

University of Massachusetts Amherst  
**ScholarWorks@UMass Amherst**

---

Mathematics and Statistics Department Faculty  
Publication Series

Mathematics and Statistics

---

2006

# Robert MacPherson and arithmetic groups

PE Gunnells

*University of Massachusetts - Amherst*, [gunnells@math.umass.edu](mailto:gunnells@math.umass.edu)

Follow this and additional works at: [https://scholarworks.umass.edu/math\\_faculty\\_pubs](https://scholarworks.umass.edu/math_faculty_pubs)

 Part of the [Physical Sciences and Mathematics Commons](#)

---

## Recommended Citation

Gunnells, PE, "Robert MacPherson and arithmetic groups" (2006). *PURE AND APPLIED MATHEMATICS QUARTERLY*. 398.  
Retrieved from [https://scholarworks.umass.edu/math\\_faculty\\_pubs/398](https://scholarworks.umass.edu/math_faculty_pubs/398)

This Article is brought to you for free and open access by the Mathematics and Statistics at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Mathematics and Statistics Department Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

## ROBERT MACPHERSON AND ARITHMETIC GROUPS

PAUL E. GUNNELLS

*Dedicated to Robert MacPherson on the occasion of his sixtieth birthday*

ABSTRACT. We survey contributions of Robert MacPherson to the theory of arithmetic groups. There are two main areas we discuss: (i) explicit reduction theory for Siegel modular threefolds, and (ii) constructions of compactifications of locally symmetric spaces. The former is joint work with Mark McConnell, the latter with Lizhen Ji.

## CONTENTS

1. Introduction	1
2. Arithmetic groups	3
3. Reduction theory for Siegel modular threefolds	7
4. Compactifications of locally symmetric spaces	17
References	34

## 1. INTRODUCTION

1.1. Arithmetic groups sit at the crossroads of many areas of mathematics. On one hand, they lead to beautiful manifolds with intricate geometry, and to moduli spaces for many important objects in arithmetic. On the other, they are conjoined with the theory of automorphic forms, and provide one path to understanding the mysteries of the absolute Galois group of the rationals. The study of arithmetic groups is a beautiful blend of algebraic topology, algebraic and differential geometry, representation theory, and number theory, and includes some of the most fascinating and inscrutable phenomena in mathematics.

---

*Date:* February 19, 2006.

*1991 Mathematics Subject Classification.* 11F23, 11F46, 11F75, 20G30, 22E40, 54D35,

*Key words and phrases.* Cohomology of arithmetic groups, reduction theory, compactifications of locally symmetric spaces.

1.2. In this survey we discuss Robert MacPherson’s contributions to arithmetic groups. We focus on two of his collaborations. The first (§3) is joint work with Mark McConnell, and appears in the papers *Explicit reduction theory for Siegel modular threefolds* (Inv. Math **111** (1993)) [39] and *Classical projective geometry and modular varieties* (Proceedings of JAMI 1988) [38]. The second (§4) is joint work with Lizhen Ji, and appears in *Geometry of compactifications of locally symmetric spaces* (Ann. Inst Fourier, Grenoble **52** (2002)) [31]. We also provide (§2) some background on arithmetic groups and their relationship to number theory and automorphic forms.

1.3. There are two other contributions of MacPherson to arithmetic groups that we unfortunately will not discuss, the *topological trace formula* and the *geometric approach to the fundamental lemma*. To compensate for this omission, we say a few words here.

In [2], using his trace formula, Arthur derived an expression for the Lefschetz number of the action of a Hecke correspondence on the  $L^2$ -cohomology of a modular variety. In view of the Zucker conjecture (proved by Saper and Stern [46] and, independently, by Looijenga [36]), which identifies this  $L^2$ -cohomology with the intersection cohomology of the Baily–Borel compactification (§4.2), Casselman and Arthur asked whether Arthur’s formula could be interpreted as a Lefschetz fixed point formula for intersection cohomology, and in particular whether each term in Arthur’s formula might correspond to a single fixed point component. The ingredients (volumes of centralizers, orbital integrals, and averaged discrete series characters) in Arthur’s formula did not, initially, look like the local contributions one might expect from a Lefschetz fixed point formula.

Nevertheless, Arthur and Casselman’s suggestion was eventually realized in a series of papers [19–21, 25–27] in which it was shown that each term in Arthur’s formula corresponds to a sum of (Lefschetz) contributions over a certain collection of fixed points. New techniques were developed including (i) a general topological formula for the local contribution to the Lefschetz number from “hyperbolic” fixed points; see also the closely related results of Kashiwara and Schapira [32]; (ii) the “weighted cohomology” of modular varieties (similar to the intersection cohomology, but involving a weight truncation rather than a degree truncation), later shown [43] to coincide with the *weighted  $L^2$ -cohomology* of Franke [17]; and (iii) a combinatorial formula for the characters of discrete series, somewhat different from the formulas [29] of Herb. For more details about the actual constructions involved in this program, we refer to the announcement [25].

For the geometric approach to the fundamental lemma, we have considerably less to say. This program is still under development, with some publications available [22, 23] and others in preparation. Recently Laumon–Ngo extended and improved the original strategy of Goresky–Kottwitz–MacPherson and completed a proof of the fundamental lemma for unitary groups, a spectacular advance [34]. It is premature to predict where

all this will lead, but some experts share a certain optimism. In any case, writing a detailed survey of this subject would require a long and complicated article of its own.

**1.4. Acknowledgments.** I thank Avner Ash, Mark Goresky, Lizhen Ji, and Mark McConnell for helpful comments. I also thank the editors for inviting me to write this article. Finally, I'm glad to thank (in print) Bob MacPherson, not only for helping to create such beautiful mathematics, but also for being a patient and benevolent teacher.

## 2. ARITHMETIC GROUPS

2.1. In this section we review background on algebraic groups, arithmetic groups, and discuss their relationships to geometry and number theory.

### 2.2. Linear algebraic groups.

2.2.1. The starting point is  $\mathbf{G}$ , a connected linear algebraic group defined over  $\mathbb{Q}$ . For our purposes, this means the following [44, §2.1.1]:

- (1) The group  $\mathbf{G}$  has the structure of an affine algebraic variety given by an ideal  $I$  in the ring  $\mathbb{C}[x_{ij}, D^{-1}]$ , where the variables  $\{x_{ij} \mid 1 \leq i, j \leq n\}$  should be interpreted as the entries of an “indeterminate matrix,” and  $D$  is the polynomial  $\det(x_{ij})$ . Both the group multiplication  $\mathbf{G} \times \mathbf{G} \rightarrow \mathbf{G}$  and inversion  $\mathbf{G} \rightarrow \mathbf{G}$  are required to be morphisms of algebraic varieties.

The ring  $\mathbb{C}[x_{ij}, D^{-1}]$  is the coordinate ring of the algebraic group  $\mathrm{GL}_n$ . Hence we are viewing  $\mathbf{G}$  as a subgroup of  $\mathrm{GL}_n$  defined by the vanishing of polynomial equations in matrix entries.

- (2) *Defined over  $\mathbb{Q}$*  means that the ideal  $I$  is generated by polynomials with rational coefficients.
- (3) *Connected* means that  $\mathbf{G}$  is connected as an algebraic variety.

Given any subring  $R \subset \mathbb{C}$ , we can consider the group of  $R$ -rational points  $\mathbf{G}(R)$ . As a set  $\mathbf{G}(R)$  is the set of solutions to the defining equations for  $\mathbf{G}$  with coordinates in  $R$ . Especially important in the following will be the groups of *real points*  $\mathbf{G}(\mathbb{R})$ , *rational points*  $\mathbf{G}(\mathbb{Q})$ , and *integral points*  $\mathbf{G}(\mathbb{Z})$ . Following a usual convention, we denote algebraic groups by bold letters and their groups of real points by the corresponding Roman letter. Hence we write  $G = \mathbf{G}(\mathbb{R})$ ,  $P = \mathbf{P}(\mathbb{R})$ , and so on.

2.2.2. A linear algebraic group  $\mathbf{G}$  is *reductive* if its maximal connected unipotent normal subgroup is trivial, and *semisimple* if its maximal connected solvable normal subgroup is trivial.

The prototypical example of a reductive group is  $\mathbf{G} = \mathrm{GL}_n$ , the *split general linear group*. For any ring  $R \subset \mathbb{C}$ , the group  $\mathrm{GL}_n(R)$  is the group of  $n \times n$  invertible matrices with entries in  $R$ .

There are two basic examples of semisimple groups that will be important for us. The first is the *split special linear group*  $\mathbf{G} = \mathrm{SL}_n$ . For any ring  $R \subset \mathbb{C}$ , we have  $\mathbf{G}(R) = \mathrm{SL}_n(R)$ , which is the group of all  $n \times n$  matrices with entries in  $R$  and with determinant equal to one. The second is the *split symplectic group*  $\mathrm{Sp}_{2n}$ . The group of real points  $\mathrm{Sp}_{2n}(\mathbb{R})$  is the automorphism group of a fixed nondegenerate alternating bilinear form on a real vector space of dimension  $2n$ .

For some nonsplit examples, let  $F$  be an algebraic number field. Then there is a  $\mathbb{Q}$ -group  $\mathbf{G}_F$  satisfying  $\mathbf{G}(\mathbb{Q}) = \mathrm{SL}_n(F)$ . The group  $\mathbf{G}_F$  is constructed via the technique of *restriction of scalars* from  $F$  to  $\mathbb{Q}$ ; the notation is  $\mathbf{G}_F = \mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_n$  [44, §2.1.2]. For example, if  $F$  is totally real, the group  $\mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_2$  plays an important role in the study of *Hilbert modular forms*.

2.2.3. Another family of examples is provided by *tori*. A torus is a linear algebraic group  $\mathbf{T}$  such that  $\mathbf{T} \simeq \mathbf{D}_n$ , where  $\mathbf{D}_n \subset \mathrm{GL}_n$  is the subgroup of diagonal matrices. We have  $\mathbf{T}(\mathbb{C}) \simeq (\mathbb{C}^\times)^n$ . The number  $n$  is called the (*absolute*) *rank* of  $\mathbf{T}$ . A torus is said to be  $\mathbb{Q}$ -*split* if  $\mathbf{T}(\mathbb{Q}) \simeq (\mathbb{Q}^\times)^n$ , where  $n$  is the absolute rank of  $\mathbf{T}$ , and if the isomorphism is defined over  $\mathbb{Q}$ . The  $\mathbb{Q}$ -*rank* of  $\mathbf{G}$  is defined to be the dimension of a maximal  $\mathbb{Q}$ -split torus. For example, the  $\mathbb{Q}$ -rank of  $\mathrm{SL}_n$ , or more generally  $\mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_n$ , is  $n - 1$ .

### 2.3. Symmetric spaces.

2.3.1. Let  $\mathbf{G}$  be connected semisimple, and let  $G = \mathbf{G}(\mathbb{R})$  be the group of real points. Then  $G$  is a connected Lie group. Let  $K \subset G$  be a maximal compact subgroup. The quotient  $X := G/K$  is called a (*global*) *Riemannian symmetric space*; it is diffeomorphic to a contractible smooth real manifold.

For example, if  $\mathbf{G}$  is the special linear group  $\mathrm{SL}_n$ , then  $G$  is the Lie group  $\mathrm{SL}_n(\mathbb{R})$ . For a maximal compact subgroup  $K$  we can take  $\mathrm{SO}(n)$ , the special orthogonal group. This is the group of  $n \times n$  real matrices  $A$  with  $A^{-1} = A^t$  and  $\det A = 1$ . Thus  $X = \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}(n)$ . It is easy to compute that  $\dim X = n(n + 1)/2 - 1$ . The most familiar example from this family of symmetric spaces is  $n = 2$ . We have  $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2) \simeq \mathfrak{H}$ , where  $\mathfrak{H} \subset \mathbb{C}$  is the *upper halfplane* of complex numbers with positive imaginary part.

If  $G = \mathrm{Sp}_{2n}(\mathbb{R})$ , then  $K = \mathrm{U}(n)$ , the *unitary group* of all  $n \times n$  complex matrices  $A$  with  $A^* = A^{-1}$ , where the star denotes conjugate transpose. One can show that the symmetric space  $X$ , called the *Siegel upper halfspace*, has real dimension  $n^2 + n$ .

2.4. The nonsplit examples lead to more complicated geometry. Consider  $\mathbf{G} = \mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_2$ , where  $F$  is a quadratic extension of  $\mathbb{Q}$ . Any such field has the form  $F = \mathbb{Q}(\sqrt{d})$ , where  $d \neq 0, 1$  is squarefree. There are two cases to consider,  $F$  real and  $F$  imaginary.

If  $F$  is real, then  $d > 0$ . There are two ways to embed  $F$  as a subfield of  $\mathbb{R}$ , corresponding to the two choices for the square root of  $d$ . Denote these embeddings

by  $a \mapsto a^{(i)}$ ,  $i = 1, 2$ . We have  $G = \mathbf{G}(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ . A maximal compact subgroup is  $K = \mathrm{SO}(2) \times \mathrm{SO}(2)$ , and the symmetric space is  $X = \mathfrak{H} \times \mathfrak{H}$ .

Now suppose  $F$  is imaginary. This time  $G = \mathrm{SL}_2(\mathbb{C})$  and  $K = \mathrm{SU}(2)$ ; the symmetric space  $X$  is the hyperbolic 3-space  $\mathfrak{H}_3$ .

## 2.5. Locally symmetric spaces.

2.5.1. Let  $\Gamma \subset \mathbf{G}(\mathbb{Q})$  be an *arithmetic group*. By definition this is a subgroup commensurable with  $\mathbf{G}(\mathbb{Z})$ , which means  $\Gamma \cap \mathbf{G}(\mathbb{Z})$  has finite index in both  $\Gamma$  and  $\mathbf{G}(\mathbb{Z})$ . For example,  $\Gamma = \mathrm{SL}_n(\mathbb{Z}) \subset \mathrm{SL}_n(\mathbb{Q})$  is arithmetic, as is any conjugate  $g\Gamma g^{-1}$ , where  $g \in \mathrm{SL}_n(\mathbb{Q})$ .

2.5.2. The left action of  $\Gamma$  on the symmetric space  $X$  is properly discontinuous. Suppose  $\Gamma$  is torsion-free. Then  $\Gamma \backslash X$  is a manifold, and is called a *locally symmetric space*.

## 2.6. Cohomology of arithmetic groups.

2.6.1. We continue to assume that  $\Gamma$  is torsion-free. Since  $X$  is contractible,  $\Gamma \backslash X$  is an Eilenberg–Mac Lane space for  $\Gamma$ . In particular  $\pi_1(\Gamma \backslash X) \simeq \Gamma$ , and all other homotopy groups vanish. The group cohomology of  $\Gamma$  (with trivial complex coefficients) is isomorphic to the complex cohomology of the quotient  $\Gamma \backslash X$ :

$$(2.6.1) \quad H^*(\Gamma; \mathbb{C}) \simeq H^*(\Gamma \backslash X; \mathbb{C}).$$

In fact (2.6.1) remains true even if  $\Gamma$  has torsion, since we are using complex coefficients.<sup>1</sup>

More generally, let  $M$  be a complex finite dimensional rational representation of  $G$ . Then  $M$  is also a  $\Gamma$ -module, and one can define the group cohomology  $H^*(\Gamma; M)$  of  $\Gamma$  with coefficients in  $M$ . One can associate to  $M$  a locally constant sheaf  $\mathcal{M}$  on  $\Gamma \backslash X$  so that

$$H^*(\Gamma; M) \simeq H^*(\Gamma \backslash X; \mathcal{M}).$$

Again, this remains true even if  $\Gamma$  has torsion, although we must use locally constant sheaves on orbifolds.

2.6.2. It turns out that the cohomology groups  $H^*(\Gamma; M)$  are intimately related to automorphic forms. The simplest interesting example is  $\mathbf{G} = \mathrm{SL}_2$ . The basic arithmetic group here is  $\mathbf{G}(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})$ , which acts on the symmetric space  $\mathfrak{H}$  by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \quad z \in \mathfrak{H}.$$

---

<sup>1</sup>The isomorphism (2.6.1) is not true for *integral* coefficients if  $\Gamma$  has torsion. Similar remarks apply to more general coefficient modules  $M$  than those we consider in this paper.

The group  $\mathrm{SL}_2(\mathbb{Z})$  is the most obvious arithmetic group in  $\mathrm{SL}_2(\mathbb{Q})$ , but there are many others. For any  $N \geq 1$ , let  $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$  be the subgroup of matrices that are upper triangular modulo  $N$ . The locally symmetric space  $Y_0(N) = \Gamma_0(N) \backslash \mathfrak{H}$  is the *(open) modular curve of level  $N$* . Topologically,  $Y_0(N)$  is a punctured surface with genus roughly  $N/12$ . The Eichler–Shimura isomorphism connects the cohomology of  $Y_0(N)$  with holomorphic modular forms. In particular, we have

$$H^1(\Gamma_0(N); \mathbb{C}) = H^1(Y_0(N); \mathbb{C}) \simeq S_2(N) \oplus \overline{S}_2(N) \oplus \mathrm{Eis}_2(N),$$

where  $S_2(N)$  is the space of weight 2 cuspidal modular forms of level  $N$ , the bar denotes complex conjugation, and  $\mathrm{Eis}_2(N)$  is the space of weight 2 Eisenstein series of level  $N$ . More generally, if  $k \geq 0$ , let  $M_k = \mathrm{Sym}^k(\mathbb{C}^2)$  be the  $k$ th symmetric power of the standard representation of  $\Gamma_0(N)$ , and let  $\mathcal{M}_k$  be the associated locally constant sheaf on  $Y_0(N)$ . Then we have, in obvious notation,

(2.6.2)

$$H^1(\Gamma_0(N); M_k) = H^1(Y_0(N); \mathcal{M}_k) \simeq S_{k+2}(N) \oplus \overline{S}_{k+2}(N) \oplus \mathrm{Eis}_{k+2}(N) \quad k \geq 0.$$

Hence the group cohomology for these modules is directly related to modular forms of level  $N$ .

2.6.3. Similar phenomena occur for our nonsplit examples. Let  $\mathbf{G} = \mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_2$ , where  $F$  is the quadratic field  $\mathbb{Q}(\sqrt{d})$ . Inside  $F$  is a certain subring  $\mathcal{O}_F$ , the ring of algebraic integers, that plays the same role for  $F$  that the ring  $\mathbb{Z}$  does for  $\mathbb{Q}$ . We have  $\mathcal{O}_F = \mathbb{Z}[\sqrt{d}]$  unless  $d \equiv 1 \pmod{4}$ , in which case  $\mathcal{O}_F = \mathbb{Z}[(1 + \sqrt{d})/2]$ . The basic arithmetic group here is  $\Gamma = \mathbf{G}(\mathbb{Z}) = \mathrm{SL}_2(\mathcal{O}_F)$ .

Suppose  $F$  is real. Then  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$  acts on the symmetric space  $X = \mathfrak{H} \times \mathfrak{H}$  by fractional linear transformations, where on the factors we use the two embeddings  $F \rightarrow \mathbb{R}$  to map  $\mathrm{SL}_2(\mathcal{O}_F)$  into  $\mathrm{SL}_2(\mathbb{R})$ :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot (z_1, z_2) = \left( \frac{a^{(1)}z_1 + b^{(1)}}{c^{(1)}z_1 + d^{(1)}}, \frac{a^{(2)}z_2 + b^{(2)}}{c^{(2)}z_2 + d^{(2)}} \right), \quad (z_1, z_2) \in \mathfrak{H} \times \mathfrak{H}.$$

The quotient  $\Gamma \backslash X$  is an algebraic surface, called a *Hilbert modular surface*. Its cohomology can be described explicitly in terms of Hilbert modular forms.

If  $F$  is imaginary, the group  $\Gamma$  is called a *Bianchi group*. The quotient  $\Gamma \backslash \mathfrak{H}_3$  is a 3-dimensional hyperbolic orbifold, and its cohomology can also be computed in terms of appropriate automorphic forms.

2.6.4. For a general arithmetic group  $\Gamma$ , the relationship between cohomology and automorphic forms is captured by the following deep theorem of Franke [17].

Assume the  $\Gamma$ -module  $M$  arises from a complex finite-dimensional rational representation of  $G$  as in §2.6.1. Then Franke’s result says that the cohomology  $H^*(\Gamma \backslash X; \mathcal{M})$  can be systematically built out of automorphic forms attached to  $G$  and to certain

subquotients of  $G$ . Specifically, we have a decomposition

$$(2.6.3) \quad H^*(\Gamma; M) = H_{\text{cusp}}^*(\Gamma; M) \oplus \bigoplus_{\{P\}} H_{\{P\}}^*(\Gamma; M),$$

where the direct sum is taken over the set of classes of *associate proper  $\mathbb{Q}$ -parabolic subgroups of  $G$*  (§4.4.3). The summand  $H_{\text{cusp}}^*(\Gamma; M)$  corresponds to the full group  $G$ , and is known as the *cuspidal cohomology*; this is the subspace of cohomology classes represented by cuspidal automorphic forms. The remaining summands constitute the *Eisenstein cohomology* of  $\Gamma$ . In particular the summand indexed by  $\{P\}$  is built of Eisenstein series and their residues attached to suitable cuspidal automorphic forms on the *Levi quotients* (§4.4.4) of elements of  $\{P\}$ ; these are the subquotients alluded to above.

For  $G = \text{SL}_2(\mathbb{R})$ , this decomposition is exactly the right side of Eichler-Shimura isomorphism (2.6.2). The cuspidal cohomology is the subspace  $S_{k+2}(N) \oplus \overline{S}_{k+2}(N)$ , and the Eisenstein cohomology is the subspace  $\text{Eis}_{k+2}(N)$ .

For an exposition of Franke's result, as well as more information about the cohomology of arithmetic groups, we highly recommend the recent survey [35].

### 3. REDUCTION THEORY FOR SIEGEL MODULAR THREEFOLDS

#### 3.1. Reduction theory.

3.1.1. Let  $\mathbf{G}$  be a connected semisimple group, let  $X = G/K$  be the associated symmetric space, and let  $\Gamma \subset \mathbf{G}(\mathbb{Q})$  be an arithmetic group. The goal of reduction theory is understanding the action of  $\Gamma$  on  $X$ . In particular, one wants to find a “nice” fundamental domain for the action of  $\Gamma$  on  $X$ . This should be an open set  $\Omega \subset X$  such that

- (1) the union of the  $\Gamma$ -translates of the closure  $\overline{\Omega}$  is all of  $X$ , and
- (2) for all  $\gamma \in \Gamma$  with  $\gamma \neq 1$ , we have  $\gamma\Omega \cap \Omega = \emptyset$ .

Of course, what “nice” means is a matter of taste, but commonly this is taken to mean that  $\Omega$  is locally homeomorphic to a polytope, and is of finite type in some sense.

3.1.2. The prototypical example is  $\Gamma = \text{SL}_2(\mathbb{Z})$  and  $X = \mathfrak{H}$ . The classical fundamental domain  $\Omega$ , shown in the left of Figure 3.1, is the set

$$\Omega = \{z \in \mathfrak{H} \mid |z| > 1, -1/2 < \Im z < 1/2\}.$$

This example is also the source of the name *reduction theory*. There is a close connection between  $\mathfrak{H}$  and the space of binary positive definite quadratic forms  $Q(x, y) = Ax^2 + Bxy + Cy^2$  (§3.3). The  $\Gamma$ -action on  $\mathfrak{H}$  corresponds to unimodular change of variables: the matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  takes  $Q(x, y)$  to  $Q(ax + cy, bx + dy)$ .



This change of variables is a natural equivalence relation on quadratic forms for the following reason. Given a positive definite binary quadratic form  $Q$ , we can form the theta series

$$\Theta(Q) = \sum_{x,y \in \mathbb{Z}} q^{Q(x,y)}, \quad \text{where } q = \exp(2\pi iz), \quad z \in \mathfrak{H}.$$

Writing  $\Theta(Q) = \sum_N a_N q^N$ , we see that the coefficient  $a_N$  is the number of integral solutions of the equation  $Q(x,y) = N$ . Hence  $\Theta(Q)$  encodes all positive integers that can be represented by  $Q$ , along with the multiplicity of each representation. If  $Q$  and  $Q'$  are related by the  $\Gamma$ -action, then  $\Theta(Q) = \Theta(Q')$ ; in particular  $Q$  and  $Q'$  represent the same set of integers, each with the same multiplicity.

Now given  $Q$ , we can use the  $\Gamma$ -action to find a form  $Q'$  equivalent to  $Q$  and such that the corresponding point  $z(Q') \in \mathfrak{H}$  lies in the domain  $\overline{\Omega}$ . Moreover  $Q'$  is uniquely determined if  $z(Q') \in \Omega$ . The quadratic form  $Q'$  is called *reduced*, since its coefficients  $A, B, C$  are small.

3.1.3. Sometimes in reduction theory one considers weaker conditions on  $\Omega$  than those above. For example, sometimes one replaces item (2) in §3.1.1 with the requirement that  $\{\gamma \mid \gamma\Omega \cap \Omega \neq \emptyset\}$  is *finite*. The middle of Figure 3.1 shows an example; this set is three times as large as the exact domain on the left of Figure 3.1, and is stabilized by a subgroup of  $\Gamma$  of order 6. Although this domain is larger than a true fundamental domain, it has the nice property that  $\gamma\Omega \cap \Omega$  is either empty or equals  $\Omega$ . One may even drop this condition, and merely require that  $\{\gamma \mid \gamma\Omega \cap \Omega \neq \emptyset\}$  is finite and that that  $\Omega$  has an easily described form. This is useful in the theory of automorphic forms, when one wants good control over certain analytic objects related to  $\Omega$  without fretting too much about the exact geometry of  $\Omega$ . The right of Figure 3.1 shows an example of this kind of domain, a *Siegel set* in  $\mathfrak{H}$  [8, I.1].

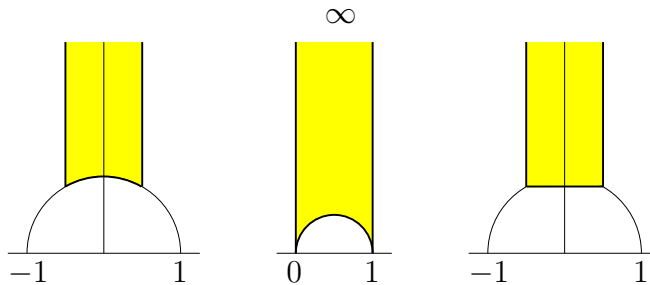


FIGURE 3.1. Various notions of reduction domains for  $\mathrm{SL}_2(\mathbb{Z})$

### 3.2. Reduction theory and cohomology.

3.2.1. There is another application of reduction theory that is less familiar: computation of the cohomology  $H^*(\Gamma; M)$ . Explicit knowledge of a precise reduction domain  $\Omega$  allows one to chop  $X$  up into subsets. These subsets pass to subsets of the quotient  $\Gamma \backslash X$ , and even  $\Gamma' \backslash X$  where  $\Gamma' \subset \Gamma$  is a finite-index subgroup. These subsets can be used with standard methods of combinatorial topology to build chain complexes to compute the cohomology of  $\Gamma$  and its finite-index subgroups.

3.2.2. We will make this discussion more precise in a moment, but the general idea can already be seen for  $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ . Let  $N \geq 1$ , and let  $\Gamma(N) \subset \mathrm{SL}_2(\mathbb{Z})$  be the *principal congruence subgroup* of matrices congruent to the identity modulo  $N$ .<sup>(2)</sup> For  $N \geq 2$  the group  $\Gamma(N)$  is torsion-free. The quotients  $Y(N) = \Gamma(N) \backslash \mathfrak{H}$  are punctured topological surfaces with genus roughly  $N^3/24$ .

Let  $\Omega$  be the ideal triangle in  $\mathfrak{H}$  with vertices at  $0, 1, \infty$ , as in the middle of Figure 3.1. Then the  $\mathrm{SL}_2(\mathbb{Z})$ -translates of  $\Omega$  pass to a “punctured triangulation of  $Y(N)$ ,” by which we mean the following. The surface  $Y(N)$  sits naturally inside a compact surface  $X(N)$  obtained by filling the punctures with points, and there is a triangulation of  $X(N)$  such that the vertices are the punctures. For example, if  $N = 3$  (respectively, 4, 5) then the genus of  $Y(N)$  is 0, and  $Y(N)$  equipped with this structure is isomorphic to a tetrahedron (resp., octahedron, icosahedron) with punctures at its vertices.

3.2.3. Here is how this structure can be used to compute  $H^*(\Gamma(N); M)$ . Consider the tessellation of  $\mathfrak{H}$  given by the  $\mathrm{SL}_2(\mathbb{Z})$ -translates of  $\Omega$ , where  $\Omega$  is as in §3.2.2. There is a regular trivalent graph  $W$  embedded in  $\mathfrak{H}$  dual to this tessellation (Figure 3.2). In other words,  $W$  has a vertex for each triangle in the tessellation, and two vertices are connected by an edge if and only if the corresponding triangles meet along an edge.

Modulo  $\Gamma(N)$ , the graph  $W$  is finite. For example, for  $N = 3$  (respectively, 4, 5), the quotient  $\Gamma(N) \backslash W$  is isomorphic to the 1-skeleton of the tetrahedron (resp., cube, dodecahedron). In fact,  $\Gamma(N) \backslash W$  is naturally a deformation retract of  $\Gamma(N) \backslash \mathfrak{H}$ : the retraction is given by pinching  $Y(N)$  along its punctures to enlarge them. Thus the cohomology of  $\Gamma(N) \backslash W$  equals that of  $Y(N)$ . Since  $\Gamma(N) \backslash W$  is a simplicial complex, it can be used with standard techniques of combinatorial topology to explicitly compute  $H^*(\Gamma(N); M)$ . Similar considerations apply to any finite-index subgroup  $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ .

3.2.4. The construction of  $W$  reveals more about the cohomology of  $\Gamma$ . The dimension of  $\mathfrak{H}$  is 2, so a priori  $H^*(\Gamma \backslash \mathfrak{H}; \mathcal{M})$  can be nonzero in degrees 0, 1, 2. But since  $\Gamma \backslash W$  is a graph, its cohomology vanishes in degree 2. This is a special case of a theorem of Borel–Serre, which implies that for any arithmetic group  $\Gamma \subset \mathbf{G}(\mathbb{Q})$  with

---

<sup>2</sup>A *congruence subgroup* of  $\mathrm{SL}_2(\mathbb{Z})$  is one that contains  $\Gamma(N)$  as a subgroup for some  $N$ . Hence the groups  $\Gamma_0(N)$  from §2.6.2 are congruence subgroups.

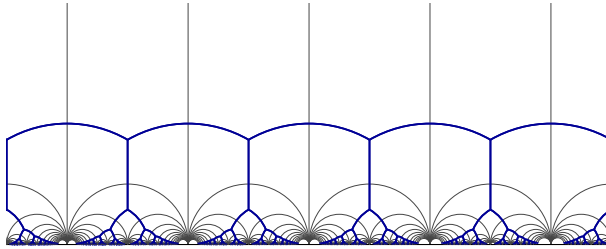


FIGURE 3.2. The graph  $W$  inside the tessellation by  $\Gamma$ -translates of  $\Omega$

$M$  as in §2.6.4, we have  $H^k(\Gamma; M) = 0$  if  $k > \nu := \dim X - q$ , where  $q$  is the  $\mathbb{Q}$ -rank of  $\mathbf{G}$ . Thus for example if  $\Gamma \subset \mathrm{SL}_n(\mathbb{Z})$ , we have  $\nu = n(n-1)/2$ ; for  $\Gamma \subset \mathrm{Sp}_{2n}(\mathbb{Z})$  we have  $\nu = n^2$ . The number  $\nu$  is called the *virtual cohomological dimension* of  $\Gamma$ . Since for  $\mathrm{SL}_2(\mathbb{Z})$  the number  $\nu$  is 1, the graph  $W$  is optimal from a computational point of view: it is exactly the right dimension to use in investigating the cohomology. It is also optimal from an aesthetic point of view, since  $W$  is beautifully embedded in  $\mathfrak{H}$ .

Hence we have the following natural problem: Given an arithmetic group  $\Gamma$ , find a subspace  $W$  of the associated symmetric space  $X$  such that the following hold:

- (1)  $W$  is a locally finite regular cell complex (a regular cell complex is a CW complex such that the attaching map from each closed cell into the complex is a homeomorphism onto its image [13]);
- (2)  $W$  admits a cellular  $\Gamma$ -action such that if  $\Gamma' \subseteq \Gamma$  is a torsion-free finite-index subgroup,  $\Gamma' \backslash W$  then is a regular cell complex;
- (3)  $W$  is a  $\Gamma$ -equivariant deformation retract of  $X$ ; and
- (4)  $W$  has dimension  $\nu(\Gamma)$ .

This is the problem solved in [38, 39] for  $\mathbf{G} = \mathrm{Sp}_4$ .

### 3.3. The cone of positive-definite quadratic forms.

3.3.1. To explain the results in [38, 39], we first need to understand how to find  $W$  for the special linear group  $\mathrm{SL}_n$ . Indeed, the constructions here play a key role for  $\mathrm{Sp}_4$  through the natural inclusion  $\mathrm{Sp}_4(\mathbb{R}) \subset \mathrm{SL}_4(\mathbb{R})$ .

One technique to construct  $W$  for  $\mathrm{SL}_n$  builds on work of Voronoï's theory of perfect quadratic forms [50]. We recall the definitions. Let  $V$  be the  $\mathbb{R}$ -vector space of all symmetric  $n \times n$  matrices with entries in  $\mathbb{R}$ , and let  $C \subset V$  be the subset of positive-definite matrices. The space  $C$  can be identified with the space of all real positive-definite quadratic forms in  $n$  variables: in coordinates, if  $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$  (column vector), then the matrix  $A \in C$  induces the quadratic form

$$x \longmapsto x^t A x.$$

It is well known that any positive-definite quadratic form arises in this way. The space  $C$  is a cone, in that it is preserved by homotheties: if  $x \in C$ , then  $\lambda x \in C$  for all  $\lambda \in \mathbb{R}_{>0}$ . Let  $D$  be the quotient of  $C$  by homotheties.

3.3.2. The case  $n = 2$  is illustrative. We can take coordinates on  $V \simeq \mathbb{R}^3$  by representing any matrix in  $V$  as

$$\begin{pmatrix} x & y \\ y & z \end{pmatrix}, \quad x, y, z \in \mathbb{R}.$$

The subset of singular matrices  $Q = \{xz - y^2 = 0\}$  is a quadric cone in  $V$  dividing the complement  $V \setminus Q$  into three connected components. The component containing the identity matrix is the cone  $C$ . The quotient  $D$  can be identified with an open 2-disk.

3.3.3. The group  $G = \mathrm{SL}_n(\mathbb{R})$  acts on  $C$  on the left by

$$(g, c) \longmapsto gcg^t.$$

This action commutes with that of the homotheties, and thus descends to a  $G$ -action on  $D$ . One can show that  $G$  acts transitively on  $D$  and that the stabilizer of the image of the identity matrix is  $K = \mathrm{SO}(n)$ . Hence we may identify  $D$  with the symmetric space  $X_{\mathrm{SL}} = \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}(n)$ . We will do this in the sequel, using the notation  $D$  when we want to emphasize the coordinates coming from the linear structure of  $C \subset V$ , and the notation  $X_{\mathrm{SL}}$  for the quotient  $G/K$ .

The  $G$ -action induces an action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $C$ . This is the unimodular change of variables action on quadratic forms as in §3.1.3. Under our identification of  $D$  with  $X_{\mathrm{SL}}$ , this is the usual action of  $\mathrm{SL}_n(\mathbb{Z})$  by left translation from §2.5.2.

### 3.4. The Voronoï polyhedron.

3.4.1. Recall that a point in  $\mathbb{Z}^n$  is said to be *primitive* if the greatest common divisor of its coordinates is 1. In particular, a primitive point is nonzero. Let  $\mathcal{P} \subset \mathbb{Z}^n$  be the set of primitive points. Any  $v \in \mathcal{P}$ , written as a column vector, determines a rank-one symmetric matrix  $q(v)$  in the closure  $\overline{C}$  via  $q(v) = vv^t$ . The *Voronoï polyhedron*  $\Pi$  is defined to be the closed convex hull in  $\overline{C}$  of the points  $q(v)$ , as  $v$  ranges over  $\mathcal{P}$ . Note that by construction,  $\mathrm{SL}_n(\mathbb{Z})$  acts on  $\Pi$ , since  $\mathrm{SL}_n(\mathbb{Z})$  preserves the set  $\{q(v)\}$  and acts linearly on  $V$ .

3.4.2. The polyhedron  $\Pi$  is quite complicated: it has infinitely many faces, and is not locally finite. However, one of Voronoï's great insights is that  $\Pi$  is actually not as complicated as it seems. To explain his insight, we need the notion of a perfect quadratic form.

For any  $A \in C$ , let  $\mu(A)$  be the minimum value attained by  $A$  on  $\mathcal{P}$ , and let  $M(A) \subset \mathcal{P}$  be the set on which  $A$  attains  $\mu(A)$ . Note that  $\mu(A)$  is positive and  $M(A)$  is finite since  $A$  is positive-definite. Then  $A$  is called *perfect* if it is recoverable

from the knowledge of the pair  $(\mu(A), M(A))$ . In other words, given  $(\mu(A), M(A))$ , we can write a system of linear equations

$$(3.4.1) \quad mZm^t = \mu(A), \quad m \in M(A),$$

where  $Z = (z_{ij})$  is a symmetric matrix of variables. Then  $A$  is perfect if and only if it is the unique solution to the system (3.4.1).

3.4.3. We can now summarize Voronoï's main results:

- (1) There are finitely many equivalence classes of perfect forms modulo the action of  $\mathrm{SL}_n(\mathbb{Z})$ . Voronoï even gave an explicit algorithm to determine all the perfect forms of a given dimension.
- (2) The facets of  $\Pi$ , in other words the codimension 1 faces, are in bijection with the rank  $n$  perfect quadratic forms. Under this correspondence the minimal vectors  $M(A)$  determine a facet  $F_A$  by taking the convex hull in  $\overline{C}$  of the finite point set  $\{q(m) \mid m \in M(A)\}$ . Hence there are finitely many faces of  $\Pi$  modulo  $\mathrm{SL}_n(\mathbb{Z})$ , and thus finitely many modulo any finite-index subgroup  $\Gamma$ .
- (3) Let  $\mathcal{V}$  be the set of cones over the faces of  $\Pi$ . Then  $\mathcal{V}$  is a *fan*, which means (i) if  $\sigma \in \mathcal{V}$ , then any face of  $\sigma$  is also in  $\mathcal{V}$ ; and (ii) if  $\sigma, \sigma' \in \mathcal{V}$ , then  $\sigma \cap \sigma'$  is a common face of each.<sup>(3)</sup>
- (4) The Voronoï fan  $\mathcal{V}$  provides a reduction theory for  $C$  in the following sense: any point  $x \in C$  is contained in a unique  $\sigma(x) \in \mathcal{V}$ , and the stabilizer subgroup  $\{\gamma \in \mathrm{SL}_n(\mathbb{Z}) \mid \gamma \cdot \sigma(x) = \sigma(x)\}$  is finite. Voronoï also gave an explicit algorithm to determine  $\sigma(x)$  given  $x$ , the *Voronoï reduction algorithm*.

The number of equivalence classes of perfect forms modulo the action of  $\mathrm{GL}_n(\mathbb{Z})$  grows rapidly with  $n$ . Voronoï computed the equivalence classes for  $n \leq 5$  [50]. Currently the largest  $n$  for which the number is known is  $n = 8$ : Dutour–Schurmann–Vallentin recently showed that there are 10916 equivalence classes [15]. For a list of perfect forms up to  $n = 7$ , see [12].

### 3.5. The Voronoï decomposition and the retract.

3.5.1. Here is how the Voronoï fan  $\mathcal{V}$  can be used to construct higher-dimensional analogues of the tessellation in Figure 3.2. The idea is to use the cones in  $\mathcal{V}$  to chop the quotient  $D$  into pieces.

For any  $\sigma \in \mathcal{V}$ , let  $\sigma^\circ$  be the open cone obtained by taking the complement in  $\sigma$  of its proper faces. Then after quotienting by homotheties, the cones  $\{\sigma^\circ \cap C \mid \sigma \in \mathcal{V}\}$  pass to locally closed subsets of  $D$ . Note that  $\sigma^\circ \cap C$  may be empty. If  $\sigma^\circ \cap C \neq \emptyset$ , and  $c$  is the image of  $\sigma^\circ$  modulo homotheties, we say  $\sigma^\circ$  *induces*  $c$ . Note that each  $c$  is a cell, in other words is homeomorphic to an open ball, since it is the quotient of an open polyhedral cone by homotheties.

---

<sup>3</sup>Strictly speaking, Voronoï actually showed that every codimension 1 cone is contained in exactly two top dimensional cones.

3.5.2. Let  $\mathcal{C}$  be the set of cells  $c$  induced from the Voronoï cones. Clearly  $C$  is the union of all cones  $\sigma$  that induce cells in  $\mathcal{C}$ . Since  $\mathcal{C}$  comes from the fan  $\mathcal{V}$ , the cells in  $\mathcal{C}$  have good incidence properties: the closure in  $D$  of any  $c \in \mathcal{C}$  can be written as a finite disjoint union of elements of  $\mathcal{C}$ . Moreover,  $\mathcal{C}$  is locally finite: by taking quotients of only the  $\sigma^\circ$  meeting  $C$ , we have eliminated the open cones lying in  $\overline{C}$ , and it is the latter cones that are responsible for the failure of local finiteness of  $\mathcal{V}$ .

We summarize these properties by saying that  $\mathcal{C}$  gives a *cellular decomposition* of  $D$ . Clearly  $\mathrm{SL}_n(\mathbb{Z})$  acts on  $\mathcal{C}$ , since  $\mathcal{C}$  is constructed using the fan  $\mathcal{V}$ . Thus we obtain a cellular decomposition of  $\Gamma \backslash D$  for any finite-index  $\Gamma \subset \mathrm{SL}_n(\mathbb{Z})$ . We call  $\mathcal{C}$  the *Voronoï decomposition* of  $D$ .

3.5.3. Now we explain how  $\mathcal{V}$  can be used to construct the cell complex  $W$ . The first step is to enlarge the cone  $C$  to a *partial Satake compactification*  $C^*$ . Let  $H$  be a hyperplane in  $V$ , and let  $\overline{C}$  be the closure of  $C$  in  $V$ . We say  $H$  is a *supporting hyperplane* of  $C$  if  $H$  is rational and  $H \cap C = \emptyset$  but  $H \cap \overline{C} \neq \emptyset$ . Since  $\overline{C}$  is convex, these conditions imply that  $\overline{C}$  lies entirely in one of the two closed half-spaces determined by  $H$ .

Given a rational supporting hyperplane  $H$  of  $C$ , let  $C' = \mathrm{Int}(H \cap \overline{C})$ , where  $\mathrm{Int}(\ )$  denotes the interior in the linear span. Then  $C'$  is called a *rational boundary component*; it is isomorphic to a smaller dimensional cone of positive definite quadratic forms. Let  $C^*$  be the union of  $C$  and all its proper rational boundary components. We can similarly form  $D^*$ , the union of  $D$  and the images of the rational boundary components modulo homotheties. One can topologize  $D$  such that  $\Gamma \backslash D^*$  is a compact Hausdorff space. In general  $\Gamma \backslash D^*$  is singular.

Again, we can consider these constructions for  $\mathrm{SL}_2$  and the principal congruence subgroup  $\Gamma(N)$ . In this case  $D^*$  is  $\mathfrak{H} \cup \mathbb{Q} \cup \{\infty\}$ . The quotient  $\Gamma(N) \backslash D^*$  is isomorphic to the surface  $X(N)$  obtained by filling in the punctures of  $Y(N)$ .

3.5.4. It turns out that formation of the Voronoï polyhedron  $\Pi$  is compatible with construction of the rational boundary components, which means the fan  $\mathcal{V}$  actually lies in  $C^*$ . Let  $B(\mathcal{V})$  be the first barycentric subdivision of  $\mathcal{V}$ . Take the cones in  $B(\mathcal{V})$  that are *dual* to the cones inducing cells in  $\mathcal{C}$ . These cones are contained entirely in  $C$ , and project to a collection of cells in  $D$ . By taking unions of these cells, we build  $W$ . This strategy was devised by Ash in [4], and used by him to construct retracts for a large class of arithmetic groups.

Figure 3.3 shows the process in  $D$  for  $\mathrm{SL}_2$ . On the left we see part of the Voronoï tessellation in  $D^*$ , displayed in the linear coordinates on  $D$ . The gray discs correspond to the points in  $D^* \setminus D$ . The middle shows the barycentric subdivision of this decomposition. The right shows  $W$ , in heavy black lines. Note that edges from the centers of the triangles to the boundary components do not appear in  $W$ , since they are not dual to cells in the original tessellation. Also note that the cells in  $W$  are

formed by taking unions of cells in the barycentric subdivision. For more about  $W$ , as well as extensions of  $W$  to other settings, we refer to [4, 5, 40, 49, 51, 52].

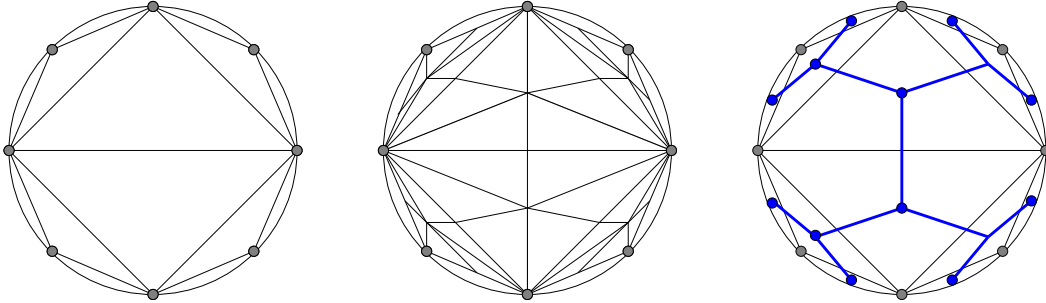


FIGURE 3.3. Forming  $W$  by taking the dual

### 3.6. The symplectic cell decomposition.

3.6.1. Finally we come to the symplectic case, and to the results of [38, 39]. Let  $\mathbf{G} = \mathrm{Sp}_{2n}$ ,  $G = \mathrm{Sp}_{2n}(\mathbb{R})$ , and  $\Gamma = \mathrm{Sp}_{2n}(\mathbb{Z})$ . Let  $X_{\mathrm{Sp}}$  be the symplectic symmetric space  $\mathrm{Sp}_{2n}/\mathrm{U}(n)$ , a real manifold of dimension  $n^2 + n$ . How can we construct  $W$ ? Unfortunately, there is no analogue of the cone  $C$  for  $X_{\mathrm{Sp}}$ : the Siegel upper halfspace has no hidden linear structure.

The main idea of [39] is extremely simple. There is an embedding  $\mathrm{Sp}_{2n}(\mathbb{R}) \subset \mathrm{SL}_{2n}(\mathbb{R})$  that induces an embedding  $\iota: X_{\mathrm{Sp}} \rightarrow X_{\mathrm{SL}}$ . In terms of the linear coordinates on  $D \simeq X_{\mathrm{SL}}$ , the image of  $X_{\mathrm{Sp}}$  is cut out by quadratic equations. In  $X_{\mathrm{SL}}$  we have the Voronoï cells  $\mathcal{C}$  that provide a cellular decomposition, and we can consider all possible intersections  $\{c \cap \iota(X_{\mathrm{Sp}}) \mid c \in \mathcal{C}\}$ . These provide candidates for the cells in a cellular decomposition  $\mathcal{C}_{\mathrm{Sp}}$  of  $X_{\mathrm{Sp}}$ .

However, there is a potential pitfall with this idea: how do we know that the intersections  $c \cap \iota(X_{\mathrm{Sp}})$  are actually *cells*? If each  $c$  met the image of  $X_{\mathrm{Sp}}$  transversely, then we would know by a general argument that the intersections would be cells. Unfortunately, this is not the case. It seems that sometimes the Voronoï cells meet the Siegel space transversely, and sometimes not. Thus we have no way of knowing a priori that the sets  $c \cap \iota(X_{\mathrm{Sp}})$  are actually nice.

3.6.2. Because of this difficulty, the authors decide to study one special case:  $\mathrm{Sp}_4$ . There are several reasons why this is a good choice. First of all, the Voronoï complex is only known in full detail for  $n \leq 5$ , so this is the only symplectic case other than  $\mathrm{Sp}_2 \simeq \mathrm{SL}_2$  that can be studied explicitly. It is also the simplest example of a nonlinear symmetric space of  $\mathbb{Q}$ -rank  $> 1$ . This is significant, since to construct the symplectic retract  $W_{\mathrm{Sp}}$  the authors plan to use a Satake compactification as in §3.5.3, and with such compactifications there are significant differences in passing

from  $\mathbb{Q}$ -rank 1 to  $\mathbb{Q}$ -rank  $> 1$ . Finally, arithmetic quotients of  $\mathrm{Sp}_4(\mathbb{R})$  have always played a special role in the literature. For one, they are moduli spaces of abelian surfaces with extra structure, and hence give interesting yet tractable examples of Shimura varieties other than the modular curves. Also the associated automorphic forms have long been of interest in arithmetic, since they are the first examples of Siegel modular forms that are not elliptic.

Thus to study  $\mathcal{C}_{\mathrm{Sp}}$  the authors resort to explicit computations in coordinates. They find that indeed the intersections  $c \cap \iota(X_{\mathrm{Sp}})$  are always cells, and that the collection  $\mathcal{C}_{\mathrm{Sp}}$  provides a cell decomposition of  $X_{\mathrm{Sp}}$ .

### 3.7. The symplectic retract.

3.7.1. After forming the cell decomposition  $\mathcal{C}_{\mathrm{Sp}}$ , McConnell and MacPherson construct the symplectic retract  $W_{\mathrm{Sp}}$  by following the strategy indicated in §3.5.3. They enlarge  $X_{\mathrm{Sp}}$  to a Satake compactification  $X_{\mathrm{Sp}}^*$ , and show that the decomposition  $\mathcal{C}_{\mathrm{Sp}}^*$  extends to  $X_{\mathrm{Sp}}^*$  in such a way that the compact quotient  $\Gamma' \backslash \mathcal{C}_{\mathrm{Sp}}^*$  is a regular cell complex for any torsion-free finite-index  $\Gamma' \subset \mathrm{Sp}_4(\mathbb{Z})$ . Then they define  $W_{\mathrm{Sp}}$  to be the Poincaré dual complex to  $\mathcal{C}_{\mathrm{Sp}}$  in the first barycentric subdivision of  $\mathcal{C}_{\mathrm{Sp}}^*$ , as indicated in Figure 3.3.

3.7.2. The difficulty in carrying out these arguments is that they must first verify that  $\mathcal{C}_{\mathrm{Sp}}^*$  is a regular cell complex. Again this doesn't follow from any general principles; they really need to show that the closure of each cell in  $\mathcal{C}_{\mathrm{Sp}}^*$  is a closed ball. Various bad things could happen when taking the closures. For example, the boundaries of the closures could be non-simply-connected homology spheres instead of true spheres.

To verify the regularity, more explicit computations are needed. They show that the boundaries of the closures are spheres by constructing explicit shellings. Roughly speaking, a shelling of a simplicial complex is a total ordering of its maximal faces satisfying certain properties that guarantee that the full complex is assembled from the maximal faces nicely. A result from combinatorics [14] states that if a finite  $n$ -dimensional simplicial complex  $\Delta$  is shellable and any  $(n-1)$ -simplex is contained in exactly two  $n$ -simplices, then  $\Delta$  is homeomorphic to an  $n$ -sphere. Using a computer, McConnell and MacPherson construct shellings of the barycentric subdivisions of the boundaries of the closures of cells with a computer; the largest complex they shelled was a simplicial  $S^5$  with 23232 faces.

3.7.3. We conclude this section by describing the geometry of the 4-dimensional cell complex  $W_{\mathrm{Sp}}$ .<sup>(4)</sup> The group  $\mathrm{Sp}_4(\mathbb{Z})$  acts cellularly on  $W_{\mathrm{Sp}}$ , and in the following we say two cells have the same *type* if they lie in the same  $\mathrm{Sp}_4(\mathbb{Z})$ -orbit.

---

<sup>4</sup>The images in Figures 3.4–3.7 were produced by the author in collaboration with Mark McConnell, as a birthday present to Bob MacPherson. Happy Birthday, Bob!



There is one type of 4-cell. Fix a 4-cell and let  $P$  be its closure. Then  $P$  can be realized as a cellular ball with 40 facets, 180 two-faces, 216 edges, and 76 vertices.

There are three types of 3-cells. In [38] their closures are called the *crystal*, the *vertebra*, and the *pyramid* (Figures 3.4 and 3.5). The pyramid has 4 triangle faces and 1 square face. The crystal has 12 square and 12 triangle faces. The vertebra has 2 hexagon, 12 square, and 12 triangle faces.

There are two types of 1-cells and two types of 0-cells modulo  $\mathrm{Sp}_4(\mathbb{Z})$ .

Now we focus on the closure  $P$  of a 4-cell, which contains 4 crystals, 4 vertebrae, and 32 pyramids. Each crystal (respectively, vertebra and pyramid) lies in the boundary of 3 (resp., 3 and 4) 4-cells. Hence  $P$  meets the closures of 112 other 4-cells.

How is  $P$  assembled together from its 3-faces? Unfortunately  $P$  has so many edges that it's difficult to draw. The structure of  $P$ , on the other hand, is relatively easy to describe.

Begin by connecting 4 vertebrae together in pairs along their hexagonal faces to form a cellular solid torus. This chain of vertebrae has 72 vertices, and thus accounts for all but 4 of the vertices of  $P$ .

Embed the chain of vertebrae in the boundary  $\partial P \simeq S^3$  as a solid torus  $T$ , and let  $C$  be an unknotted circle disjoint from  $T$  but nontrivially linking  $T$  with linking number 1. For example, if we identify  $S^3$  with the polydisc  $\{(x, y) \in \mathbb{C}^2 \mid |x| \leq 1, |y| \leq 1\}$ , then we can take  $T = \{|x| = 1, |y| \leq 1\}$  and  $C = \{x = 0, |y| = 1\}$ . The remaining 4 vertices of  $P$  can be taken to lie along the circle  $C$ . If we place 4 vertices along  $C$ , then any two adjacent vertices  $a, a'$ , together with certain vertices in the vertebrae, determine a crystal.

More precisely, each vertebra meets every other crystal in  $P$ , meeting each in three 2-faces (2 squares and a triangle). This is indicated in Figure 6(a); the shaded 2-faces are those that meet the crystals. Each triple of shaded 2-faces meets one of the crystals, and as we move around the chain  $T$  adjacent sets of triples meet the same crystal. It follows that each crystal meets every vertebra. This is indicated in Figure 6(b) by the triples of shaded 2-faces (note that in this figure, one vertex of the crystal appears at infinity in the Schlegel diagram).

Figure 6(b) also shows which two of the vertices of a crystal don't lie in any vertebrae; in the figure these vertices appear connected to the rest with light-shaded edges. Hence to build a crystal in  $P$  one takes adjacent pairs  $a, a'$  of vertices along  $C$  and uses each of them with 8 vertices of the chain  $T$  to flesh out a crystal.

This shows how to find all the crystals and vertebrae in  $P$ . The pyramids simply plug up the gaps to fill out all of  $\partial P$ .

Figure 3.7 shows the 1-skeleton of  $\partial P$ . The torus  $T$  is the central block of dark vertices and edges; the circle  $C$  is represented by a vertical line (i.e. we've used stereographic projection to identify  $S^3 \setminus \{\text{pt}\}$  with  $\mathbb{R}^3$ ), and the light vertices and edges are those arising from the "fleshing out" construction described above.

For more about the structure of  $W_{\text{Sp}}$ , as well as a beautiful way to index its cells in terms of configurations of points and lines in  $\mathbb{P}^3$ , we refer to [38].

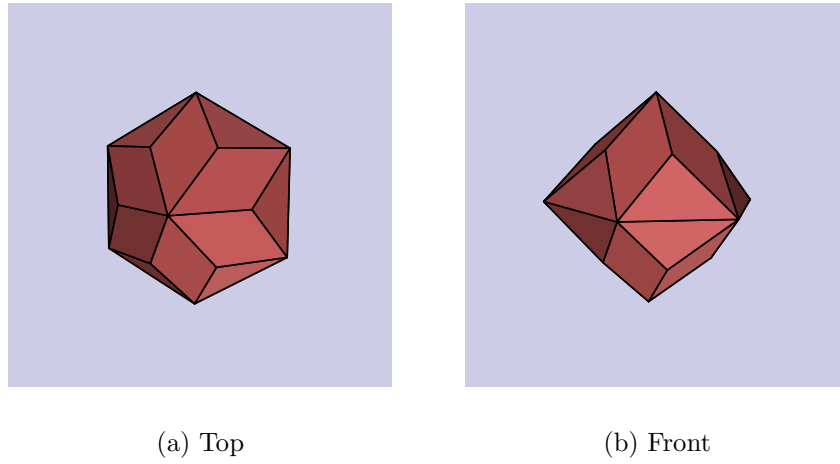


FIGURE 3.4. Two views of the crystal

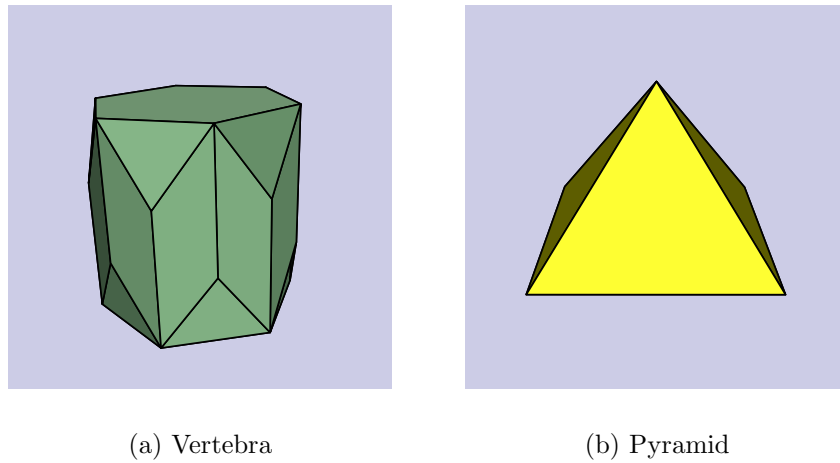


FIGURE 3.5. The vertebra and the pyramid

## 4. COMPACTIFICATIONS OF LOCALLY SYMMETRIC SPACES

### 4.1. Introduction.

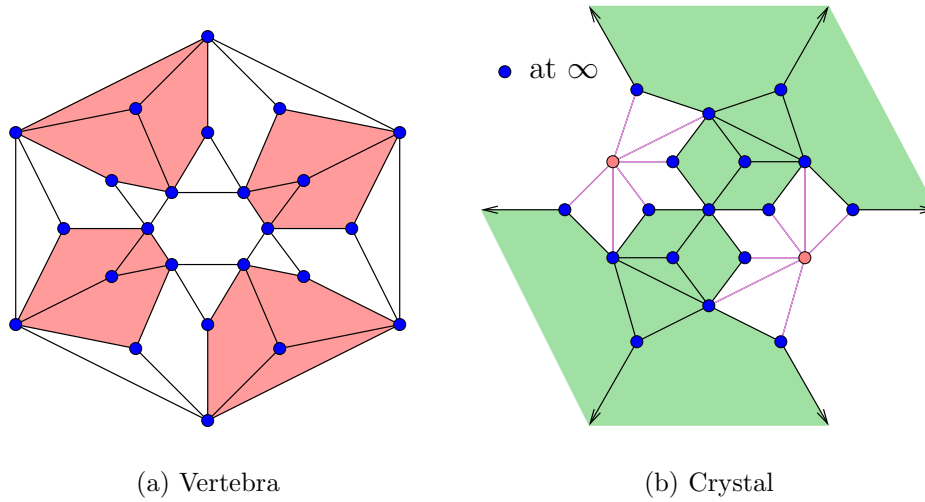


FIGURE 3.6. Schlegel diagrams for the vertebra and the crystal. A colored 2-face of the vertebra (respectively crystal) also meets another crystal (resp., vertebra) in  $\partial P$ . The light vertices and edges in the crystal are those not contained in any vertebra (cf. Figure 3.7).

4.1.1. We return to the general setting. Let  $\mathbf{G}$  be a connected semisimple  $\mathbb{Q}$ -group with group of real points  $G$ , and let  $K$  be a maximal compact subgroup of  $G$ . Let  $X = G/K$  be the associated symmetric space, and let  $\Gamma \subset \mathbf{G}(\mathbb{Q})$  be a torsion-free arithmetic group. Then as we have discussed, the quotient  $\Gamma \backslash X$  is a smooth manifold.

Suppose the  $\mathbb{Q}$ -rank of  $\mathbf{G}$  is positive. Then the quotient  $\Gamma \backslash X$  is noncompact. The basic problem we now address is the following: *Can one find a good compactification of  $\Gamma \backslash X$ ?*

4.1.2. This problem has a long history, and there are many beautiful and ingenious solutions. Before we talk about some of them, it's useful say a few words about why this question is worth asking. Why do we care that  $\Gamma \backslash X$  is noncompact?

Indeed, for many applications we don't care. The space  $X$  carries a natural  $G$ -invariant complete Riemannian metric that induces a  $G$ -invariant volume form. It turns out that even though  $\Gamma \backslash X$  is noncompact, its volume is finite (one says that  $\Gamma$  is *cofinite*). For many applications this is sufficient; the noncompactness of  $\Gamma \backslash X$  doesn't matter.

4.1.3. Nevertheless, there are reasons why one wants to compactify  $\Gamma \backslash X$ :

- (1) There are many tools from algebraic and differential topology (for example, Lefschetz fixed point theorem, Morse theory) that only work on compact

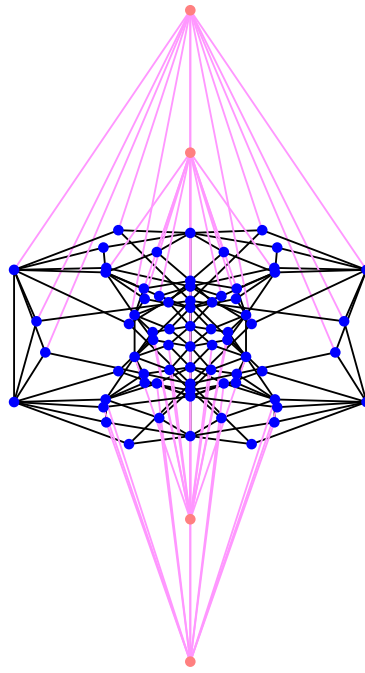


FIGURE 3.7. The 1-skeleton of the closure of a 4-cell in  $W_{\text{Sp}}$ . The dark edges lie in the chain of 4 vertebrae, which one should picture as being on a torus lying flat on its side. The light vertices are the vertices of the 4 crystals not already appearing in the vertebrae; in  $S^3$  they lie on a circle passing through the hole of the torus. If one labels these vertices  $0, 1, 2, 3$  from top to bottom, with labels taken mod 4, then the pair  $(i, i + 1)$  lies in a single crystal (cf. the right of Figure 3.6)

spaces. Since the cohomology of  $\Gamma \backslash X$  is important in number theory, we would like to avail ourselves of such tools to study it.

- (2) For certain groups  $\mathbf{G}$  the spaces  $\Gamma \backslash X$  can be interpreted as the complex points of a moduli space, for example for  $\mathbf{G} = \text{Sp}_{2n}$ . Here the space  $\Gamma \backslash X$  parametrizes  $n$ -dimensional abelian varieties with additional structure determined by  $\Gamma$ .

When  $\Gamma \backslash X$  is a moduli space, the points at infinity of a compactification often correspond to degenerations of the objects parametrized by  $\Gamma \backslash X$ . The higher the  $\mathbb{Q}$ -rank, the more intricate these degenerations are. One is interested in understanding these possible degenerations and their interrelationships.

- (3) Even if  $\Gamma \backslash X$  is not a moduli space, it has additional structure coming from the group theory of  $\mathbf{G}$ , structure that affects computations on  $\Gamma \backslash X$  in subtle ways. Perhaps the most definitive example of this phenomenon is Langlands's pièce de résistance, the determination of the spectral decomposition of  $L^2(\Gamma \backslash G)$  [33, 41]. Other manifestations of this principle can be seen in Franke's theorem (§2.6.4), Arthur's trace formula [1], and the computation of the spectrum of the Laplace operator on  $L^2(\Gamma \backslash X)$  (§4.7.2). This group-theoretic structure also arises in the construction of compactifications of  $\Gamma \backslash X$ . Hence it is natural to understand such computations in the fuller context of different compactifications of arithmetic quotients of  $X$ .

## 4.2. Examples of compactifications.

4.2.1. We hope the reader is convinced that compactifying  $\Gamma \backslash X$  is profitable. Because of the wealth of different applications of these compactifications, there is a whole zoo of them in the literature. Excellent overviews can be found in [9, 18, 28]. Here we content ourselves with a brief synopsis of some of the most useful. To simplify notation, we write  $Y = \Gamma \backslash X$ .

4.2.2. *Borel–Serre compactification  $Y^{BS}$  [10].* From a topological perspective, this compactification is the simplest: it is differentially a manifold with corners, by which we mean a Hausdorff space smoothly modeled on the generalized halfspaces  $\mathbb{R}^k \times (\mathbb{R}_{\geq 0})^{n-k}$ ,  $k = 0, \dots, n$ . Topologically  $Y^{BS}$  is just a manifold with boundary. The boundary  $\partial Y^{BS} := Y^{BS} \setminus Y$  is assembled from certain fiber bundles over locally symmetric spaces of lower rank (§4.6.2)

4.2.3. *Reductive Borel–Serre compactification  $Y^{RBS}$  [19, 53].* In contrast to the Borel–Serre,  $Y^{RBS}$  is usually singular. It is obtained as a quotient of  $Y^{BS}$ ; one collapses the fibers in the bundles above to points. Hence the boundary  $\partial Y^{RBS}$  is glued together from locally symmetric spaces of lower rank (§4.6.2)

Why should one ever use  $Y^{RBS}$  instead of  $Y^{BS}$ ? After all,  $Y^{BS}$  is nearly a manifold itself, and  $Y^{RBS}$  is quite singular. There are two answers. One is that the singularities of  $Y^{RBS}$  are intricate yet manageable, so little serviceability is lost in passing from  $Y^{BS}$  to  $Y^{RBS}$ . This feature plays an important role in the topological trace formula [27, §1.1.4]. The other is that  $Y^{BS}$  is deficient from a differential-geometric perspective: the complete metric on  $Y$  becomes degenerate on  $Y^{BS}$ . This is not so on  $Y^{RBS}$ ; the metric on  $Y$  extends to be nondegenerate on  $Y^{RBS}$ .

4.2.4. *Satake compactifications  $Y_\tau^{Sat}$  [47, 48].* Here there is not one compactification, but rather a whole collection of them. The construction is very similar to what we did in §3.5.3 in creating  $C^*$  from  $C$ , except that now we don't have a linear model for  $X$ . One begins by choosing an irreducible locally faithful representation  $\tau: G \rightarrow \mathrm{GL}(V)$  on a finite dimensional real vector space  $V$ . Not just any such  $\tau$  will work:  $\tau$  must

be *geometrically rational* (cf. [11]). Let  $\mathbb{P}\mathrm{Sym}^2 V$  be the symmetric square of  $V$  modulo homotheties. Then one embeds the global symmetric space  $X$  in  $\mathbb{P}\mathrm{Sym}^2(V)$  by  $g \mapsto \tau(g)^t \tau(g)$  mod homotheties. One then takes the closure  $\overline{X}$  of the image of  $X$  in  $\mathbb{P}\mathrm{Sym}^2(V)$  and identifies a certain subset of rational boundary components  $X_\tau^{\mathrm{Sat}} \subset \overline{X}$ .<sup>5</sup> After appropriately topologizing  $X_\tau^{\mathrm{Sat}}$ , the quotient  $Y_\tau^{\mathrm{Sat}} = \Gamma \backslash X_\tau^{\mathrm{Sat}}$  is compact and Hausdorff. This compactification is also usually singular, in general even more singular than the reductive Borel–Serre  $Y^{\mathrm{RBS}}$ .

4.2.5. *Baily–Borel compactification  $Y^{\mathrm{BB}}$  [6]*. This compactification, and the *toroidal compactifications* that follow, are defined for *Hermitian symmetric spaces*  $X$ . This means that  $X$  carries a  $G$ -invariant complex structure; accordingly the quotient  $Y$  does as well. Such  $X$  may be realized as a bounded domain in a complex vector space. One takes the closure  $\overline{X}$  of  $X$  there and again identifies a certain subset  $X^{\mathrm{BB}}$  of rational boundary components. Again after defining an appropriate topology, the quotient  $Y^{\mathrm{BB}} = \Gamma \backslash X^{\mathrm{BB}}$  is compact and Hausdorff, and in fact is a normal analytic space. After further investigation of the functions giving  $Y^{\mathrm{BB}}$  its ringed structure, namely certain Poincaré–Eisenstein series, one finds that  $Y^{\mathrm{BB}}$  is actually a projective variety. Topologically  $Y^{\mathrm{BB}}$  can be obtained as one of the Satake compactifications ([53], cf. [11, §7]). This compactification is usually extremely singular.

4.2.6. *Toroidal compactifications  $Y^\Sigma$  [3]*. These compactifications were originally constructed to resolve the singularities of the Baily–Borel  $Y^{\mathrm{BB}}$ . The construction depends on some extra (noncanonical) combinatorial data  $\Sigma$ ; this data is a collection of rational polyhedral fans in certain self-adjoint homogeneous cones attached to  $\mathbf{G}$ , fans that are geometrically very similar to the Voronoï fan  $\mathcal{V}$  in the cone of positive definite quadratic forms  $C$  (§3.4.2). Further assumptions on  $\Sigma$  guarantee that  $Y^\Sigma$  is a smooth projective variety with  $\partial Y^\Sigma$  a divisor with normal crossings. An overview of the construction for Siegel modular varieties can be found in [42].

4.2.7. The simplest example where there are differences between some of these compactifications is  $\mathbf{G} = \mathrm{SL}_2$ . Consider  $Y(N)$ , the modular curve  $\Gamma(N) \backslash \mathfrak{H}$  (§3.2.2). As described before,  $Y(N)$  is homeomorphic to a punctured topological surface. There are two obvious compactifications of  $Y(N)$ , one filling each puncture with a single point, and the other lining each puncture with a circle  $S^1$ . The first is the reductive Borel–Serre  $Y(N)^{\mathrm{RBS}}$ , as well as the Baily–Borel  $Y(N)^{\mathrm{BB}}$ . It is also the only possible Satake and only possible toroidal compactification of  $Y(N)$ . The second is the Borel–Serre  $Y(N)^{\mathrm{BS}}$ . In this case the quotient map  $\partial Y^{\mathrm{BS}} \rightarrow \partial Y^{\mathrm{RBS}}$  defined in §4.2.3 simply collapses each  $S^1 \subset \partial Y^{\mathrm{BS}}$  to a point.

---

<sup>5</sup> $X$  itself is also considered to be a boundary component, called the *improper* boundary component.

4.2.8. For a more revealing example, let  $\mathbf{G} = \mathbf{R}_{F/\mathbb{Q}} \mathrm{SL}_2$ , where  $F$  is real quadratic. Let  $\Gamma \subset \mathrm{SL}_2(\mathcal{O}_F)$  be torsion-free and of finite index. The locally symmetric space  $Y = \Gamma \backslash \mathfrak{H} \times \mathfrak{H}$  has real dimension 4. Each connected component of the boundary of the Borel–Serre  $\partial Y^{\mathrm{BS}}$  is a three manifold  $Z$  that is naturally the total space of a torus bundle over a circle:  $T^2 \hookrightarrow Z \rightarrow S^1$ .<sup>(6)</sup>

In the boundary of the reductive Borel–Serre  $\partial Y^{\mathrm{RBS}}$ , the tori in these bundles are collapsed to points; hence the boundary components are circles. This already shows that  $Y^{\mathrm{RBS}}$  is singular, since the link of any point in the boundary is a 2-torus.<sup>(7)</sup> The Baily–Borel  $Y^{\mathrm{BB}}$  is obtained by collapsing these circles to points; thus the link of any point in  $\partial Y^{\mathrm{BB}}$  is a 3-manifold  $Z$  as above. For more details, see [45].

A toroidal compactification resolving the cusp singularities of  $Y^{\mathrm{BB}}$  was first constructed by Hirzebruch [30]. Indeed, Hirzebruch’s technique was one of the principal motivations for the general theory of toroidal compactifications [3]. Since  $F$  is quadratic, there is a canonical minimal toroidal compactification. This is not true for higher degree  $F$ ; compactifications in this case were first constructed by Ehlers [16].

### 4.3. Building compactifications using geometry.

4.3.1. The compactifications above all use the structure theory of  $\mathbf{G}$ , the group underlying the symmetric space  $X$ , in an essential way. The basic questions addressed by Ji–MacPherson are

- What natural compactifications of  $Y = \Gamma \backslash X$  can be constructed using its intrinsic geometry? Here by intrinsic geometry we mean objects such as sets of geodesics on  $Y$ .
- What is the relationship of such compactifications with the standard ones from §4.2?

4.3.2. To illustrate the methods Ji–MacPherson have in mind, we consider  $M = \mathbb{R}^n$  equipped with the standard metric [31, §1.1]. How can we compactify  $M$ ?

We first need to know what the points at infinity  $M(\infty)$  should be. Let  $m \in M$  be a fixed point, and let  $\gamma: \mathbb{R}_{\geq 0} \rightarrow M$  be a geodesic ray with  $\gamma(0) = m$ . We say  $\gamma$  is *based* at  $m$ . Let  $d(\cdot, \cdot)$  be the distance function on  $M$  induced from the metric. Any sequence of points  $\{m_i\}_{i \geq 1} \subset \gamma$  with  $d(m, m_i) \rightarrow \infty$  should converge to a point on  $M(\infty)$ , since the  $m_i$  go arbitrarily far from  $m$ . Moreover, all such sequences along  $\gamma$  should tend to the same point at infinity.

<sup>6</sup>The different connected components of  $Y^{\mathrm{BS}}$  need not be homeomorphic.

<sup>7</sup>In a stratified space  $X \subset \mathbb{R}^n$ , the link  $L(p)$  of a point  $p$  in a stratum  $S$  is by definition  $L(p) = \partial B_\delta(p) \cap N \cap X$ , where (i)  $N$  is a submanifold through  $p$  meeting all strata of  $X$  transversely, and with  $\dim N + \dim S = n$ , and (ii)  $B_\delta(p)$  is a closed ball centered at  $p$  with radius  $0 < \delta \ll 1$ . It is known that the homeomorphism type of the  $L(p)$  is independent of the above choices for a wide class of stratified spaces, and that the normal slice  $B_\delta(p) \cap N \cap X$  is homeomorphic to the cone on  $L(p)$ . For this same class of spaces, the link is an invariant of the stratum  $S$ . For more details, we refer to [24, I.1.4].

Now consider two geodesic rays  $\gamma, \gamma'$  with distinct basepoints  $m \neq m'$ . If  $\gamma$  and  $\gamma'$  are not parallel, then for any two sequences  $\{m_i\}_{i \geq 1}$  and  $\{m'_i\}_{i \geq 1}$  with  $d(m, m_i), d(m', m'_i) \rightarrow \infty$ , we have  $d(m_i, m'_i) \rightarrow \infty$ . Hence these sequences should converge to different points at infinity. On the other hand if  $\gamma$  and  $\gamma'$  are parallel, we can find two such sequences with  $d(m_i, m'_i)$  bounded above. Hence it is reasonable to require that these sequences converge to the *same* point at infinity.

This leads to the following definition. As a set,  $M(\infty)$  is the set of geodesic rays in  $M$  modulo the equivalence relation of parallelism. For this example the topology on  $M(\infty)$  is clear: if the angle between two rays is small, the corresponding points on  $M(\infty)$  should be close. With this topology,  $M \cup M(\infty)$  is homeomorphic to the closed ball  $B^n$ , with  $M(\infty)$  homeomorphic to  $S^{n-1}$ .

4.3.3. Now the authors want to apply this idea to a locally symmetric space  $Y = \Gamma \backslash X$ . The geometry here is more intricate: there are geodesic rays that don't go cleanly off to infinity, but close up to form immersed loops. Even worse, there are geodesics that reenter a fixed compact set infinitely often. This is already visible when  $Y = \Gamma \backslash \mathfrak{H}$  is a modular curve. If  $\gamma \subset \mathfrak{H}$  is a geodesic ray tending to an ideal point  $\alpha \in \mathbb{R}$ , then the behavior of the image of  $\gamma$  in  $Y$  depends subtly on the basepoint  $\gamma(0)$  and the arithmetic nature of  $\alpha$ , in particular the continued fraction expansion of  $\alpha$ .

Because of this phenomenon, the authors define a *distance minimizing ray* (DM ray) to be a geodesic ray  $\gamma: \mathbb{R}_{\geq 0} \rightarrow Y$  that gives an isometric embedding of  $\mathbb{R}_{\geq 0}$  in  $Y$ . More generally, they introduce *eventually distance minimizing geodesics* (EDM geodesics); by definition  $\gamma: \mathbb{R} \rightarrow Y$  is EDM if there exists  $t_0 \gg 0$  such that  $\gamma$  restricted to  $\mathbb{R}_{\geq t_0}$  is DM. These will be the basic objects determining points on  $Y(\infty)$ .

Next they need an appropriate equivalence relation on geodesics. Since parallelism makes no sense in  $Y$ , they generalize the characterization of parallel rays in  $\mathbb{R}^n$  through distances as in §4.3.2. They define two DM rays  $\gamma, \gamma'$  to be *equivalent* if

$$\limsup_{t \rightarrow \infty} d(\gamma(t), \gamma'(t)) < \infty.$$

In other words, as  $t \rightarrow \infty$ , the distances between the corresponding points on  $\gamma, \gamma'$  remain bounded above.

Finally they define  $Y(\infty)$  as a set to be the set of DM rays modulo equivalence. After defining an appropriate topology on  $Y \cup Y(\infty)$ , they obtain a compact Hausdorff space such that each DM ray converges to the point on  $Y(\infty)$  corresponding to its equivalence class. The resulting space is called the *geodesic compactification* of  $Y$ , and is denoted  $Y^{\text{geo}}$ .

4.3.4. The basic idea of Ji–MacPherson's construction makes sense for any complete noncompact Riemannian manifold, not just a locally symmetric space, and in fact



the geodesic compactification appears in the literature in different guises. For example, the geodesic compactification of a Hadamard manifold is known as the *conic compactification* [7].<sup>(8)</sup>

For complete noncompact Riemannian manifolds  $M$ , one has a very general compactification technique due to Gromov [7]: one embeds  $M$  in a space of continuous functions using the distance function associated to the metric, and then takes the closure. Ji and MacPherson show that, if  $M$  is a locally symmetric space, the geodesic compactification coincides with Gromov's.

#### 4.4. More structure theory.

4.4.1. Our next goal is a description of the geometry of  $Y^{\text{geo}}$ , as well as the relationship of  $Y^{\text{geo}}$  to  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$ . This requires substantially more notation from the theory of algebraic groups (§§4.4.2–4.4.6). We give examples of most of the following in §4.4.7 for  $\mathbf{G} = \text{SL}_n$ ; the inexpert reader may wish to skip ahead.

4.4.2. Let  $\mathbf{S}$  be a maximal  $\mathbb{Q}$ -split torus of  $\mathbf{G}$ , with character lattice (respectively, dual lattice of one-parameter subgroups)  $X(\mathbf{S})_{\mathbb{Q}}$  (resp.,  $X^{\vee}(\mathbf{S})_{\mathbb{Q}}$ ). We denote by  $\langle \cdot, \cdot \rangle: X(\mathbf{S})_{\mathbb{Q}} \times X^{\vee}(\mathbf{S})_{\mathbb{Q}} \rightarrow \mathbb{Z}$  the natural unimodular pairing.

Let  $\mathfrak{g}$  be the Lie algebra of  $\mathbf{G}$ , and let  $\Phi = \Phi(\mathbf{G}, \mathbf{S})$  be the roots of  $\mathbf{S}$  in the adjoint action on  $\mathfrak{g}$ . Then  $\Phi$  is a root system in the vector space  $X(\mathbf{S})_{\mathbb{R}} := X(\mathbf{S})_{\mathbb{Q}} \otimes \mathbb{R}$ . This root system determines a hyperplane arrangement  $\{H_{\alpha} \mid \alpha \in \Phi\}$  in  $X^{\vee}(\mathbf{S})_{\mathbb{R}}$  by

$$H_{\alpha} = \{y \in X^{\vee}(\mathbf{S})_{\mathbb{R}} \mid \langle \alpha, y \rangle = 0\}.$$

The connected components of the complement of this arrangement are called *Weyl chambers*. The Weyl group  $W = W(\mathbf{G})$  acts on  $X^{\vee}(\mathbf{S})_{\mathbb{R}}$  by transitively permuting the chambers. If we fix a chamber  $C$ , we determine a subset  $\Phi^+ \subset \Phi$  of positive roots, and a subset  $\Delta \subset \Phi^+$  of simple roots.

4.4.3. A closed subgroup  $\mathbf{P} \subset \mathbf{G}$  is called *parabolic* if it contains a maximal connected solvable subgroup, and is called  *$\mathbb{Q}$ -parabolic* (or *rational parabolic*) if it is defined over  $\mathbb{Q}$ .

The *spherical Tits building*  $\mathcal{B} = \mathcal{B}(\mathbf{G})$  is the simplicial complex constructed as follows. Simplices of  $\mathcal{B}$  are in bijection with proper rational parabolic subgroups of  $\mathbf{G}$ . The vertices of  $\mathcal{B}$  correspond to the maximal parabolic subgroups; given a set of such  $\mathbf{P}_1, \dots, \mathbf{P}_k$ , if the intersection  $\mathbf{Q} = \mathbf{P}_1 \cap \dots \cap \mathbf{P}_k$  is a rational parabolic subgroup, then there is a simplex in  $\mathcal{B}$  corresponding to  $\mathbf{Q}$  with vertices corresponding to  $\mathbf{P}_1, \dots, \mathbf{P}_k$ . It is known that  $\mathcal{B}$  has pure dimension equal to  $q - 1$ , where  $q$  is the  $\mathbb{Q}$ -rank of  $\mathbf{G}$ .

The group of rational points  $\mathbf{G}(\mathbb{Q})$  acts on  $\mathcal{B}$  through conjugation of rational parabolic subgroups. If  $\Gamma \subset \mathbf{G}(\mathbb{Q})$  is an arithmetic subgroup, then  $\Gamma$  acts on  $\mathcal{B}$ , and

---

<sup>8</sup>A manifold is *Hadamard* if it is simply-connected, nonpositively curved, and complete [7]. For example,  $\mathfrak{H} \times \mathbb{R}$ ,  $\mathfrak{H} \times \mathfrak{H}$ , and  $\mathfrak{H}_3$  are all Hadamard.

the quotient  $\Gamma \backslash \mathcal{B}$  has simplices in bijection with the  $\Gamma$ -conjugacy classes of rational parabolic subgroups. It is known that  $\Gamma \backslash \mathcal{B}$  is always finite.

4.4.4. Let  $\mathbf{N}_{\mathbf{P}} \subset \mathbf{P}$  be the unipotent radical of  $\mathbf{P}$ . The quotient  $\mathbf{L}_{\mathbf{P}} := \mathbf{P}/\mathbf{N}_{\mathbf{P}}$  is called the *Levi quotient*. It is a reductive group, defined over  $\mathbb{Q}$  if  $\mathbf{P}$  is.

Let  $\mathbf{S}_{\mathbf{P}}$  be a maximal  $\mathbb{Q}$ -split torus in the center of  $\mathbf{L}_{\mathbf{P}}$ . Let  $A_{\mathbf{P}} = S_{\mathbf{P}}^0$ , the connected component of the identity in the group  $S_{\mathbf{P}}$  of real points  $\mathbf{S}_{\mathbf{P}}(\mathbb{R})$ . Put

$$\mathbf{M}_{\mathbf{P}} = \bigcap_{\alpha \in X(\mathbf{L}_{\mathbf{P}})_{\mathbb{Q}}} \ker \alpha^2.$$

The group  $\mathbf{M}_{\mathbf{P}}$  is semisimple, in general not connected.

There is a unique lift  $i: \mathbf{L}_{\mathbf{P}} \rightarrow \mathbf{P}$  compatible with our choice of maximal compact subgroup  $K \subset G$ . This induces lifts of  $\mathbf{S}_{\mathbf{P}}$  and  $\mathbf{M}_{\mathbf{P}}$  to  $\mathbf{P}$  that allow us to view the groups of real points  $A_{\mathbf{P}}$  and  $M_{\mathbf{P}}$  as subgroups of  $P$ . This leads to the *Langlands decomposition* of  $P$ :

$$P = N_{\mathbf{P}} A_{\mathbf{P}} M_{\mathbf{P}}.$$

One can show that the map  $N_{\mathbf{P}} \times A_{\mathbf{P}} \times M_{\mathbf{P}} \rightarrow P$  induced by multiplication is a diffeomorphism.

4.4.5. The Langlands decomposition leads to coordinates on the global symmetric space  $X = G/K$  as follows. The group  $P$  acts transitively on  $X$ . Any  $x \in X$  can be written as  $x = uamx_0$ , where  $u \in N_{\mathbf{P}}$ ,  $a \in A_{\mathbf{P}}$ ,  $m \in M_{\mathbf{P}}$  are uniquely determined, and where  $x_0 \in X$  is the basepoint determined by  $K$ . Let  $K_{\mathbf{P}} = M_{\mathbf{P}} \cap K$  and put  $X_{\mathbf{P}} = M_{\mathbf{P}}/K_{\mathbf{P}}$ . Then  $X_{\mathbf{P}}$  is the product of a global symmetric space of noncompact type with a possible Euclidean factor. The decomposition  $x = uamx_0$  induces a diffeomorphism

$$N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}} \longrightarrow X$$

by

$$(u, mK_{\mathbf{P}}, a) \longmapsto uamx_0.$$

4.4.6. Let  $\mathbf{P}$  be a rational parabolic subgroup, and let  $\mathfrak{a}_{\mathbf{P}}$ ,  $\mathfrak{n}_{\mathbf{P}}$  be the Lie algebras of  $A_{\mathbf{P}}$ ,  $N_{\mathbf{P}}$ . We denote by  $\Phi^+(\mathbf{P}, A_{\mathbf{P}})$  the positive roots of the adjoint action of  $\mathfrak{a}_{\mathbf{P}}$  on  $\mathfrak{n}_{\mathbf{P}}$ . Put

$$(4.4.1) \quad A_{\mathbf{P}}^+(\infty) = \{H \in \mathfrak{a}_{\mathbf{P}} \mid \alpha(H) > 0, \quad (H, H) = 1, \quad \alpha \in \Phi^+(\mathbf{P}, A_{\mathbf{P}})\},$$

where  $(\ , \ )$  is the Killing form on  $\mathfrak{a}_{\mathbf{P}}$ . Let  $\overline{A}_{\mathbf{P}}^+(\infty)$  be the closure of  $A_{\mathbf{P}}^+(\infty)$  obtained by replacing the conditions  $\alpha(H) > 0$  in (4.4.1) with  $\alpha(H) \geq 0$ . Then  $\overline{A}_{\mathbf{P}}^+(\infty)$  is homeomorphic to a closed simplex. Define a simplicial complex  $B(X)$  by

$$(4.4.2) \quad B(X) = \bigcup_{\mathbf{P}} \overline{A}_{\mathbf{P}}^+(\infty) / \sim,$$

where the union is taken over all proper rational parabolic subgroups, and we identify  $\overline{A}_{\mathbf{P}}^+(\infty)$  with a face of  $\overline{A}_{\mathbf{Q}}^+(\infty)$  if the two parabolic subgroups  $\mathbf{P}, \mathbf{Q}$  satisfy  $\mathbf{Q} \subset \mathbf{P}$ . Then  $B(X) \simeq \mathcal{B}$ , in other words the complex  $B(X)$  is a realization of the Tits building  $\mathcal{B}$ . If  $\Gamma$  is an arithmetic subgroup, we similarly define

$$(4.4.3) \quad B(\Gamma \backslash X) = \bigcup_{\mathbf{P}} \overline{A}_{\mathbf{P}}^+(\infty) / \sim,$$

where the union is now taken over all  $\Gamma$ -conjugacy classes of proper rational parabolic subgroups. We have  $B(\Gamma \backslash X) \simeq \Gamma \backslash \mathcal{B}$ .

4.4.7. We consider the example  $\mathbf{G} = \mathrm{SL}_n$ .

The maximal  $\mathbb{Q}$ -split torus  $\mathbf{S}$  is the subgroup of all diagonal matrices. The standard choice of positive roots is the set  $\Phi^+ = \{e_i - e_j \mid 1 \leq i < j \leq n\}$ , where  $\{e_i\}$  is the standard basis of  $\mathbb{R}^n$ . The simple roots  $\Delta \subset \Phi^+$  are the points  $\{e_i - e_{i+1} \mid i = 1, \dots, n-1\}$ , and the root lattice  $X(\mathbf{S})_{\mathbb{Q}}$  can be identified with  $\{(x_1, \dots, x_n) \in \mathbb{Z}^n \mid \sum x_i = 0\}$ .

Any proper  $\mathbb{Q}$ -parabolic subgroup is conjugate over  $\mathrm{SL}_n(\mathbb{Q})$  to a standard proper  $\mathbb{Q}$ -parabolic subgroup. The latter are indexed by ordered positive partitions  $\pi$  of  $n$  with at least 2 parts. Given such a partition  $\pi = (\pi_1, \dots, \pi_k)$ , the corresponding parabolic subgroup has real points

$$P = \left\{ \left( \begin{array}{ccc} P_1 & \cdots & * \\ & \ddots & \vdots \\ 0 & & P_k \end{array} \right) \mid P_i \in \mathrm{GL}_{\pi_i}(\mathbb{R}), \prod \det(P_i) = 1 \right\}.$$

The Langlands decomposition  $P = N_{\mathbf{P}} A_{\mathbf{P}} M_{\mathbf{P}}$  is given as follows:

- $M_{\mathbf{P}} \subset P$  is the subgroup of block diagonal matrices such that each block is an element of  $\mathrm{SL}_{\pi_i}^{\pm}(\mathbb{R})$ . Here the  $\pm$  means to take matrices with determinant  $\pm 1$ .
- $A_{\mathbf{P}} \subset P$  is the subgroup of block diagonal matrices such that each  $i$ th block has the form  $a_i I_{\pi_i}$ , where  $a_i > 0$  and  $I_{\pi_i}$  is the  $\pi_i \times \pi_i$  identity matrix.
- $N_{\mathbf{P}} \subset P$  is the subgroup such that each  $i$ th block equals  $I_{\pi_i}$ .

The parabolic subgroup  $\mathbf{P}_0$  corresponding to the partition  $(1, \dots, 1)$  is called the Borel subgroup. The Lie algebra  $\mathfrak{a}_{\mathbf{P}_0}$  of  $A_{\mathbf{P}_0}$  can be identified with  $X^{\vee}(\mathbf{S})_{\mathbb{R}}$ . There are  $n!$  Weyl chambers, each of which is an open simplicial cone of dimension  $n-1$ . The subset  $\overline{A}_{\mathbf{P}_0}^+(\infty)$  is the intersection of the closure of one of these cones with a sphere. For any other standard rational parabolic subgroup  $\mathbf{Q}$ , we have an inclusion  $\mathfrak{a}_{\mathbf{Q}} \subset \mathfrak{a}_{\mathbf{P}_0}$  such that  $\overline{A}_{\mathbf{Q}}^+(\infty)$  is identified with a proper face of  $\overline{A}_{\mathbf{P}_0}^+(\infty)$ .

Figure 4.1 shows the situation for  $\mathrm{SL}_4$ . The left shows  $\overline{A}_{\mathbf{P}_0}^+(\infty)$  as the dark spherical triangle topping off the simplicial cone. In this figure we have used the Killing form to identify  $\mathfrak{a}_{\mathbf{P}_0}$  with its dual, which allows us to view the simple roots  $\alpha_i$  alongside

the cone. Each simple root is orthogonal to a facet of the cone. We have also rescaled the form in the definition (4.4.1) to make the picture clearer.

There are 7 partitions  $\pi$ : 1111, 211, 121, 112, 31, 22, 13. Here we abbreviate  $(\pi_1, \dots, \pi_k)$  by  $\pi_1 \cdots \pi_k$ , and in what follows denote a parabolic subgroup by its partition and drop  $(\infty)$  from the notation. The right of Figure 4.1 shows how the faces of  $\overline{A}_{1111}^+ = \overline{A}_{\mathbf{P}_0}^+(\infty)$  are indexed by the partitions. (To go from the left figure to the right rotate the back of the dark triangle forward.) The edges of  $\overline{A}^+$  correspond to partitions with 3 parts; the bottom is  $\overline{A}_{121}^+$ , the left  $\overline{A}_{112}^+$ , and the right  $\overline{A}_{211}^+$ . The vertices correspond to partitions with 2 parts; clockwise from the top they are  $\overline{A}_{22}^+$ ,  $\overline{A}_{31}^+$ , and  $\overline{A}_{13}^+$ .

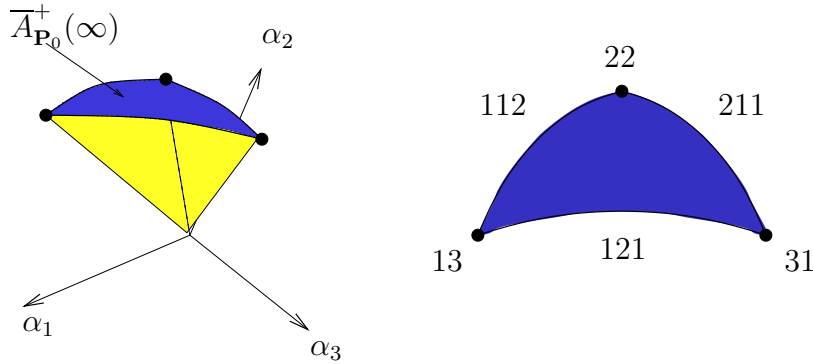


FIGURE 4.1. The region  $\overline{A}_{1111}^+ = \overline{A}_{\mathbf{P}_0}^+(\infty)$  for  $\mathrm{SL}_4$ , where  $\mathbf{P}_0$  is a Borel subgroup.

#### 4.5. The geodesic compactification.

4.5.1. We are now ready to discuss the geometry of  $Y^{\mathrm{geo}}$ . Choose a rational parabolic subgroup  $\mathbf{P}$ , and write  $X = N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}}$  as in §4.4.5. If we choose  $u \in N_{\mathbf{P}}$ ,  $z \in X_{\mathbf{P}}$ ,  $a \in A_{\mathbf{P}}$ , and  $H \in A_{\mathbf{P}}^+(\infty)$ , we get a geodesic  $\gamma$  on  $X$  by

$$(4.5.1) \quad \gamma(t) = (u, z, a \exp(tH)) \subset N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}}, \quad t \in \mathbb{R},$$

where  $\exp: \mathfrak{a}_{\mathbf{P}} \rightarrow A_{\mathbf{P}}$  is the exponential map. The authors show that the image of  $\gamma$  in  $Y = \Gamma \backslash X$  is an EDM geodesic, and that in fact *any* EDM geodesic on  $Y$  has the form (4.5.1) for appropriately chosen  $\mathbf{P}$ . Moreover, two geodesics  $\gamma, \gamma'$  of the form (4.5.1) project to the same geodesic in  $Y$ , up to reparametrization, if and only if  $H = H'$ ,  $\log a - \log a'$  is a multiple of  $H$ , and  $(u, z) = g(u', z')$  for some  $g \in \Gamma_{\mathbf{P}} := \Gamma \cap \mathbf{P}(\mathbb{Q})$ . Hence all EDM geodesics on  $Y$  have an especially simple form.

4.5.2. Next the authors investigate equivalence. It turns out that the equivalence class of the geodesic  $\gamma(t) = (u, z, a \exp(tH))$  depends only on the  $H$  component, and not  $u$ ,  $z$ , or  $a$ . This, together with the explicit realizations

$$B(X) \simeq \mathcal{B}, \quad B(\Gamma \backslash X) \simeq \Gamma \backslash \mathcal{B}$$

from §4.4.6, shows that the boundary  $\partial Y^{\text{geo}}$  is homeomorphic to  $\Gamma \backslash \mathcal{B}$ .

This result motivates Ji and MacPherson to define another natural compactification of  $Y$ : the *Tits compactification*  $Y^{\text{T}}$ . By definition, the (partial) Tits compactification of  $X$  is the union  $X \cup \mathcal{B}$ , appropriately topologized so that the quotient  $Y \cup (\Gamma \backslash \mathcal{B})$  is compact and Hausdorff. In  $Y^{\text{T}}$ , the image of the geodesic  $(u, z, a \exp(tH))$  converges to the point in  $\Gamma \backslash \mathcal{B}$  corresponding to  $H$  via the realization (4.4.3).

#### 4.6. The Borel–Serre and the reductive Borel–Serre compactifications.

4.6.1. The Tits compactification is a new compactification of  $Y$  constructed using the group theory of  $\mathbf{G}$ , and as such properly joins the list in §4.2. As we shall shortly see,  $Y^{\text{T}}$  is as far from  $Y^{\text{BS}}$  as possible: if the  $\mathbb{Q}$ -rank of  $\mathbf{G}$  is  $> 1$ , then the greatest common quotient of  $Y^{\text{T}}$  and  $Y^{\text{BS}}$  is the one-point compactification of  $Y$ ! With hindsight, it is clear that this must be the case, since the boundary of  $Y^{\text{T}} \simeq Y^{\text{geo}}$  is constructed using rays that converge to each other at infinity, and on  $Y^{\text{BS}}$  the metric degenerates. Hence equivalence classes of EDM geodesics cannot detect points in  $\partial Y^{\text{BS}}$ .

4.6.2. To understand the exact relationship between  $Y^{\text{geo}}$  and  $Y^{\text{BS}}$ ,  $Y^{\text{RBS}}$ , we must first say more about the construction of the latter spaces. We begin with the globally symmetric space  $X$  and construct partial compactifications  $X^{\text{BS}}$ ,  $X^{\text{RBS}}$  by gluing on boundary components for each proper rational parabolic subgroup. Given such a subgroup  $\mathbf{P}$ , write  $X = N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}}$ . We define boundary components  $e^{\text{BS}}(\mathbf{P})$ ,  $e^{\text{RBS}}(\mathbf{P})$  by

$$e^{\text{BS}}(\mathbf{P}) = N_{\mathbf{P}} \times X_{\mathbf{P}}, \quad e^{\text{RBS}}(\mathbf{P}) = X_{\mathbf{P}},$$

and set

$$(4.6.1) \quad X^{\text{BS}} = X \bigcup_{\mathbf{P}} e^{\text{BS}}(\mathbf{P}), \quad X^{\text{RBS}} = X \bigcup_{\mathbf{P}} e^{\text{RBS}}(\mathbf{P}),$$

where the unions are taken over all proper rational parabolic subgroups.

The topologies on (4.6.1) can be defined by specifying the convergence properties of sequences of points in  $X$  and in the boundary components. For instance, to converge to a point on  $e^{\text{BS}}(\mathbf{P})$  starting from inside  $X$ , take a sequence  $\{y_n\} = \{(u_n, z_n, \exp(H_n))\} \subset N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}}$ , and suppose that

- (1)  $u_n \rightarrow u_{\infty} \in N_{\mathbf{P}}$ ,
- (2)  $z_n \rightarrow z_{\infty} \in X_{\mathbf{P}}$ , and
- (3) for any  $\alpha \in \Phi^+(\mathbf{P}, A_{\mathbf{P}})$ , we have  $\alpha(H_n) \rightarrow \infty$  as  $n \rightarrow \infty$ .

Then in  $X^{\text{BS}}$ , the sequence  $y_n$  converges to  $(u_\infty, z_\infty) \in e^{\text{BS}}(\mathbf{P})$ . Similarly, in the reductive Borel–Serre  $X^{\text{RBS}}$ , the same sequence converges to the point  $z_\infty \in e^{\text{RBS}}(\mathbf{P})$ . Other sequences can be constructed that converge from one boundary component to a point in another. For a full list of all the sequences that need to be specified to produce the correct topologies on  $X^{\text{BS}}$  and  $X^{\text{RBS}}$ , we refer to [31, §7].

Appropriately topologized, the quotients by the action of any arithmetic group  $\Gamma$  become the compact Hausdorff compactifications  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$ . We have a commutative diagram

$$\begin{array}{ccc} X^{\text{BS}} & \longrightarrow & Y^{\text{BS}} \\ \downarrow & & \downarrow \\ X^{\text{RBS}} & \longrightarrow & Y^{\text{RBS}} \end{array} .$$

The left vertical map is the identity on  $X$ , and for each proper  $\mathbb{Q}$ -parabolic subgroup  $\mathbf{P}$  is the projection  $N_{\mathbf{P}} \times X_{\mathbf{P}} \rightarrow X_{\mathbf{P}}$ . The right vertical map is the identity on  $Y$ ; to understand it on the boundary components, we need more notation. Let  $\Gamma_{N_{\mathbf{P}}} = \Gamma \cap N_{\mathbf{P}}$ , and let  $U_{\mathbf{P}}$  be the nilmanifold  $\Gamma_{N_{\mathbf{P}}} \backslash N_{\mathbf{P}}$ . Let  $\Gamma_{\mathbf{P}} = \Gamma \cap P$ , and let  $\Gamma_{M_{\mathbf{P}}} \subset M_{\mathbf{P}}$  be the discrete group obtained by projecting  $\Gamma_{\mathbf{P}}$  to  $M_{\mathbf{P}}$  via  $P = N_{\mathbf{P}} \times M_{\mathbf{P}} \times A_{\mathbf{P}} \rightarrow M_{\mathbf{P}}$ . Write  $Z_{\mathbf{P}}^{\text{BS}}$  (respectively,  $Z_{\mathbf{P}}^{\text{RBS}}$ ) for the boundary component of  $Y^{\text{BS}}$  (resp.,  $Y^{\text{RBS}}$ ) corresponding to  $\mathbf{P}$ . Then  $Z_{\mathbf{P}}^{\text{RBS}}$  is isomorphic to the locally symmetric space  $Y_{\mathbf{P}} = \Gamma_{M_{\mathbf{P}}} \backslash X_{\mathbf{P}}$ . In  $Y^{\text{BS}}$ , the component  $Z_{\mathbf{P}}^{\text{BS}}$  has the structure of a fiber bundle over  $Y^{\mathbf{P}}$  with fiber  $U_{\mathbf{P}}$ . For each  $\mathbf{P}$ , the map  $Z_{\mathbf{P}}^{\text{BS}} \rightarrow Z_{\mathbf{P}}^{\text{RBS}}$  collapses the nilmanifold  $U_{\mathbf{P}}$  to a point.

4.6.3. From the description of EDM geodesics on  $Y$  given in §4.5.1, it is clear that all points in the boundaries  $\partial Y^{\text{BS}}$  and  $\partial Y^{\text{RBS}}$  can be reached by following an EDM geodesic to its limit point. Hence there should be some construction of  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$  by putting a suitable equivalence relation on the EDM geodesics. This is indeed the case, and is carried out here.

If the  $\mathbb{Q}$ -rank of  $\mathbf{G}$  is 1, the construction is quite simple. For the Borel–Serre, Ji and MacPherson prove that there is a bijection between the set of all EDM geodesics and points in  $\partial Y^{\text{BS}}$ , given by taking the endpoint of an EDM geodesic  $\gamma$ :

$$\gamma \longmapsto \lim_{t \rightarrow \infty} \gamma(t)$$

The reductive Borel–Serre  $Y^{\text{RBS}}$  is a quotient of the Borel–Serre, so we need an equivalence relation on the EDM geodesics. We say two EDM geodesics  $\gamma, \gamma'$  are *N-equivalent* (notation:  $\gamma \stackrel{N}{\sim} \gamma'$ ) if

$$\lim_{t \rightarrow \infty} d(\gamma(t), \gamma'(t)) = 0$$

for appropriate parametrizations of  $\gamma, \gamma'$ . Denote the  $N$ -equivalence class of  $\gamma$  by  $[\gamma]_N$ . Then Ji–MacPherson show that there is bijection between the set of EDM

geodesics modulo  $N$ -equivalence and the points in  $\partial Y^{\text{RBS}}$ , again given by taking the endpoint:

$$(4.6.2) \quad [\gamma]_N \longmapsto \lim_{t \rightarrow \infty} \gamma(t).$$

Here the limit in (4.6.2) is taken in  $Y^{\text{RBS}}$ .

4.6.4. For higher  $\mathbb{Q}$ -ranks, similar results hold, although we need more equivalence relations to state them. Let  $\gamma$  be an EDM geodesic. We define two sets  $C(\gamma)$ ,  $F(\gamma)$  of EDM geodesics related to  $\gamma$  by

$$\begin{aligned} C(\gamma) &= \{ \gamma' \mid d(\gamma(t), \gamma'(t)) \text{ is constant for } t \gg 0 \} \\ F(\gamma) &= \{ \gamma' \mid \limsup_{t \rightarrow \infty} d(\gamma(t), \gamma'(t)) < \infty \}. \end{aligned}$$

The set  $C(\gamma)$  is called the *congruence bundle* of  $\gamma$ . It consists of all the EDM geodesics that are eventually at constant distance to  $\gamma$ . The set  $F(\gamma)$  is called the *finite bundle* of  $\gamma$ . Note that  $F(\gamma)$  is actually the set of all EDM geodesics that are *equivalent* to  $\gamma$  in the sense used to construct the geodesic compactification (§4.3.3).

4.6.5. The congruence bundle  $C(\gamma)$  can be turned into a metric space with metric  $\delta$  as follows. First we put  $\delta(\gamma', \gamma) = c$ , where  $c$  is the constant in the definition of  $C(\gamma)$ . This constant can also be recovered as  $\lim_{t \rightarrow \infty} \bar{d}(\gamma(t), \gamma')$ , where  $\bar{d}(\gamma(t), \gamma')$  is defined by

$$\inf \{ d(\gamma(t), \gamma'(s)) \mid s \in \mathbb{R} \}.$$

Via this description, we can also extend  $\delta$  to all of  $C(\gamma)$ . The authors show that  $(C(\gamma), \delta)$  is complete, and in fact has the following concrete form. If

$$\gamma(t) = (u, z, a \exp(tH)) \in N_{\mathbf{P}} \times X_{\mathbf{P}} \times A_{\mathbf{P}},$$

then

$$(4.6.3) \quad C(\gamma) \simeq Y_{\mathbf{P}} \times \text{span}(H)^\perp.$$

Here  $Y_{\mathbf{P}}$  is as in §4.6.2, and  $\text{span}(H)^\perp$  is the orthogonal complement to the line through  $H$  in  $\mathfrak{a}_{\mathbf{P}}$ .

4.6.6. Now we are ready to define our next equivalence relation. Let  $\gamma$  be EDM. We define the *rank*  $r(\gamma)$  of  $\gamma$  by

$$r = r(\gamma) = \{ k \in \mathbb{Z} \mid \text{there exists a faithful isometric action of } \mathbb{R}^{k-1} \text{ on } C(\gamma) \}.$$

We then say  $\gamma' \in C(\gamma)$  is *L-related* to  $\gamma$  (notation:  $\gamma \stackrel{L}{\sim} \gamma'$ ) if  $\gamma, \gamma'$  belong to the same  $\mathbb{R}^{r-1}$ -orbit. Here the  $L$  stands for *linear*; one pictures the  $\mathbb{R}^{r-1}$ -action as linearly sliding the geodesics in  $C(\gamma)$  around in the  $\text{span}(H)^\perp$  factor from (4.6.3).

The authors show that  $L$ -equivalence extends to an equivalence relation on the set of all EDM geodesics. In terms of (4.6.3),  $L$ -equivalence can be written as follows. If  $\gamma_1 \stackrel{L}{\sim} \gamma_2$ , then there exists  $\gamma$  with  $\gamma_i \in C(\gamma)$ . Write  $C(\gamma)$  as in (4.6.3), and write

$\gamma_i(t) = (u_i, z_i, a_i \exp(tH_i))$ . Then  $\gamma_1 \stackrel{L}{\sim} \gamma_2$  means that the  $u_i$  project to the same point in the nilmanifold  $U_{\mathbf{P}}$  and the  $z_i$  to the same point in the locally symmetric space  $Y_{\mathbf{P}}$ . From this it also follows that  $r(\gamma)$  is the  $\mathbb{Q}$ -rank of  $\mathbf{P}$ .

4.6.7. The restriction of  $L$ -equivalence to the finite bundle  $F(\gamma)$  also induces an equivalence relation. The dimension of the quotient  $F(\gamma)/\stackrel{L}{\sim}$  is called the *mobility degree* of  $\gamma$ , and is denoted  $\mu(\gamma)$ .

Using the mobility degree we can define another equivalence relation on EDM geodesics, called *R-equivalence*. The  $R$  stands for *rotation*. Let  $\gamma_0, \gamma_1$  be EDM and let  $[\gamma_i]_L$  be their  $L$ -equivalence classes. We say  $[\gamma_0]$  is  $R$ -equivalent to  $[\gamma_1]$ , and write  $[\gamma_0]_L \stackrel{R}{\sim} [\gamma_1]_L$ , if there exists a family  $\gamma_s$ ,  $s \in [0, 1]$  of EDM geodesics interpolating  $\gamma_0$  and  $\gamma_1$  with the following properties:

- (1)  $d(\gamma_{s_1}(t), \gamma_{s_2}(t)) = c|s_1 - s_2|t$  for  $t \geq 0$ , for some constant  $c$  and for all  $s_1, s_2 \in [0, 1]$ , and
- (2) the mobility degree  $\mu$  restricted to  $\gamma_s$  is constant.

Two EDM geodesics  $\gamma, \gamma'$  are said to be *RL-related* (notation:  $\gamma \stackrel{RL}{\sim} \gamma'$ ) if their  $L$ -classes are  $R$ -related:  $[\gamma]_L \stackrel{R}{\sim} [\gamma']_L$ .

For the reductive Borel–Serre, we will need to combine  $RL$ -equivalence with  $N$ -equivalence from §4.6.3. We say  $\gamma$  and  $\gamma'$  are *NRL-equivalent* if there exists an EDM geodesic  $\gamma''$  such that  $\gamma \stackrel{RL}{\sim} \gamma''$  and  $\gamma' \stackrel{N}{\sim} \gamma''$ .

4.6.8. We can finally explain how to construct  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$  using EDM geodesics. If the  $\mathbb{Q}$ -rank of  $\mathbf{G}$  is bigger than 1, then  $\partial Y^{\text{BS}}$  is in bijection with the  $RL$ -equivalence classes of EDM geodesics, via the endpoint map:

$$[\gamma]_{RL} \longmapsto \lim_{t \rightarrow \infty} \gamma(t),$$

where the limit is taken in  $Y^{\text{BS}}$ . For the reductive Borel–Serre, the boundary  $\partial Y^{\text{RBS}}$  is in bijection with the  $NRL$ -equivalence classes of EDM geodesics by the endpoint map:

$$[\gamma]_{NRL} \longmapsto \lim_{t \rightarrow \infty} \gamma(t),$$

where the limit is now taken in  $Y^{\text{RBS}}$ .

4.6.9. We can now finally explain the relationship between  $Y^{\text{geo}}$  and the compactifications  $Y^{\text{BS}}$ ,  $Y^{\text{RBS}}$ . First we need an explicit realization of  $RL$ -equivalence.

Suppose  $\gamma(t) = (u, z, a \exp(tH))$  and  $\gamma \stackrel{RL}{\sim} \gamma'$ . Then it turns out we can write  $\gamma'(t)$  in the form

$$\gamma'(t) = (u, z, a' \exp(tH'))$$

for the same parabolic subgroup  $\mathbf{P}$ . Hence two geodesics are  $RL$ -equivalent if their  $N_{\mathbf{P}}$  and  $X_{\mathbf{P}}$  components coincide; the  $A_{\mathbf{P}}$  part is irrelevant.



Notice that this is exactly the *opposite* of the basic equivalence relation used to construct  $Y^{\text{geo}}$ : for the geodesic compactification, the  $A_{\mathbf{P}}$  component is the *only* component of an EDM geodesic that matters; the  $N_{\mathbf{P}}$  and  $X_{\mathbf{P}}$  components play no role (§4.5.2). This is the sense in which  $Y^{\text{geo}}$  is as far as possible from  $Y^{\text{BS}}$ .

4.6.10. We conclude our discussion by giving fanciful pictures of the local structure of compactifications  $Y^{\text{BS}}$ ,  $Y^{\text{RBS}}$ , and  $Y^{\text{geo}}$  for  $\mathbf{G} = \text{SL}_3$  near the boundary components corresponding to the  $\Gamma$ -conjugacy classes of the standard rational parabolic subgroups.

Figure 4.2 shows the Borel–Serre compactification. This figure is based on one by MacPherson [37]. We use the notation of §4.4.7, so that partitions are 111, 21, and 12. If  $\pi$  is a partition corresponding to the standard rational parabolic subgroup  $\mathbf{P}$  with Langlands decomposition  $P = N_{\mathbf{P}}A_{\mathbf{P}}M_{\mathbf{P}}$ , we write  $Y_{\pi}$  for the locally symmetric space  $\Gamma_{M_{\mathbf{P}}}\backslash X_{\mathbf{P}}$  and  $U_{\pi}$  for the nilmanifold  $\Gamma_{N_{\mathbf{P}}}\backslash N_{\mathbf{P}}$  (§4.6.2). The codimension 1 corners are both torus bundles over locally symmetric spaces for  $\text{SL}_2$ . In these cases the unipotent radicals  $N_{12}$ ,  $N_{21}$  are both isomorphic to  $\mathbb{R}^2$  with the subgroups  $\Gamma_{12}$ ,  $\Gamma_{21}$  isomorphic to  $\mathbb{Z}^2$  acting by translation. The alignment of the tori in the figure indicates the structure of the unipotent radicals. For example, the unipotent radical  $N_{12}$  consists of all real matrices of the form

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In the codimension 2 corner  $Z_{111}$  the locally symmetric space  $Y_{111}$  is a point. Hence all the topology of  $Z_{111}$  is contained in the nilmanifold  $U_{111}$ . This 3-manifold is known as the *Heisenberg manifold*. It can be written as a bundle  $T^2 \rightarrow Z_{111} \rightarrow S^1$  in two different ways, reflecting the two subgroups  $N_{12}, N_{21} \subset N_{111}$ .<sup>(9)</sup>

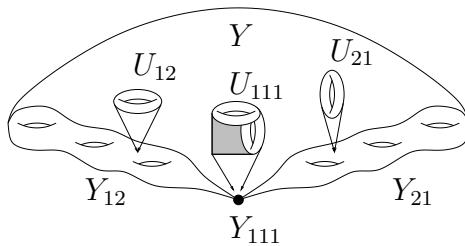
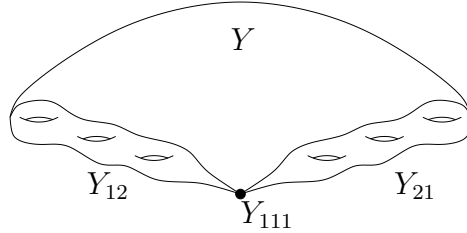


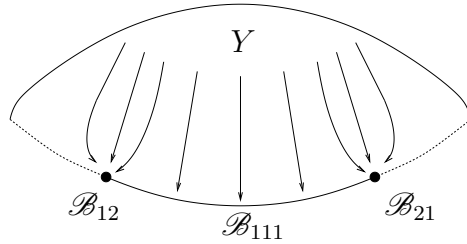
FIGURE 4.2. Borel–Serre compactification for  $\text{SL}_3$

Figure 4.3 shows the reductive Borel–Serre compactification. In this figure the nilmanifold fibers have been collapsed to points, and all that remains in the boundary are the lower-rank locally symmetric spaces.

<sup>9</sup>Incidentally, both of these bundles are different from the  $T^2$ -bundle in §4.2.8.


 FIGURE 4.3. Reductive Borel–Serre compactification for  $SL_3$ 

Finally, Figure 4.4 shows the geodesic compactification with some geodesics suggestively converging to boundary components. Here we write  $\mathcal{B}_\pi$  for the simplex in the building corresponding to the standard parabolic subgroup  $P_\pi$ . As expected from §4.6.9, the boundary components here are in some sense totally opposite to those of  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$ : as the codimension increases in the boundaries of  $Y^{\text{BS}}$  and  $Y^{\text{RBS}}$ , so does the *dimension* of the corresponding boundary components of  $Y^{\text{geo}}$ . This illustrates why, if the  $\mathbb{Q}$ -rank is  $> 1$ , the greatest common quotient of  $Y^{\text{BS}}$  and  $Y^{\text{geo}}$  is the one-point compactification.


 FIGURE 4.4. Geodesic compactification for  $SL_3$ 

#### 4.7. Complements.

4.7.1. We conclude by briefly summarizing some of the other material in [31].

4.7.2. Let  $M$  be a complete Riemannian manifold, and let  $\Delta$  be the Laplace operator on  $L^2(M)$ . It is known that if  $M$  is noncompact, then the continuous spectrum of  $\Delta$  (if it exists) cannot change under compact perturbation of  $M$ . Hence one has the natural problem of trying to understand the connection between the continuous spectrum of  $\Delta$  and compactifications of  $M$ .

In the case that  $M$  is a locally symmetric space  $Y = \Gamma \backslash X$ , it is known that  $\Delta$  has continuous spectrum thanks to Langlands’s study of Eisenstein series [33, 41]. Ji and MacPherson are able to describe the continuous spectrum of  $\Delta$  in terms of the geometry of  $Y^{\text{geo}}$  by reinterpreting Langlands’s fundamental work. This provides a very accessible and geometric introduction to the intricate constructions in [33, 41].

4.7.3. Let  $M$  be a complete Riemannian manifold and let  $\lambda$  be any real number less than the bottom of the spectrum of the Laplace operator  $\Delta$ . Associated to this data is a certain compactification of  $M$ , the *Martin compactification*  $M_\lambda^{\text{Mar}}$ . The precise definition is somewhat involved; we refer to [28] and [31, §15] for details. Using Eisenstein series, Ji and MacPherson show that for a locally symmetric space  $Y$  the geodesic compactification  $Y^{\text{geo}}$  canonically injects into the Martin compactification  $Y_\lambda^{\text{Mar}}$ . They also conjecture that this injection is in fact a homeomorphism.

## REFERENCES

- [1] J. Arthur, *Introduction to the trace formula*, Harmonic Analysis, The Trace Formula, And Shimura Varieties (J. Arthur, D. Ellwood, and R. Kottwitz, eds.), CMI/AMS Publications, 2005, Proceedings of the 2003 Clay Mathematics Institute Summer School.
- [2] J. Arthur, *The  $L^2$ -Lefschetz numbers of Hecke operators*, Invent. Math. **97** (1989), no. 2, 257–290.
- [3] A. Ash, D. Mumford, M. Rapaport, and Y. Tai., *Smooth compactifications of locally symmetric varieties*, Math. Sci. Press, Brookline, Mass., 1975.
- [4] A. Ash, *Deformation retracts with lowest possible dimension of arithmetic quotients of self-adjoint homogeneous cones*, Math. Ann. **225** (1977), no. 1, 69–76.
- [5] ———, *Small-dimensional classifying spaces for arithmetic subgroups of general linear groups*, Duke Math. J. **51** (1984), no. 2, 459–468.
- [6] W. L. Baily, Jr. and A. Borel, *Compactification of arithmetic quotients of bounded symmetric domains*, Ann. of Math. (2) **84** (1966), 442–528.
- [7] W. Ballmann, M. Gromov, and V. Schroeder, *Manifolds of nonpositive curvature*, Progress in Mathematics, vol. 61, Birkhäuser Boston Inc., Boston, MA, 1985.
- [8] A. Borel, *Introduction aux groupes arithmétiques*, Hermann, Paris, 1969, Publications de l’Institut de Mathématique de l’Université de Strasbourg, XV. Actualités Scientifiques et Industrielles, No. 1341.
- [9] A. Borel and L. Ji, *Compactifications of symmetric and locally symmetric spaces*, Mathematics: Theory & Applications, Birkhäuser, 2005.
- [10] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comm. Math. Helv. **48** (1973), 436–491.
- [11] W. A. Casselman, *Geometric rationality of Satake compactifications*, Algebraic groups and Lie groups, Austral. Math. Soc. Lect. Ser., vol. 9, Cambridge Univ. Press, Cambridge, 1997, pp. 81–103.
- [12] J. H. Conway and N. J. A. Sloane, *Low-dimensional lattices. III. Perfect forms*, Proc. R. Soc. Lond. A. **418** (1988), 43–80.
- [13] G. E. Cooke and R. L. Finney, *Homology of cell complexes*, Based on lectures by Norman E. Steenrod, Princeton University Press, Princeton, N.J., 1967.
- [14] G. Danaraj and V. Klee, *Shellings of spheres and polytopes*, Duke Math. J. **41** (1974), 443–451.
- [15] M. Dutour, F. Vallentin, and A. Schürmann, *Classification of perfect forms in dimension 8*, talk at Oberwolfach meeting *Sphere packings: Exceptional structures and relations to other fields*, November 2005.
- [16] F. Ehlers, *Eine Klasse komplexer Mannigfaltigkeiten und die Auflösung einiger isolierter Singularitäten*, Math. Ann. **218** (1975), no. 2, 127–156.
- [17] J. Franke, *Harmonic analysis in weighted  $L_2$ -spaces*, Ann. Sci. École Norm. Sup. (4) **31** (1998), no. 2, 181–279.

- [18] M. Goresky, *Compactifications and cohomology of modular varieties*, Harmonic Analysis, The Trace Formula, And Shimura Varieties (J. Arthur, D. Ellwood, and R. Kottwitz, eds.), CMI/AMS Publications, 2005, Proceedings of the 2003 Clay Mathematics Institute Summer School.
- [19] M. Goresky, G. Harder, and R. MacPherson, *Weighted cohomology*, Invent. Math. **116** (1994), no. 1-3, 139–213.
- [20] M. Goresky, R. Kottwitz, and R. MacPherson, *Discrete series characters and the Lefschetz formula for Hecke operators*, Duke Math. J. **89** (1997), no. 3, 477–554.
- [21] ———, *Correction to: “Discrete series characters and the Lefschetz formula for Hecke operators”*, Duke Math. J. **92** (1998), no. 3, 665–666.
- [22] ———, *Homology of affine Springer fibers in the unramified case*, Duke Math. J. **121** (2004), no. 3, 509–561.
- [23] ———, *Regular points in affine Springer fibers*, Michigan Math. J. **53** (2005), no. 1, 97–107.
- [24] M. Goresky and R. MacPherson, *Stratified Morse theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 14, Springer-Verlag, Berlin, 1988.
- [25] ———, *Lefschetz numbers of Hecke correspondences*, The zeta functions of Picard modular surfaces, Univ. Montréal, Montreal, QC, 1992, pp. 465–478.
- [26] ———, *Local contribution to the Lefschetz fixed point formula*, Invent. Math. **111** (1993), no. 1, 1–33.
- [27] ———, *The topological trace formula*, J. Reine Angew. Math. **560** (2003), 77–150.
- [28] Y. Guivarc’h, L. Ji, and J. C. Taylor, *Compactifications of symmetric spaces*, Progress in Mathematics, vol. 156, Birkhäuser Boston Inc., Boston, MA, 1998.
- [29] R. A. Herb, *Characters of averaged discrete series on semisimple real Lie groups*, Pacific J. Math. **80** (1979), no. 1, 169–177.
- [30] F. E. P. Hirzebruch, *Hilbert modular surfaces*, Enseignement Math. (2) **19** (1973), 183–281.
- [31] L. Ji and R. MacPherson, *Geometry of compactifications of locally symmetric spaces*, Ann. Inst. Fourier (Grenoble) **52** (2002), no. 2, 457–559.
- [32] M. Kashiwara and P. Schapira, *Sheaves on manifolds*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 292, Springer-Verlag, Berlin, 1994, With a chapter in French by Christian Houzel, Corrected reprint of the 1990 original.
- [33] R. P. Langlands, *On the functional equations satisfied by Eisenstein series*, Springer-Verlag, Berlin, 1976, Lecture Notes in Mathematics, Vol. 544.
- [34] G. Laumon and B. C. Ngo, *Le lemme fondamental pour les groupes unitaires*, arXiv:math.AG/0404454.
- [35] J.-S. Li and J. Schwermer, *Automorphic representations and cohomology of arithmetic groups*, Challenges for the 21st century (Singapore, 2000), World Sci. Publishing, River Edge, NJ, 2001, pp. 102–137.
- [36] E. Looijenga,  *$L^2$ -cohomology of locally symmetric varieties*, Compositio Math. **67** (1988), no. 1, 3–20.
- [37] R. MacPherson, *Modular varieties*, course given at MIT, 1990.
- [38] R. MacPherson and M. McConnell, *Classical projective geometry and modular varieties*, Algebraic analysis, geometry, and number theory (Baltimore, MD, 1988), Johns Hopkins Univ. Press, Baltimore, MD, 1989, pp. 237–290.
- [39] ———, *Explicit reduction theory for Siegel modular threefolds*, Invent. Math. **111** (1993), no. 3, 575–625.

- [40] M. McConnell, *Generalizations of Voronoï I to symmetric spaces and arithmetic groups*, Voronoï's Impact on Modern Science (Kyiv, Ukraine), vol. I, National Academy of Sciences of Ukraine, 1998, pp. 155–185.
- [41] C. Mœglin and J.-L. Waldspurger, *Décomposition spectrale et séries d'Eisenstein*, Progress in Mathematics, vol. 113, Birkhäuser Verlag, Basel, 1994, Une paraphrase de l'Écriture. [A paraphrase of Scripture].
- [42] D. Mumford, *A new approach to compactifying locally symmetric varieties*, Discrete subgroups of Lie groups and applications to moduli (Internat. Colloq., Bombay, 1973), Oxford Univ. Press, Bombay, 1975, pp. 211–224.
- [43] A. Nair, *Weighted cohomology of arithmetic groups*, Ann. of Math. (2) **150** (1999), no. 1, 1–31.
- [44] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics, vol. 139, Academic Press Inc., Boston, MA, 1994, Translated from the 1991 Russian original by Rachel Rowen.
- [45] L. Saper, *On the cohomology of locally symmetric spaces and of their compactifications*, Current developments in mathematics, 2002, Int. Press, Somerville, MA, 2003, pp. 219–289.
- [46] L. Saper and M. Stern,  *$L_2$ -cohomology of arithmetic varieties*, Ann. of Math. (2) **132** (1990), no. 1, 1–69.
- [47] I. Satake, *On compactifications of the quotient spaces for arithmetically defined discontinuous groups*, Ann. Math. **72** (1960), 555–580.
- [48] ———, *On representations and compactifications of symmetric Riemannian spaces*, Ann. Math. **71** (1960), 77–110.
- [49] C. Soulé, *Cohomologie de  $SL_3(\mathbb{Z})$* , C. R. Acad. Sci. Paris Sér. A-B **280** (1975), no. 5, A1, A251–A254.
- [50] G. Voronoï, *Nouvelles applications des paramètres continus à la théorie des formes quadratiques, I. Sur quelques propriétés des formes quadratiques positives parfaites*, J. Reine Angew. Math. **133** (1908), 97–178.
- [51] D. Yasaki, *On the cohomology of  $SU(2,1)$  over the Gaussian integers*, preprint, 2005.
- [52] ———, *On the existence of spines for  $\mathbb{Q}$ -rank 1 groups*, preprint, 2005.
- [53] S. Zucker, *Satake compactifications*, Comment. Math. Helvetici **58** (1983), 312–343.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MASSACHUSETTS, AMHERST, MA 01003, USA

*E-mail address:* gunnells@math.umass.edu