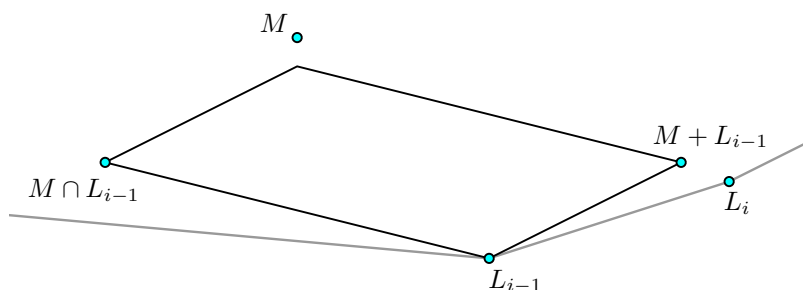


Suppose that $M \subseteq L_i$ with $i > 1$. Then $M + L_{i-1}$ is contained in $L_i$ and contains $L_{i-1}$, hence its plot lies on or above the segment $(L_{i-1}, L_i)$. By induction, the plot of the intersection $M \cap L_{i-1}$ also lies on or above $P$. The parallelogram rule thus implies that the plot of $M$ also lies on or above $P$. □

An isomorphism of two lattices takes the canonical flag of one into that of the other. The canonical plot and profile of a lattice are therefore invariants of the isomorphism class of a lattice, as is the $GL(L)$-conjugacy class of the canonical flag.

In general, if $p$ is a function defined on the integer interval $[0, n]$, I'll call its profile the polygon that starts at $(0, \infty)$, then follows segments $(i, p(i))$ in increasing order of $i$, and finally goes up to $(n, \infty)$. The convex ones among these are the profiles of lattices. The polygons obtained in this way I'll call **profile polygons**.

**5. The geometry of acute cones.** This section is largely a self-contained account of a simple geometrical construction first applied in this subject in [Langlands:1989]. The new feature here is the connection with Grayson's diagrams.

Suppose $\Delta$ to be a set of linearly independent vectors in a Euclidean space $V$. Let $P$ be a basis of $V$. I define a **weight map** to be a map $\alpha \mapsto \varpi_\alpha$ from $\Delta$ to $P$ satisfying the condition that for all $\varpi$ in $P$

$$\varpi \bullet \alpha = \begin{cases} 1 & \text{if } \varpi = \varpi_\alpha \\ 0 & \text{otherwise.} \end{cases}$$

Fix $V$, $\Delta$, $P$, and a weight map for the rest of this section. The subspace of $V$ perpendicular to $\Delta$ is complementary to the subspace spanned by the image of the weight map; let $P$ be a basis of $V$ extending that image and containing a basis of that complement.

For each $\Theta \subseteq \Delta$ let $\Theta^\perp$ be the subset of the $\varpi$ in $P$ that are perpendicular to the $\alpha$ in $\Theta$, and let

$$\nu_\Theta = \text{ orthogonal projection onto the subspace spanned by } \Theta\,.$$

The map $\nu_\Theta$ will be referred to as **normalization**. The set $\Theta^\perp$ is also the complement of the $\varpi_\alpha$ for $\alpha$ in $\Theta$. The vectors $\varpi_\alpha$ are by no means unique—any of them may be translated by a vector orthogonal to all the $\alpha$ in $\Delta$. They may be made unique by imposing the condition that the $\varpi_\alpha$ all lie in the subspace spanned by the $\alpha$. I'll not impose this condition on a weight map, because then we would lose the very useful feature feature that

*The restriction of a weight map to a subset of $\Delta$ is still a weight map.*

I'll fix $V$, $\Delta$, and a weight map for the rest of this section.

Let $\mathcal{C} = \mathcal{C}^\Delta$ be the open cone dual to the $\alpha$ in $\Delta$—the $v$ in $V$ such that $\alpha \bullet v > 0$ for all $\alpha$. The vector $\sum_{\varpi \in P} c_\varpi \varpi$ lies in $\mathcal{C}$ if and only if $c_\varpi > 0$ whenever $\varpi = \varpi_\alpha$ for some $\alpha$ in $\Delta$. The cone $\mathcal{C}$ is invariant under translation by elements of $\Delta^\perp$.
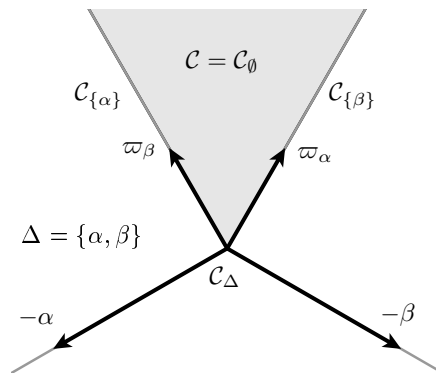


The faces of $\mathcal{C}$ are parametrized by subsets of $\Delta$—to $\Theta$ corresponds the face $\mathcal{C}_\Theta$ of $v$ such that $\alpha \bullet v = 0$ for $\alpha$ in $\Theta$ and $\alpha \bullet v > 0$ for $\alpha$ not in $\Theta$. In addition, let $V_\Theta$ be the linear subspace spanned by $\mathcal{C}_\Theta$, that of all $v$ with $\alpha \bullet v = 0$ for $\alpha$ in $\Theta$. Thus $\mathcal{C}_\emptyset$ is $\mathcal{C}$ itself and $\mathcal{C}_\Delta = V_\Delta$, the face of lowest dimension, the linear space spanned by the $\varpi$ in $\Delta^\perp$. Let

$$\pi_\Theta = \text{ orthogonal projection onto } V_\Theta\,.$$

Thus $\pi_\Theta$ and $\nu_\Theta$ are complementary, in the sense that they add up to the identity operator and their images are orthogonal.

The cone $\mathcal{C}$ also determines a partition of the whole space $V$—to each of its open faces $F$ associate the set $V_F^{\mathcal{C}}$ of points $p$ for which the point of $\overline{\mathcal{C}}$ closest to $p$ lies on $F$. This partition is shown above in dimension 2. In general, $V_{\mathcal{C}}^{\mathcal{C}}$ is just $\mathcal{C}$ itself. I write $V_\Theta^\Delta$ for $V_{\mathcal{C}_\Theta}^{\mathcal{C}}$.

In summary:

$$\mathcal{C}^\Delta = \{v \in V \,|\, \alpha \bullet v > 0 \text{ for all } \alpha \in \Delta\}$$

$$\mathcal{C}_\Theta^\Delta = \{v \in \overline{\mathcal{C}}^\Delta \,|\, \alpha \bullet v = 0 \text{ for all } \alpha \in \Theta\}$$

$$V_\Theta = \{v \in V \,|\, \alpha \bullet v = 0 \text{ for all } \alpha \in \Theta\}$$

$$V_\Theta^\Delta = \{v \in V \,|\, \mathcal{C}_\Theta^\Delta \text{ contains the nearest point to } v \text{ in } \overline{\mathcal{C}}^\Delta \, \}$$

LEMMA 5.1. *Suppose $v$ to be a point of $V$ not in $\overline{\mathcal{C}}$, $\overline{v}$ to be a point of $\overline{\mathcal{C}}$, and $H$ the hyperplane containing $\overline{v}$ and perpendicular to $v - \overline{v}$. Then $\overline{v}$ is the nearest point in $\overline{\mathcal{C}}$ to $v$ if and only if all points of $\mathcal{C}$ lie on the side of $H$ opposite to $v$.*
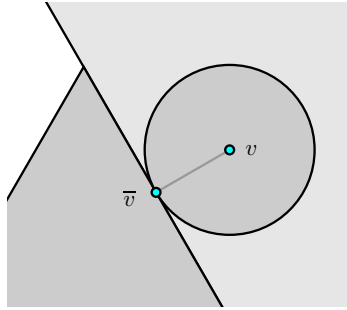


This is because of the convexity of the sphere centred at $v$ and passing through $\overline{v}$.

PROPOSITION 5.2. *The points in $V_\Theta^\Delta$ are those of the form*

$$v = \sum_\Theta c_\alpha \alpha + \overline{v}$$

*where $\overline{v}$ lies in $\mathcal{C}_\Theta$ and each $c_\alpha \leq 0$. In particular*

$$V_\Delta^\Delta = \left\{ \sum_\Delta c_\alpha \alpha + \sum_{\Delta^\perp} c_\varpi \varpi \,\middle|\, \text{ all } c_\alpha \leq 0 \right\},$$

*or, equivalently, it is the inverse image under orthogonal projection of the closed cone spanned by the $-\Delta$.*

*Proof.* Suppose $v$ in $V$ but not in $\overline{\mathcal{C}}$. Then according to the Lemma $v$ lies in $V_\Theta^\Delta$ with nearest point $\overline{v}$ if and only if the hyperplane $H$ perpendicular to $v - \overline{v}$ contains $v$ on one side and $\mathcal{C}$ on the other. If $\overline{v}$ lies in $\mathcal{C}_\Theta$ then it is easy to see that $H$ must contain a neighbourhood of $\overline{v}$ in $\mathcal{C}_\Theta$, hence all of $\mathcal{C}_\Theta$. Therefore

$$v - \overline{v} = \sum_{\alpha \in \Theta} c_\alpha \, \alpha \,,$$

and since the $\varpi$ in $\Theta$ are on the other side from $v$, all $c_\alpha \leq 0$. The converse is also straightforward.   ▯

Because the vectors in $\Theta$ are orthogonal to the face $\mathcal{C}_\Theta$, this says that the set $V_\Theta^\Delta$ is the product of two sets, and these turn out to be rather easy to describe. The set $V_\Theta^\Delta$ is the intersection of $\pi_\Theta^{-1}(\mathcal{C}_\Theta)$ and $V_\Theta^\Theta$, and $V_\Theta^\Theta$ is itself the inverse image under $\nu_\Theta$ of the closed cone spanned by $-\Theta$.

Now I take up the class of examples that we'll be interested in later on. Let $E = \mathbb{R}^n$ with orthogonal basis $\varepsilon_i$ and coordinates $s_i$. For $1 \leq i \leq n-1$ let $\alpha_i = \varepsilon_{i+1} - \varepsilon_i$, so that

$$\alpha_i \bullet \left( \sum s_j \varepsilon_j \right) = s_{i+1} - s_i \ .$$

Then let $\Delta = \{\alpha_i \,|\, 1 \leq i \leq n-1\}$. The subspace spanned by the $\alpha_i$ is that where the sum of coordinates vanishes. If for $i \leq n$

$$\varpi_i = -\varepsilon_1 - \cdots - \varepsilon_i$$

then

$$\alpha_i \bullet \varpi_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

so that the space orthogonal to $\alpha_i$ is spanned by the $\varpi_j$ with $i \neq j$. Since $\alpha_i \bullet \alpha_j = -1$ if $|i - j| = 1$ and otherwise vanishes, the cone spanned the $\alpha_i$ is obtuse. That spanned by the $\varpi_i$ is acute. The projection of $\varpi_i$ onto the space spanned by the $\alpha_i$ is $\nu_\Delta(\varpi_i) = \varpi_i - (i/n)\varpi_n$, since $\varpi_i \bullet \varpi_n = i$. In these circumstances the cone $\mathcal{C}^\Delta$

$$\mathcal{C} = \left\{ \sum s_i \varepsilon_i \,\Big|\, s_1 < \ldots < s_n \right\} \ .$$

If $(s_i)$ is a point of $\mathbb{R}^n$, I define its **profile** to be the profile polygon that moves from $x = i$ to $x = i + 1$ along a segment of slope $s_i$. If it passes through the points $(i, y_i)$ then we must have

$$y_0 = 0, \ y_i - y_{i-1} = s_i \text{ or } y_i = y_{i-1} + s_i$$

so that

$$y_i = s_1 + \cdots + s_i = -\varpi_i \bullet \left( \sum s_j \varepsilon_j \right) \ .$$

PROPOSITION 5.3. *The map taking a point of $\mathbb{R}^n$ to its profile is a bijection of $\mathbb{R}^n$ with the set of profile polygons. A point of $\mathbb{R}^n$ lies in $\overline{\mathcal{C}}$ if and only if its profile is convex.*

The last is true because slopes $s_i$ of a profile are non-decreasing if and only if it is convex. Going backwards, given a profile polygon I define its **slope** to be the point $(s_i)$ of $\mathbb{R}^n$ whose profile it is.

The cone $V_\Delta^\Delta$ spanned by $\varpi_n$ and the $-\alpha_i$ is that of all $v$ such that $\nu_\Delta(\varpi_i) \bullet v \leq 0$ for $i \leq n-1$, or equivalently where $y_i - (i/n)y_n \geq 0$ for $1 \leq i < n$. Since $(i/n)y_n$ is the linearly interpolated $y$-value at $x = i$ on the line from $(0,0)$ to $(n, y_n)$, this implies:

PROPOSITION 5.4. *A profile $(y_i)$ corresponds to a point of $V_\Delta^\Delta$ if and only if it lies completely above the straight line from $(0,0)$ to $(n, y_n)$.*

In this case, its the bottom of its convex hull is just a straight line segment.

For each $\ell \geq 0$ let $I_\ell = \{1, \ldots, \ell\}$ and for each $I \subseteq I_{n-1}$ let $\Theta_I$ be the set of $\alpha_i$ with $i$ in $I$. For each such $I$, the set $\Theta_I^\perp$ is the set of $\varpi_j$ with $j$ in $I_n - I$.

PROPOSITION 5.5. *A point $(s_i)$ lies in the subspace spanned by $\Theta_I$ if and only if the coordinates $y_i$ vanish whenever $i$ is in $I_n - I$.*

Projection onto the linear subspace $V_\Theta$ (where $\alpha \bullet v = 0$ for $\alpha$ in $\Theta$) is very nicely described in terms of profiles.

PROPOSITION 5.6. *Let $\Theta = \Theta_I$ be a subset of $\Delta$. If $\Pi$ is the profile of a point $(s_i)$ in $\mathbb{R}^n$, the profile of the projection $\pi_\Theta(s)$ of $s$ onto $V_\Theta$ is the polygon obtained from $\Pi$ by skipping along in straight line segments among the vertices of $\Pi$ whose $x$-coordinate is not in $I$.*

In the following picture, $I = \{1, 2, 4\}$, so it skips from $x = 0$ to $x = 3$ and then to $x = 5$.



$\Theta$-projection
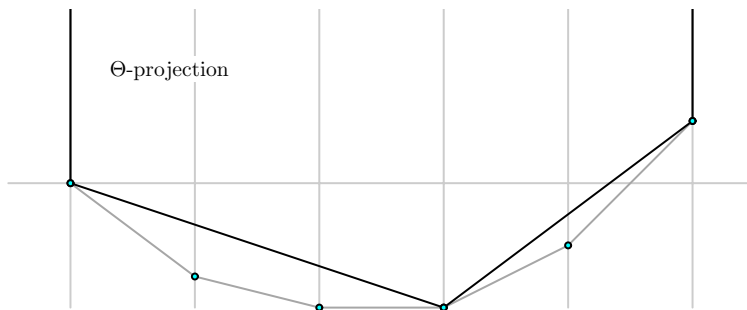
*Proof of the Proposition.* The second profile certainly satisfies the condition that $s_i = s_{i+1}$ for $i$ in $I$, which means that it lies in $V_\Theta$. The two profiles agree at the $i$ not in $I$, which means that their difference is orthogonal to the $\varpi_i$ with $i$ not in $I$. But this means in turn that the difference is a linear combination of the $\alpha$ in $\Theta$.

In other words, if the original profile is $(y_i)$ then the projected one $y_*$ has $y_{*,i} = y_i$ for $i$ not in $\Theta$, and for $i$ in between two successive integers $d_k$ and $d_{k+1}$ not in $I$ the values of $y$ are linearly interpolated:

$$y_{*,i} = y_{d_k} + (y_{d_{k+1}} - y_{d_k}) \left( \frac{i - d_k}{d_{k+1} - d_k} \right) .$$

I recall that a profile polygon is **normalized** by shearing it so as to place its final vertex on the $x$-axis. The vertical coordinate $y_i$ is replaced by $y_i^* = y_i - (i/n)y_n$. In terms of the slope, this is the same as $\nu_\Delta$.

Suppose $I$ to be a subset of $I_n$, with lacunae $d_k$. That is to say that $d_k$ and $d_{k+1}$ do not lie in $I$ but all the $i$ with $d_k < i < d_{k+1}$ do. Let $\Theta = \Theta_I$. The $\Theta$-normalization of a profile shears each of the segments in the range $[d_k, d_{k+1}]$ so as to normalize it— i.e. so as to place its endpoints on the $x$-axis. In a formula: for $d_k < i \leq d_{k+1}$ the new vertical coordinate becomes

$$y_{*,i} = (y_i - y_{d_k}) - (y_{d_{k+1}} - y_{d_k}) \left( \frac{i - d_k}{d_{k+1} - d_k} \right) .$$
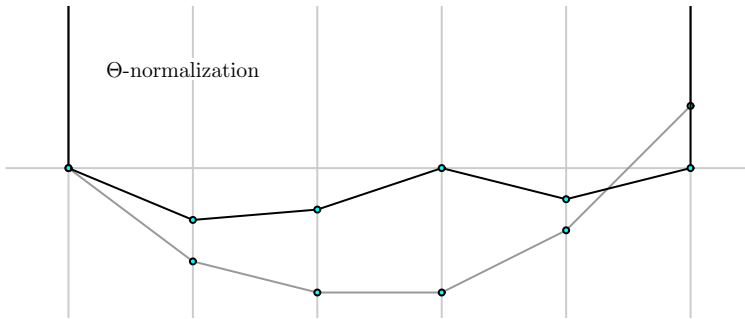
$\Theta$-normalization

PROPOSITION 5.7. *If $(s_i)$ is the slope of a profile, then its $\Theta$-normalization is its orthogonal projection onto the linear subspace perpendicular to $V_\Theta$.*

*Proof.* The formulas show that it is the complement of $\pi_\Theta$, the orthogonal projection onto $V_\Theta$.  $\square$

Here is a few figures, illustrating the comparison between profiles and slopes. First projection:
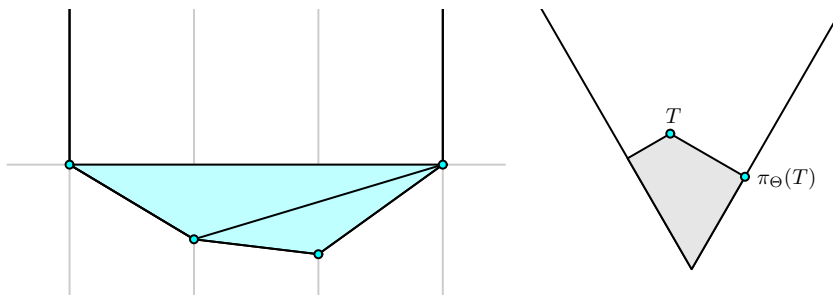


and then normalization:
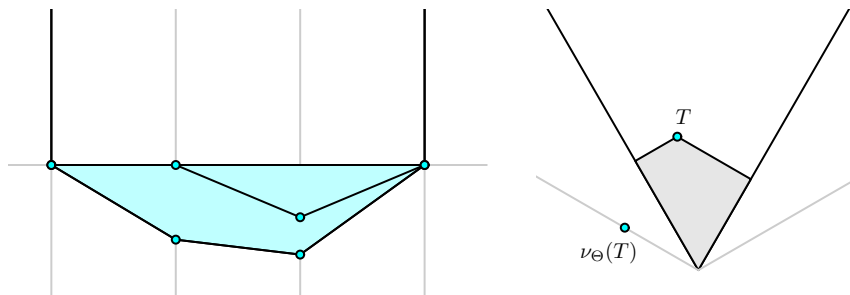


A partition of $V$ gives rise to other partitions by translating the original one. Grayson and Arthur describe partitions of $\overline{\mathcal{C}}$ of two different kinds, and each of these gives rise in turn, as we shall see, to a partition of $\Gamma\backslash X$. Grayson starts with the partition of $V$ by the signs of coordinates, and then shifts it by an element $T$ of $\mathcal{C}$ to give one of $\overline{\mathcal{C}}$:
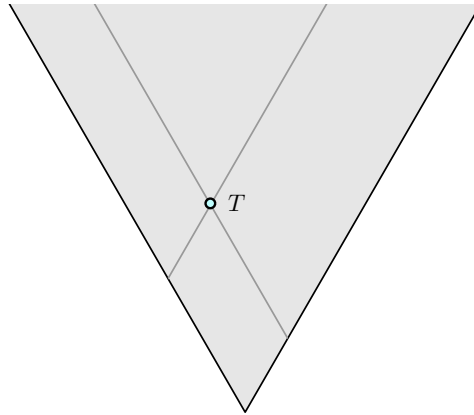
This was adequate for Grayson's purposes, but the partition used by Arthur fits more nicely into applications to automorphic forms. It just shifts the Langlands partition by an element $T$.
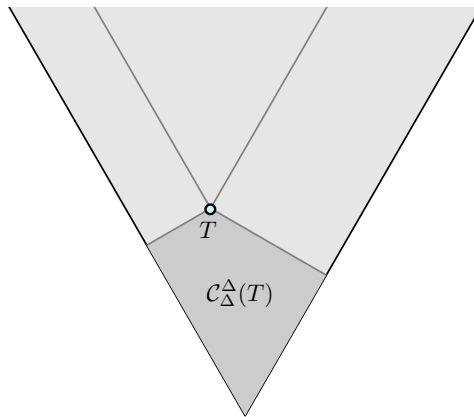


Let $\mathcal{C}_\Theta^\Delta(T)$ be the intersection of $\overline{\mathcal{C}}$ with the translation by $T$ of $V_\Theta^\Delta$.
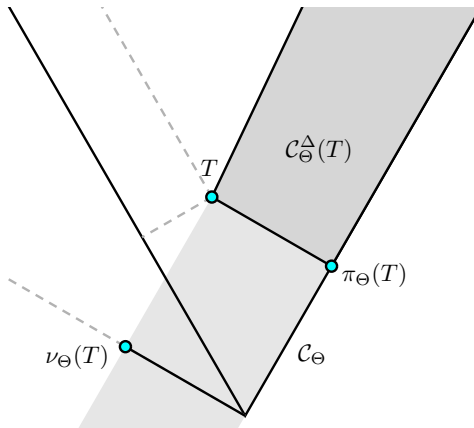


Like $V_\Theta^\Delta$ itself, it has a relatively simple product structure, one that has a useful description in terms of profiles. It is because of this product structure, and the con-

sequent structure induced on a corresponding subset of $X$, that Arthur's partition is more useful in automorphic forms.

First let's look at $\mathcal{C}_\Delta^\Delta$. I'll say that one point $T_*$ in $\mathbb{R}^n$ **dominates** another point $T$ if $T_* + V_\Delta^\Delta$ contains $T$. The points dominated by the origin, for example, are exactly those in $V_\Delta^\Delta$. What does this mean in terms of profiles? First of all, it is independent of normalization, since $V_\Delta^\Delta$ is invariant under translation by $\varpi_n$.

PROPOSITION 5.8. *If $T^*$ and $T$ are both points in the plane $\varpi_n \bullet v = 0$, the point $T_*$ dominates the point $T$ if and only if the profile $\Pi$ of $T$ lies entirely above the profile $\Pi_*$ of $T_*$.*
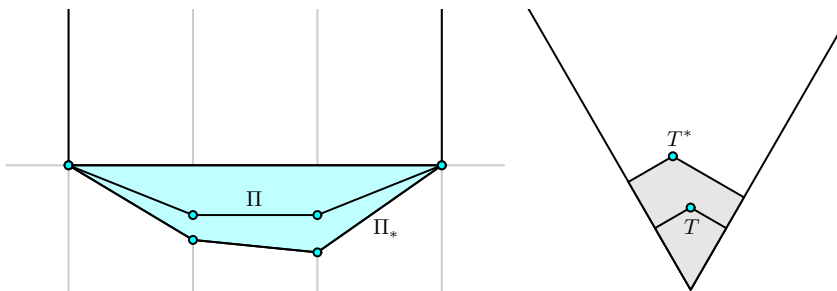


The proof is straightforward, given the description of $V_\Delta^\Delta$ in Proposition 5.6.

This might be informally phrased as saying that the points in $\mathcal{C}_\Delta^\Delta(T)$ are **T-stable**. As for the other $\mathcal{C}_\Theta^\Delta$, it is easy to see that its orthogonal projection onto the face spanned by any other $\mathcal{C}_\Theta$ is equal to the translation by $\pi_\Theta(T)$ of $\mathcal{C}_\Theta$. What about the perpendicular projection? This is onto the $\Theta$-normalized points whose profiles are convex in the segments $[d_k, d_{k+1}]$ and dominated by $\nu_\Theta(T)$. If $T$ has this profile:
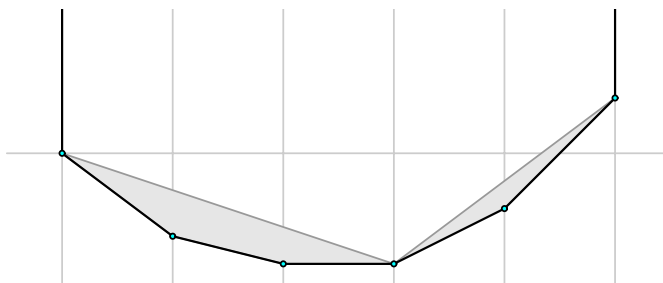


and $\Theta = \{\alpha_3, \alpha_5\}$ then the $\Theta$-normalization of $\mathcal{C}_\Theta$ contains the points whose profiles lie in this region:
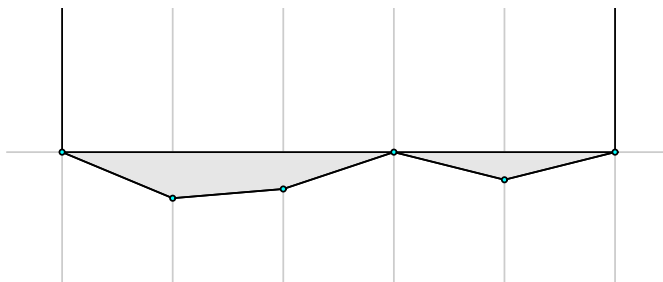
Finally, the following is a geometric formulation of an observation in the paper [Aubert-Howe:1992].

PROPOSITION 5.9. *The point of $\overline{\mathcal{C}}$ nearest to $v = (s_i)$ is that point $\overline{v}$ whose profile is the convex hull of the profile of $v$.*

*Proof.* It must be shown that $v - \overline{v}$ lies in the span of the $-\alpha_i$. If $(i, \overline{y}_i)$ lies on the hull then this means neither more nor less than that $y_i - \overline{y}_i \geq 0$. This is immediate.

There is a well known algorithm to find the convex hull of any finite set of 2D points which is particularly effective here (see Chapter 1 of [de Berg et al.:1997]). It can be roughly described as scanning from left to right, adjusting to avoid concave regions.
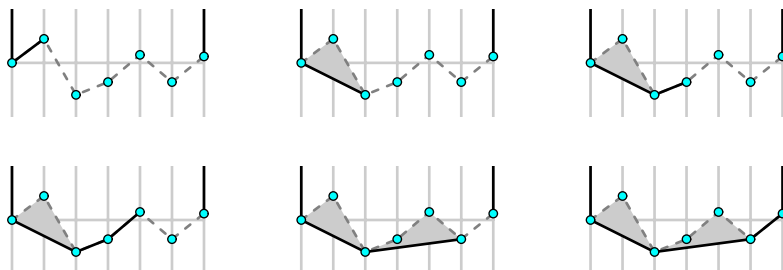


This is ridiculously efficient, since each vertex is touched only twice, and the whole process is simply proportional to the number of points in the polygon. I do not know of an algorithm of comparable efficiency for finding nearest points on an arbitrary convex subset of Euclidean space, even for arbitrary simplicial cones. Grayson's discussion of the orthogonal and symplectic groups suggests a similar algorithm for the classical root systems, but each family is dealt with in an apparently different fashion.

**6. Lattice flags.** Suppose $L$ to be a free abelian group of rank $n$ and $V = L \otimes \mathbb{R}$.

A **flag** $\mathcal{F}$ in $V$ is an increasing sequence of real vector spaces

$$V_0 = \{0\} \subset V_1 \subset \ldots \subset V_k = V .$$

I define the **dimension** of the flag to be the array $(d_i)$ of dimensions of its components, and set $\Theta = \Theta_{\mathcal{F}}$ to be the complement of these dimensions in $\{1, \ldots, n\}$. Thus for the trivial flag $\{0\} \subset V$ we have $\Theta = \{1, \ldots, n-1\}$. The stabilizer in $\mathrm{GL}(V)$ of a flag $\mathcal{F}$ is a parabolic subgroup $P = P_{\mathcal{F}}$, and the subspaces $V_i$ are called its **components**. If $\Gamma = \mathrm{GL}(L)$, we know that the quotient $\Gamma \backslash X_V$ parametrizes isomorphism classes of lattices. What does the quotient $\Gamma \cap P \backslash X_V$ parametrize?

The group $L$ induces a rational structure on $V$, and a flag is called **rational** if its components are rational. If $\mathcal{F} = (V_i)$ is a rational flag then the filtration

$$L_0 = \{0\} \subset L_1 = L \cap V_1 \subset \ldots \subset L_k = V$$

is called a **lattice flag**. Because each $V_i$ is rational, each intersection $L_i$ is a free subgroup of $L$ of rank equal to the dimension of $V_i$. If $x$ is a positive definite quadratic form on $V$ then each $L_i$ becomes a sublattice. Two lattice flags obtained from forms $x_1$ and $x_2$ and the same rational flag $\mathcal{F}$ are are isomorphic if and only if $x_2 = \gamma x_1$ with $\gamma$ in the stabilizer of $P$ as well as $\mathrm{GL}(L)$. Therefore

*The quotient $\Gamma \cap P \backslash X_V$ parametrizes lattice flags based on $\mathcal{F}$.*

The structure of this quotient is related to isomorphism classes of lattices of lower rank. An element of $P$ induces an action on each quotient $V_i/V_{i-1}$. The map from $P$ to $\prod \mathrm{GL}(V_i/V_{i-1})$ is surjective, and the kernel is the unipotent radical $N_P$ of $P$. This map therefore identifies the reductive quotient $M_P$ of $P$ with $\prod M_{P,i}$ where $M_{P,i} = \mathrm{GL}(V_i/V_{i-1})$. If $x$ is a quadratic form on $V$ then on each $V_i/V_{i-1}$ the linear isomorphism of $V_i/V_{i-1}$ with $V_i \cap V_{i-1}^\perp$ induces a quadratic form $x_i$ on $V_i/V_{i-1}$. Every $x$ in $X_V$ thus also gives rise to an orthogonal decomposition of $V$ into subspaces

$$\overline{V}_i = V_i \cap V_{i-1}^\perp \cong V_i/V_{i-1} .$$

The reductive component $M_P$ of $P$ may be canonically identified with the stabilizer of the decomposition $V = \oplus \overline{V}_i$, effecting a splitting of the canonical surjection from $P$ to $M_P$. The group $A_P$, the centre of $M_P$, may be identified with the matrices acting as scalars on each $\overline{V}_i$. By choosing a basis of $V$ compatible with the orthogonal decomposition $V = \oplus \overline{V}_i$, we represent $a$ in $A_P$ as a diagonal matrix $(a_j)$, with $a$ acting on $\overline{V}_i$ by $a_j$ if $d_{i-1} < j \leq d_i$. The map

$$\sigma_P \colon a \longmapsto (a_j)$$

is a canonical identification of $A_P$ with the subgroup $(a_j)$ of $\mathbb{R}^n$ with $a_j = a_{j+1}$ whenever $j$ is not one of the $d_i$.

The image of $L_i/L_{i-1}$ in $V_i/V_{i-1}$ is a free discrete group of maximal rank There exists also, therefore, a canonical map from $X_V$ to $\prod X_{V_i/V_{i-1}}$ induces a canonical map from lattices in $V$ to an array of lattices in the quotients $V_i/V_{i-1}$. This is $P$-covariant, and the fibres are the $N_P$-orbits in $X_V$. The quotient $\Gamma \cap P \backslash X_V$ therefore maps onto $\prod \Gamma_i \backslash X_i$ with fibres isomorphic to $\Gamma \cap N_P \backslash N_P$, where $M_i = \mathrm{GL}(V_i/V_{i-1})$, $\Gamma_i$ is the image of $\Gamma \cap P$ in $M_i$, and $X_i = X_{V_i/V_{i-1}}$.

If $P = gQg^{-1}$ are two conjugate parabolic subgroups, there is a canonical isomorphism of $A_P$ with $A_Q$, since a parabolic subgroup is its own normalizer. To each each element $a$ of $A_P$ corresponds a profile polygon—its bottom is the unique polygonal path whose slope from $x = j-1$ to $j$ is $\log|a_j|$ where $(a_j) = \sigma_P(a)$. The slopes of such polygons make up the linear subspace of $\mathbb{R}^n$ where $\alpha_i = 0$ for $i$ not in the dimension of $\mathcal{F}$.

The **canonical plot of a lattice flag** $(L_i)$ is the set of all the two-dimensional points $(\dim M, \log M))$ where $L_{i-1} \subseteq M \subseteq L_i$ for some $i$. The **profile of a lattice flag** is the unique polygon which in the range $[\dim L_{i-1}, \dim L_i]$ is equal to the convex hull of this plot. The polygon in this range, translated back to the origin, is the canonical profile $\Pi_i$ of $L_i/L_{i-1}$, which is called its $i$-th **segment**. The map taking a flag profile $\Pi$ to the sequence $(\Pi_i)$ of its segments is a bijection between the set of all flag profiles and sequences of polygons $\Pi_i$ satisfying the condition that $\Pi_i$ be the profile of a lattice of rank $\dim L_i - \dim L_{i-1}$.

The profile contains at least the points $\lambda_i = (\dim L_i, \log L_i))$. It need not be overall convex, nor do the vertices of the profile have to be points where the profile bends. Here is a typical flag profile:
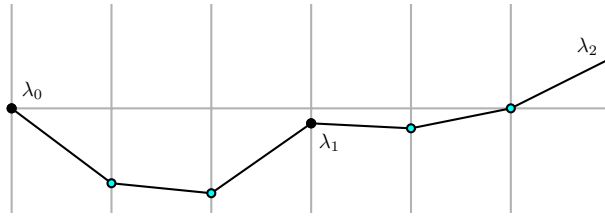
These definitions are consistent with the earlier one in a trivial sense, since the profile of a lattice $L$ is clearly the same as that of the flag $\{0\} \subset L$ determined by $L$ alone. But we also have a more interesting consistency. I say that one lattice is **subordinate** to another if its components are components of the other. This is straightforward to prove:

LEMMA 6.1. *The profile of a lattice is the same as the profile of any flag subordinate to its canonical flag.*

Any lattice may be scaled by a constant $a$, simply multiplying its metric by $|a|$. The **normalization** of a lattice is the one we get by scaling it so as to have unit volume. The effect of scaling by $a$ on the profile of a lattice is to shear it, moving each point $(d, \ell)$ to $(d, \ell + d \log |a|)$. The **geodesic action** of Borel-Serre generalizes this operation. Suppose given a lattice flag $\mathcal{F}$ and an element $a$ of $A_P$. Suppose that $a$ acts as $a_i$ on $V_i/V_{i-1}$, and therefore corresponds to an operator on all of $V$ that acts as multiplication by $a_i$ in $\overline{V}_i$. We can define a new lattice flag by changing the metric on $V$ in the natural way—if $x$ has the orthogonal decomposition $x = \oplus x_i$ with $x_i$ in $\overline{V}_i$ determined by $\mathcal{F}$, with norm $\sum \|x_i\|^2$, then the new norm of $x$ is $\sum a_i^2 \|x_i\|^2$. For example, if $a = (c, 1/c)$ in dimension 2 then $R_a$ takes $z = x + iy$ to $x + ic^2y$, whereas the usual fractional linear transformation takes $z$ to $c^2 z$. One important thing to realize about the geodesic action is that it doesn't preserve convexity of a profile, as this portrait of the profiles of a lattice under transformation by the geodesic action shows:
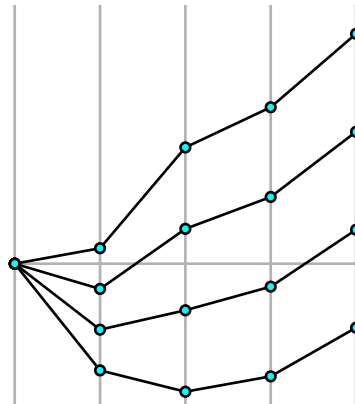


The normalization of a lattice flag $\mathcal{F}$ is the lattice obtained by normalizing each of the components in its associated graded lattice. If $v_i = \overline{L}_i$) then this normalization is also $R_a\mathcal{F}$ where $a = (v_i^{-1})$. The map taking $\mathcal{F}$ to $\nu(\mathcal{F}) = (v_i^{-1})$ defines a canonical map $\nu$ from lattice flags $\mathcal{F}$ to the connected component $A_P^0$, where $P = P_\mathcal{F}$. A flag $\mathcal{F}$ is normalized if and only if $\nu(\mathcal{F}) = 1$.

**7. The parabolic decomposition.**   I review the situation before going on.

Suppose $L$ to be a free finitely generated group, say of rank $n$, and $V = L \otimes \mathbb{R}$. The space of lattices based on $L$, that is to say that of Euclidean metrics on $L$, may be identified with $X_V$, the space of all positive definite quadratic forms on $V$. The group $\mathrm{GL}(V)$ acts on $X_V$ according to the formula

$$gx(v) = x(g^{-1}v) \,.$$

Two points $x_1$ and $x_2$ give rise to isomorphic lattices if and only if $x_1 = \gamma x_2$ with $\gamma$ in $\Gamma_L = \mathrm{GL}(L)$. If $\mathfrak{X}_V$ is the subset of lattices of discriminant 1, or equivalently those with $V/L = 1$, then $X_V \cong \mathfrak{X}_V \times \mathbb{R}^{\mathrm{pos}}$. If a form $x$ has discriminant $D$ then the projection takes $x$ to $(x/\sqrt{D}, \sqrt{D})$.

To each point $x$ of $X_V$, we associate its lattice, its profile, its canonical flag $\mathcal{F} = \mathcal{F}_x$, the stabilizer $P_x$ of that flag, hence also the slope (array) $s = s_x$ of its profile. Let $\Theta_{P_x}$ be the set of $i$ in $[0, n-1]$ such that $s_i = s_{i+1}$. The profiles of lattices are precisely the convex profile polygons, so that the image of the slope map from $X_V$ to $\mathbb{R}^n$ is precisely the closed cone $\overline{\mathcal{C}} = \mathcal{C}^\Delta$ where all $s_i \leq s_{i+1}$ for $1 \leq i \leq n-1$.

In summary, each $x$ in $X_V$ gives rise to a parabolic subgroup $P = P_x$ and a point $s = s_x$ in $\mathcal{C}_\Theta$ where $\Theta = \Theta_P$. As pointed out in [Ji-MacPherson 2002], the set of all such points $(P, s)$ with $P$ a rational parabolic subgroup of $G$ and $s$ a point of $\mathcal{C}_{\Theta_P}$ make up the interior of the cone $C(|T|)$ on the Tits complex $|T|$ of $G$. We therefore have in this case a canonical map from $X_V$ into this cone. Following Ji and MacPherson I call this cone the **rational skeleton** of $X_V$. I'll call the canonical map the **canonical skeletal projection** $\sigma$. Leslie Saper has pointed out to me that this cone occurs already, in a related manner, in [Borel-Serre 1973].

If $P$ is the rational parabolic subgroup stabilizing the flag $\mathcal{F}$, then define

$$X_P = \{x \,|\, \mathcal{F}_x = \mathcal{F}\} \,.$$

There is a canonical projection from this onto $\mathcal{C}_{\Theta_P}$. The space $X_V$ is the disjoint union of the $X_P$ as $P$ varies over all rational parabolic subgroups. The set $X_G$, in particular, parametrizes stable lattices. For any $\gamma$ in $\Gamma_L$, $\gamma X_P = X_{\gamma P \gamma^{-1}}$. The action of $\Gamma_L$ does not change the slope. Hence the first part of this:

PROPOSITION 7.1.   *The set $X_P$, and more particularly the inverse image with respect to the skeletal projection of any point of $\mathcal{C}_{\Theta_P}$, is stable under $\Gamma \cap P$ as well as $N_P$.*

We have already seen the second part proven.

From now on let $\Gamma$ be any subgroup of $\Gamma_L$ of finite index.

COROLLARY 7.2.   *The canonical map from $\Gamma \cap P \backslash X_P$ to $\Gamma \backslash X$ is an embedding. The images of $\Gamma \cap P \backslash X_P$ and $\Gamma \cap Q \backslash X_Q$ overlap if and only if $P$ and $Q$ are $\Gamma$-conjugate and in that case they are equal.*

The skeletal projection $\sigma$ maps $X_P$ onto $\mathcal{C}_{\Theta_P}$, and each fibre $\sigma^{-1}(s)$ of this map is stable with respect to $\Gamma \cap P$ and $N_P$. What is the structure of the quotient $(\Gamma \cap P)N_P \backslash \sigma^{-1}(s)$?

Suppose $\mathcal{F} = (V_i)$ to be a rational flag and $P$ its stabilizer, so that $M_P = \prod M_i$. The projection from $V_i / V_{i-1}^\perp$ to $V_i / V_{i-1}$ to gives rise to a Euclidean metric on $V_i / V_{i-1}$,

hence a point of $X_{V_i/V_{i-1}}$. These all together give rise to a canonical map from $X_V$ to $\prod X_{V_i/V_{i-1}}$. The fibres of this map are the orbits of $N_P$. According to the notation introduced above, for each $i$ the space $X_{M_i}$ is the space of semi-stable lattices in $V_i/V_{i-1}$. The image of $X_P$ in $\prod X_{V_i/V_{i-1}}$ is $X_{M_P} = \prod X_{M_i}$. To state it formally:

PROPOSITION 7.3. *If $P$ is a rational parabolic subgroup of* $\mathrm{GL}(V)$, *there exists a canonical map from $X_V$ onto $X_{M_P}$ whose fibres are the orbits of $N_P$.*

Now let $M$ be the reductive component of a rational parabolic subgroup $P$ of $\mathrm{GL}(V)$.

PROPOSITION 7.4. *There exists a canonical isomorphism $X_M \cong \mathfrak{X}_M \times V_\Theta$.*

PROPOSITION 7.5. *For any rational parabolic subgroup $P$ the set $X_P$ is stable under $N_P$. The canonical projection identifies the quotient of $X_P$ by $N_P$ with the subset of $X_{M_P}$ whose slope lies in $\mathcal{C}_{\Theta_P}$.*

We therefore understand the structure of $X_V$ reasonable well. In effect, we have reduced the question of describing it to that of describing the structure of stable unimodular lattices for all dimensions at most $n$. We have little hope of understanding the space of such lattices in any non-trivial way, but this at least is true:

PROPOSITION 7.6. *The quotient $\Gamma\backslash\mathfrak{X}_G$ is compact.*

Vectors in lattices in $\mathfrak{X}_G$ have bounded minimal length and the lattices have unit volume. This therefore follows from Mahler's criterion, which I'll recall in the next section.

From the canonical skeletal projection $\sigma$ a whole family of skeletal projections can be constructed. They are parametrized by points of $\mathcal{C} = \mathcal{C}^\Delta$.
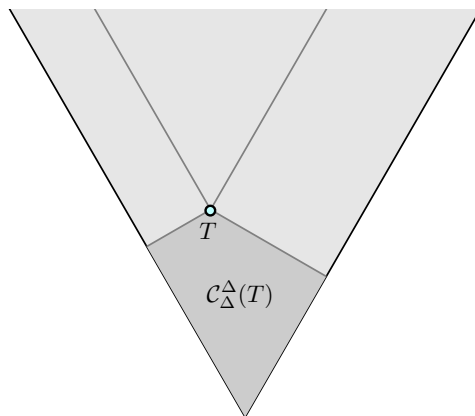


Let $T$ be an arbitrary point of $\mathcal{C}$. First of all define $X_G(T)$ to be the set of all $x$ in $X$ for which the slope $s_x$ lies in $\mathcal{C}_\Delta^\Delta(T)$, the points of $\mathcal{C}$ dominated by $T$. If $T = 0$ these are just the usual semi-stable lattices, and in general I'll call them **T-stable**. Let $\mathfrak{X}_G(T)$ be the intersection of $X_G(T)$ with $\mathfrak{X}$. Both of these are stable under $\Gamma_L$. If $T$ lies in the face $\mathcal{C}_\Theta$ then $X_G(T)$ contains points in all the $X_P$ with $\Theta \subset \Theta_P$.

PROPOSITION. 7.7. *The quotient $\Gamma\backslash\mathfrak{X}_G(T)$ is compact.*

This also follows from Mahler's criterion.
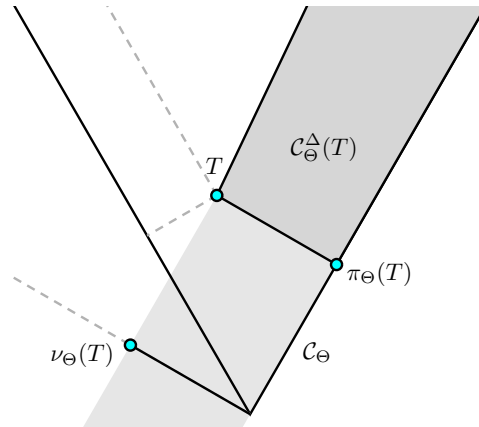
Let $P$ be a rational parabolic subgroup and $\Theta = \Theta_P$. Let $X_P(T)$ be those $x$ in $X_Q$ with $Q \subseteq P$ for which the slope lies in $\mathcal{C}_\Theta(T)$. This agrees with the earlier definition of $X_G(T)$. The product structure of $\mathcal{C}_\Theta(T)$ described in the earlier section on the Langlands decomposition shows that this has the product structure $A_P(T)/A \cap K$ times $\prod \mathfrak{X}_{V_i/V_{i-1}}(T)$, where $A_P(T)$ is the inverse image in $A_P^0$ of $\pi_\Theta(T) + \mathcal{C}_\Theta$. Note that $A_P^0 \cong \mathcal{C}_\Theta$ *via* logarithms, and the right action of $A_P^0$ is compatible with this. The skeletal projection $\sigma_T$ associated to $T$ takes a point in $X_P(T)$ to the point $x\, \pi_\Theta(s_x) - \pi_\Theta(T)$, which lies in $C_\Theta$. The previous results for the sets $X_P$ have straightforward analogues for $X_P(T)$.

There is one final useful remark. Suppose $P$ to be a maximal rational parabolic subgroup of $\mathrm{GL}(V)$. $\mathcal{F}$ the corresponding flag $V_0 \subset V_1 \subset V$. Let $d$ be the dimension of $V_1$, $I$ the complement of $d$ in $I_n$, $\Theta = \Theta_I$.

To each $x$ in $X_V$ corresponds a lattice flag $\mathcal{F} \cap L$, and hence to each $x$ the profile of this flag, a point in $\mathbb{R}^n$ in the region $\alpha_{d_P} \mathrm{ge} 0$, and then the projection onto the line containing $\mathcal{C}_\Theta$. Define $X_P^+(T)$ to be the inverse image of $\pi_\Theta(T) + \mathcal{C}_\Theta$ in $X$. The following is a basic fact of reduction theory, but as far as I can say it was first observed by Arthur.

PROPOSITION 7.8. *For any $T$ in $\overline{C}$ the $X_G(T)$ is the complement of the union of $X_P^+(T)$ as $P$ ranges over the maximal rational parabolic subgroups of $G$.*

On the one hand the region $\mathcal{C}_\Delta(T)$ is the complement in $\mathcal{C}$ of the projections onto the lines $\pi_\Theta(T) + \mathcal{C}_\Theta$. This means that $X_G(T)$ is contained in the intersection. On the other, if $x$ lies in one of these regions then the profile of $x$ has to lie below the profile of $\pi_\Theta(T)$, and $x$ cannnot lie in $X_G(T)$.

**8. Mahler's criterion.** I formulate here a variant of Mahler's criterion for the relative compactness of a set of lattices.

Suppose $A$ and $B$ to be positive numbers, $E$ a Eucidean real vector space of dimension $n$. I define a **weakly reduced frame** in $E$ with respect to $A$ and $B$ to be any subset of $n$ vectors $v_i$ satisfying the following conditions:

- $\|v_i\| \leq B$ *for all $i$;*
- *for each $j$ the projection of $v_j$ onto the subspace perpendicular to $v_1, \ldots, v_{j-1}$ has length at least $A$.*

Since projections do not increase length, the second condition implies also that the projection of any $v_i$ with $i \geq j$ onto the subspace perpendicular to $v_1, \ldots, v_{j-1}$ has length at least $A$. Recursively, this definition amounts to requiring that (1) $A \leq \|v_i\| \leq B$ for all $i$ and (2) the projections $v_i^\perp$ $(i \geq 2)$ perpendicular to $v_1$ form a weakly reduced frame of dimension one less for $A$ and $B$.

In these circumstances the volume of the parallelopiped spanned by the $v_i$ is at least $A^n$. As a consequence, any weakly reduced frame is actually a frame—i.e. a basis of $E$—and for a given $A$ and $B$ the set of all associated frames is a compact subset of frames.

THEOREM. *There exist for every $a$, $K > 0$ and positive integer $n$ constants $A$ and $B$ such that if $L$ is any lattice of dimension $n$, with volume at most $K$ and all its vectors of length $a$ or more, then $L$ possesses a basis which is weakly reduced with respect to $A$ and $B$.*

From this it follows immediately that every $\Gamma \backslash X_n(T)$ is compact.

*Proof.* For $n = 1$ the Theorem is clear, since volume and length are the same.

The proof continues by induction on $n$. In the proof it will be shown that one may choose $A$ to be $(\sqrt{3}/2)^{n-1}a$, and I'll take this to be part of the induction assumption. Let $\mu_n$ be the volume of the unit ball $B_n(1)$ in Euclidean space $\mathbb{R}^n$. A classic theorem of Minkowski asserts that if we choose $r$ so that

$$B_n(r)) = \mu_n r^n \geq 2^n L) \text{ or } r \geq 2 \, (L)/\mu_n)^{1/n}$$

then $L$ will contain a vector inside $B(r)$. Since the volume of $L$ is bounded by $K$, if we choose $r = b = 2(K/\mu_n)^{1/n}$ we can find a vector $v_1$ of length at most $r$ inside $L$. We may assume it to be a vector of least length in $L$, and in particular that it be primitive in $L$. The vector $v_1$ now satisfies the conditions

$$a \leq \|v_1\| \leq b \, .$$

I claim that the quotient $L_* = L/\mathbb{Z}v_1$ satisfies the same conditions as those on $L$, but of course with possibly different constants $a_*$, $K_*$. First of all, the volume of $L_*$ is equal to $L)/\|v_1\|$, which is at most $K_* = K/a$. It remains to show that the lengths of vectors in $L_*$ are bounded from below by a suitable constant. This result is made more explicit in the following result, which is an easy consequence of Lagrange's Lemma.

LEMMA 8.1. *If*
$$\|v\| \geq a$$

*for all non-zero vectors in $L$ then*

$$\|v_*\| \geq a_* = \frac{\sqrt{3}}{2}a$$

*for all non-zero vectors $v_*$ in $L_*$.*

*Proof.* Choose $v$ representing $v_*$ as suggested by Lagrange's Lemma. Thus $v = v_* + \alpha v_1$, with $|\alpha| \leq 1/2$, and

$$\|v_*\|^2 + \alpha^2 \|v_1\|^2 = \|v\|^2$$

Since $v_1$ has least length, we also have

$$\|v\|^2 \geq \|v_1\|^2, \quad \|v_*\|^2 \geq (1 - \alpha^2) \|v_1\|^2 \geq (3/4) \|v_1\|^2 \geq (3/4) a^2$$

which concludes the proof of the Lemma.

To conclude the proof of the Theorem, note that by the induction assumption we can find a basis $(v_{*,i})$ (for $i \geq 2$) of $L_*$, $A_* = (\sqrt{3}/2)^{n-1} a$ and $B_* > 0$ satisfying its conclusion. We may lift each $v_{*,i}$ to a vector $v_i$ with $|v_i \bullet v_1| \leq \|v_1\|^2/2$. The $v_i$ form a basis of $L$. But now we have

$$\|v_i\|^2 = \|v_{*,i}\|^2 + \|u_i\|^2 \leq (B_*)^2 + b^2/4$$

if $u_i$ is the projection of $v_i$ onto the line through $v_1$. This proves the theorem, with $A = A_*$ and $B = \sqrt{(B_*)^2 + b^2/4}$.  $\square$

This proof is (of course) not much different in substance from either of the proofs found in §1 of [Borel 1972] or Chapters V and VIII of [Cassels 1959], but is perhaps somewhat more direct.

## REFERENCES

[1]   J. ARTHUR, *A trace formula for reductive groups I: terms associated to classes in $G(\mathbb{Q})$*, Duke Math. Jour., 45 (1978), pp. 911–952.

[2]   ANNE-MARIE AUBERT AND ROGER HOWE, *Géometrie des cônes aigus et application à la projection euclidienne sur la chambre de Weyl positive*, Journal of Algebra, 149 (1992), pp. 472–493.

[3]   M. DE BERG, M. VAN KREVELD, M. OVERMARS, AND O. SCHWARZKOPF, *Computational Geometry: algorithms and applications*, Springer-Verlag, 1997.

[4]   A. BOREL, *Introduction aux groupes arithmétiques*, Hermann, Paris, 1969.

[5]   A. BOREL AND J-P. SERRE, *Corners and arithmetic groups*, Comm. Math. Helv., 48 1973, pp. 436.

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

[7]   J. W. S. CASSELS, *Introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1959.

[8]   D. GRAYSON, *Reduction theory using semi-stability*, Comment. Math. Helvetici, 59 (1984), pp. 600–634.

[9]   D. GRAYSON, *Reduction theory using semi-stability II*, Comment. Math. Helvetici, 61 (1986), pp. 661–676.

[10]  G. HARDER, *Minkowski Reduktionstheorie über Funktionenkörpern*, Inv. Math., 7 (1969), pp. 33–54.

[12]  G. HARDER AND M. NARASIMHAN, *On the cohomology groups of moduli spaces of vector bundles on algebraic curves*, 212 (1975), pp. 215.

[13]  G. HARDER AND U. STUHLER, *Reduction theory*, preprint, University of Bonn, (1997). This used to be available on the Internet as the file `redap5.ps`, but this seems no longer to be true. A preliminary version can be found at the web site for the Trieste conference (which you can locate by googling `Harder Stuhler reduction theory`).

[14]  G. H. HARDY AND E. WRIGHT, *Number theory*, Oxford Press, 1959.

[15]  R. P. LANGLANDS, *On the classification of irreducible representations of real algebraic groups*, in Representation Theory and Harmonic Analysis on Semisimple Lie Groups, edited by P. Sally and D. Vogan, pp. 101–170. American mathematical Society, 1989. (This circulated as a preprint for several years before publication.)

[16]  E. LEUZINGER, *Exhaustion of locally symmetric spaces by compact submanifolds with corners*, Invent. Math., 121 (1995), pp. 389–410.

[17]  LIZHEN JI AND R. MACPHERSON, *Geometry of compactifications of locally symmetric spaces*, Ann. Inst. Fourier Grenoble, 52 (2002), pp. 457–559.

[18]  Y. Manin, *Three-dimensional hyperbolic geometry as $\infty$-adic Arakelov geometry*, Invent. Math., 104 (1991), pp. 223–243.

[19]  M. S. Osborne and G. Warner, *Partition, truncation, reduction*, Pac. Jour. Math., 106 (1983), pp. 307–495.

[20]  L. Saper, *Tilings and finite energy retractions of locally symmetric spaces*, Comm. Math. Helv., 72 (1997), pp. 167–202.

[21]  I. Satake, *On compactifications of the quotient spaces for arithmetically defined discontinuous groups*, Ann. Math., 72 (1960), pp. 555–580.

[22]  U. Stuhler, *Eine Bermerkung zur Reduktionstheorie quadratischen Formen*, Archiv der Math., 27 (1976), pp. 604.