Quantum Ergodicity on Bruhat-Tits Buildings

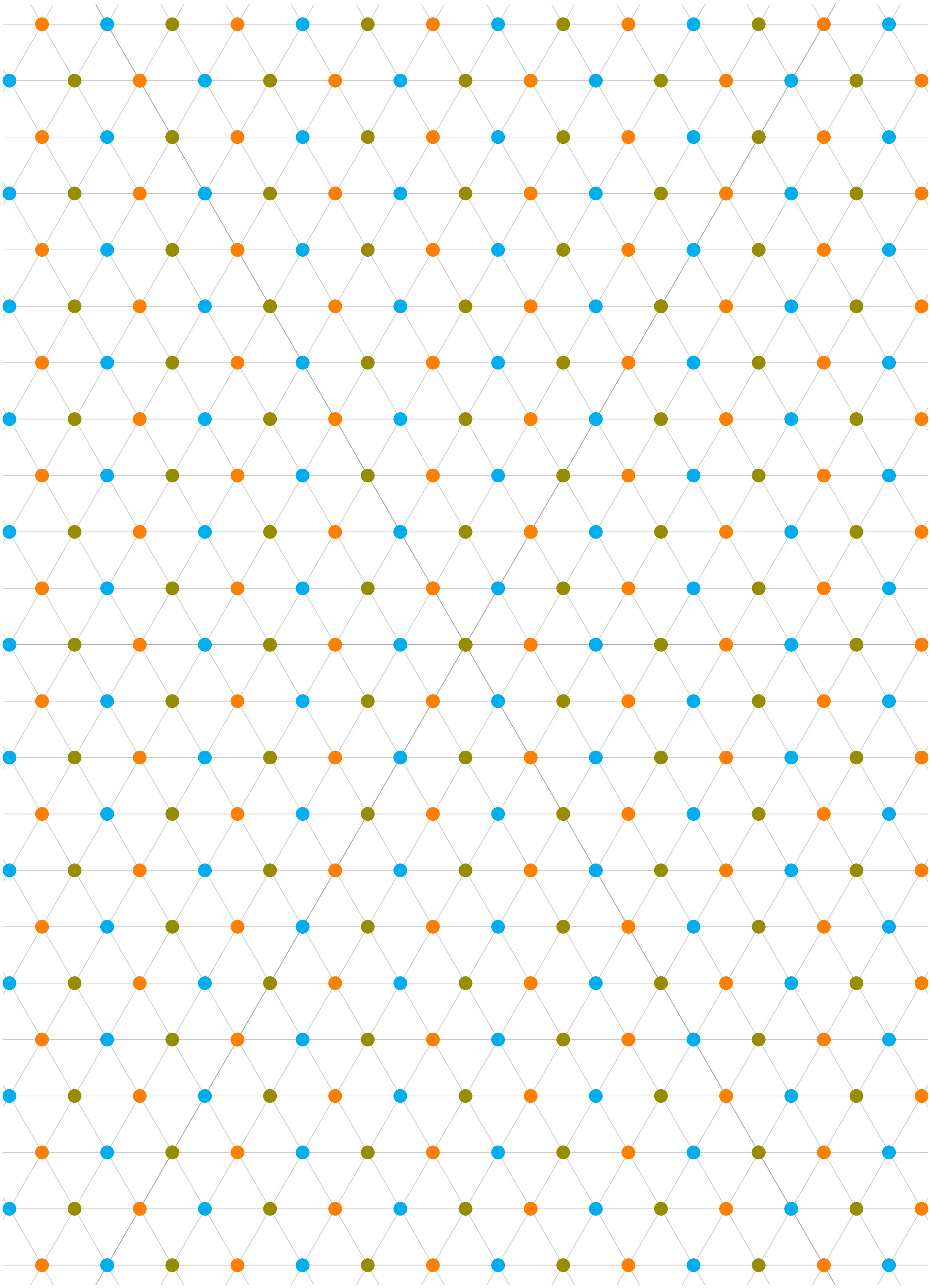
by

Carsten Peterson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2023

Doctoral Committee:

Professor Ralf J. Spatzier, Chair
Maître de Conférences Farrell Brumley, Université Sorbonne Paris Nord
Professor Stephen M. DeBacker
Professor Tasho Kaletha
Professor James T. Liu



Carsten Peterson
carstenp@umich.edu
ORCID iD: 0000-0002-8088-9000

To Mom, Dad, Soren, Maiya, Silja, and Eva

ACKNOWLEDGEMENTS

First off I would like to thank my advisor, Ralf Spatzier. His incredible support, encouragement, wisdom, and generosity have been invaluable in my PhD journey. Furthermore he has fostered my mathematical independence and confidence, and greatly shaped who I am as a mathematician.

I would like to thank my committee members for serving on my thesis committee. Additionally, I would like to thank Farrell Brumley, Stephen DeBacker, and Simon Marshall for listening to my ideas and providing thoughtful feedback.

I have had the privilege of having many incredible math teachers, professors, and mentors who have encouraged and inspired me over the years. In particular I would like to thank Dan Teague, Stefan Steinerberger, Nathan Kaplan, and Steven J. Miller.

I am deeply grateful to my parents for their unwavering love and support. I also thank my siblings for their companionship and camaraderie, both in childhood and beyond.

Lastly, I would like to thank Karen Butt, Peter Dillery, Nick Wawrykow, Yueqiao Wu, and Bradley Zykoski, who have all greatly enhanced my grad school experience. We did it, guys!

This work was partially supported by the National Science Foundation Grant DMS-2003712 and RTG Grant DMS-1840234.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

We study eigenfunctions of the spherical Hecke algebra acting on $L^2(\Gamma_n \backslash G/K)$ where $G = \mathrm{PGL}(3, F)$ with F a non-archimedean local field of characteristic zero, $K = \mathrm{PGL}(3, \mathcal{O})$ with $\mathcal{O}$ the ring of integers of F, and (Γ_n) is a sequence of cocompact torsionfree lattices. We prove a form of equidistribution on average for eigenfunctions whose spectral parameters lie in the tempered spectrum when the associated sequence of quotients of the Bruhat-Tits building Benjamini-Schramm converges to the building itself.

CHAPTER I
Introduction

I.1: Quantum ergodicity in the large eigenvalue limit

Originally *quantum ergodicity* concerned eigenfunctions of the Laplacian on a closed Riemannian manifold (M, g) with ergodic geodesic flow. The geodesic flow being *ergodic* implies that a "classical particle" moving along a generic geodesic is equally likely to end up everywhere on M in the long run. On the other hand, a "quantum particle" on M with wave function ψ has probability measure $|\psi|^2 d\mathrm{vol}_g$ of being "observed" in a given region of M. If the geodesic flow is ergodic, we expect the quantum particle is equally likely to be observed everywhere.

Through a procedure known as *quantization*, one is led to studying the Laplacian on M in place of the geodesic flow. The Laplacian has an orthonormal basis of L^2-eigenfunctions $\{\psi_j\}$. The associated eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq ...$ are all non-negative and go to infinity as j goes to infinity. Šnirelman [Sni74], Zelditch [Zel87], and Colin de Verdière [CdV85] proved the *quantum ergodicity theorem* which says that for a density-one subsequence of the eigenfunctions (ψ_{j_k}), the associated measures $|\psi_{j_k}|^2 d\mathrm{vol}_g$ weak-* converge to $\frac{d\mathrm{vol}_g}{\mathrm{vol}(M)}$, the normalized volume measure. In fact this statement is a consequence of the following:

Theorem I.1 (Quantum ergodicity theorem [Sni74, Zel87, CdV85])**.** *Suppose (M, g) is a closed Riemannian manifold with ergodic geodesic flow. Then for every smooth test function $a \in C^\infty(M)$,*

$$\lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} \left| \langle \psi_j, a\psi_j \rangle - \frac{1}{\mathrm{vol}(M)} \int_M a \ d\mathrm{vol}_g \right|^2 = 0, \tag{I.1.1}$$

where $N(\lambda) = \#\{i : \lambda_i \leq \lambda\}$.

In a certain sense letting the eigenvalue go to infinity is analogous to letting Planck's constant go to zero, so we expect to recover "classical mechanical" results (e.g. equidistribution of generic orbits when the geodesic flow is ergodic) in such a limit.

Remark I.2. In fact Theorem I.1 can be strengthened to replacing the summands in (I.1.1) by

$$\left|\langle\psi_j, \mathcal{T}\psi_j\rangle - \int_{S^*M} \sigma^0(\mathcal{T})dL\right|^2, \tag{I.1.2}$$

where $\mathcal{T}$ is an order 0 pseudodifferential operator on M, S^*M is the unit (co)tangent bundle of M (which is the space on which the geodesic flow occurs and is ergodic), $\sigma^0(\mathcal{T})$ is the principal symbol of $\mathcal{T}$ (which is a function on S^*M), and dL is the Louiville measure (normalized so that the total measure of S^*M is 1).

Remark I.3. Rudnick-Sarnak [RS94] conjectured *quantum unique ergodicity*, namely that, if M has negative curvature, then the only weak-* limit of the eigenfunction measures (or their *microlocal lifts* as in (I.1.2)) is the normalized volume measure on M (or the Liouville measure on S^*M). Lindenstrauss [Lin06] proved quantum unique ergodicity for joint eigenfunctions of the Laplacian and the Hecke operators on compact arithmetic hyperbolic surfaces (see also [Lin01, Sou10, BL14]).

I.2: Quantum ergodicity in the Benjamini-Schramm limit

As opposed to the original quantum ergodicity theorem, which concerns eigenfunctions of the Laplacian on a fixed manifold, more recently many authors have considered *quantum ergodicity in the Benjamini-Schramm limit*, which concerns eigenfunctions of Laplacian-like operators for sequences of spaces "converging" to their common universal cover. More specifically, we say that a sequence of spaces *Benjamini-Schramm converges* to their common (contractible) universal cover if asymptotically almost every point has arbitrarily large injectivity radius [BS01]. Benjamini-Schramm convergence may be viewed as a probabilistic version of Gromov-Hausdorff convergence of metric spaces.

Many authors have studied the relationship between Benjamini-Schramm convergence and the distribution of eigenvalues (for example, Kesten [Kes59] and Mckay [McK81] for regular graphs, or the more recently Abert-Bergeron-Biringer-Gelander-Nikolov-Raimbault-Samet [ABB+17] for locally symmetric spaces). Such results are often closely related to representation theory in which case they often appear in the literature as *limit multiplicity* theorems. Quantum ergodicity in the Benjamini-Schramm limit may be seen as an extension of such questions to eigenfunctions; on the other hand it arose out of attempts to extend quantum ergodicity ideas to different contexts.

One such precedent for quantum ergodicity in the Benjamini-Schramm limit is modular forms on congruence coverings of the modular surface. Modular forms have two natural parameters: weight and level. *Weight* is roughly analogous to Laplacian eigenvalue and *level* relates to the congruence covering. Holowinsky and Soundararajan [HS10] proved quantum unique ergodicity for fixed level and letting weight go to infinity. Fixing weight and letting level go to infinity results in a sequence of hyperbolic surfaces which Benjamini-Schramm converge to the hyperbolic plane as proven by Fraczyk [Fra21], and quantum unique ergodicity in this context was proven by Nelson [Nel11] and Nelson-Pitale-Saha [NPS14]. These latter results are referred to in the literature as *quantum ergodicity in the level aspect*, and consequently this terminology is also used in place of quantum ergodicity in the Benjamini-Schramm limit.

Another motivating example of quantum ergodicity in the Benjamini-Schramm limit comes from the work of Anantharaman-Le Masson [ALM15] who proved that for sequences of regular graphs Benjamini-Schramm converging to the infinite regular tree, the eigenfunctions of the adjacency operator are, on average, equidistributed in a weak sense. In particular their results include the following:

Theorem I.4 (Quantum ergodicity on large regular graphs [ALM15]; see also [BLML16, Ana17, AS19])**.** *Suppose (G_n) is a sequence of $(q+1)$-regular graphs which Benjamini-Schramm converge to the $(q+1)$-regular tree. Suppose also that there is a uniform spectral gap for the adjacency operator, namely all non-trivial eigenvalues are uniformly bounded away from $\pm(q+1)$. Let a_n be a function on the vertices of G_n such that $||a_n||_\infty \leq 1$, and let* $\mathrm{card}(G_n)$ *denote the number of vertices of G_n. Let I be a closed subinterval of $[-2\sqrt{q}, 2\sqrt{q}]$ with non-empty interior. Let $\psi_1^{(n)}, \ldots, \psi_{\mathrm{card}(G_n)}^{(n)}$ be an orthonormal basis of eigenfunctions of the adjacency operator on G_n with associated eigenvalues $\lambda_1^{(n)}, \ldots, \lambda_{\mathrm{card}(G_n)}^{(n)}$. Then*

$$\lim_{n\to\infty} \frac{1}{N(I,G_n)} \sum_{\psi_j^{(n)}:\lambda_j^{(n)}\in I} \left| \langle \psi_j^{(n)}, a_n \psi_j^{(n)} \rangle - \frac{1}{\mathrm{card}(G_n)} \sum_{\text{vertices } v\in G_n} a_n(v) \right|^2 = 0, \tag{I.2.1}$$

where $N(I,G_n) = \#\{j : \lambda_j^{(n)} \in I\}$.

Le Masson-Sahlsten [LMS17] noted the similarities between quantum ergodicity in the level aspect for modular forms and quantum ergodicity on large regular graphs, and they proved that for sequences of compact hyperbolic surfaces which Benjamini-Schramm converge to the hyperbolic plane, eigenfunctions of the Laplacian with eigenvalue in a fixed compact subinterval of $[1/4, \infty)$ are, on average, equidistributed in a weak sense. Parts of their technique were adapted from that of Brooks-Le Masson-Lindenstrauss [BLML16] who reproved Theorem I.4. This technique was further adapted by Abert-Bergeron-Le Masson [ABLM18] to prove analogous results for sequences of compact rank one locally symmetric spaces converging to the their common universal cover; these authors also connect the original quantum ergodicity theorem (Theorem I.1) with quantum ergodicity in the Benjamini-Schramm limit. More recently Le Masson and Sahlsten [LMS20] have further adapted the technique to handle finite-volume non-compact hyperbolic surfaces.

I.3: Quantum ergodicity in higher rank

Much of the research in quantum ergodicity has been focused on the case of hyperbolic surfaces. Such manifolds are very special: their universal cover is a *symmetric space*. More precisely, the *hyperbolic plane* $\mathbb{H}$ may be realized as $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$, and each hyperbolic surface X may be realized as $\Gamma\backslash\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ for some discrete subgroup Γ (which must be a cocompact lattice in case X is compact). Furthermore, the geodesic flow on the unit tangent bundle of X may be identified with the right action of the subgroup $\left(\begin{smallmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{smallmatrix}\right)$ on $\Gamma\backslash\mathrm{SL}(2,\mathbb{R})/\{\pm I\}$.

Now suppose G is a non-compact semisimple Lie group with Lie algebra $\mathfrak{g}$. Using a Cartan involution, we may write $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ where $K = \exp(\mathfrak{k})$ is a maximal compact subgroup. Let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal toral subalgebra, and let $A = \exp(\mathfrak{a})$. Then G/K has a natural G-invariant metric with respect to which it is a contractible manifold with non-positive curvature; such spaces are called symmetric spaces of non-compact type, and K is clearly the stabilizer of the point $1K$. The orbit of $1K$ under A is a flat subspace; it is an example of a *maximal flat* in G/K.

Let $\Gamma < G$ be a lattice. The manifold $\Gamma\backslash G/K$ is called a *locally symmetric space*. If the (real) rank of G, or equivalently the dimension of $\mathfrak{a}$, is greater than one, then the geodesic flow on any associated locally symmetric space is never ergodic [BM00]. On the other hand, we may consider the A-action on the space $\Gamma\backslash G/M$ where $M = Z_K(A)$; this double coset space may be viewed as the "bundle of oriented flats with basepoint" over the underlying locally symmetric space (see, e.g., Section 5.3 of [SV07]). This A-action is ergodic and reduces to the geodesic flow in rank one [BM00].

The validity of the strengthened version of Theorem I.6 as in Remark I.2 is in fact known to be equivalent to the ergodicity of the geodesic flow [Zel06]. Hence to extend the ideas of quantum ergodicity to a higher rank setting, it no longer suffices to simply consider eigenfunctions of the Laplacian. Lindenstrauss [Lin01] suggested eigenfunctions of $\mathcal{D}(G/K)$, the *algebra of invariant differential operators on* G/K, as the object of study for quantum ergodicity on higher rank locally symmetric spaces. This perspective was taken up in [SV07, AS13, SV19, BM21]. In particular, Brumley-Matz [BM21] investigated quantum ergodicity in the Benjamini-Schramm limit result for sequences of locally symmetric spaces $\Gamma_n\backslash\mathrm{SL}(d,\mathbb{R})/\mathrm{SO}(d)$ which Benjamini-Schramm converge to $\mathrm{SL}(d,\mathbb{R})/\mathrm{SO}(d)$.

In rank one, the algebra $\mathcal{D}(G/K)$ is generated by the Laplacian [Kna86]. More generally, via the Harish-Chandra isomorphism, one may identify $\mathcal{D}(G/K)$ with a polynomial ring whose associated variety may be identified with $\mathfrak{a}_{\mathbb{C}}^*/W$ where $W = N_G(\mathfrak{a})/Z_G(\mathfrak{a})$ is the Weyl group. Given a joint eigenfunction ψ of $\mathcal{D}(G/K)$ acting on $\Gamma\backslash G/K$, we get an associated homomorphism $\chi \in \mathrm{Hom}_{\mathbb{C}\text{-alg.}}(\mathcal{D}(G/K),\mathbb{C})$ via $D\psi = \chi(D)\psi$, and hence an associated point $\nu \in \mathfrak{a}_{\mathbb{C}}^*/W$ which is called the *spectral parameter* of ψ. The locus $i\mathfrak{a}^*/W \subset \mathfrak{a}_{\mathbb{C}}^*/W$ plays a distinguished role and is called the *tempered spectrum.* In case $G = \mathrm{SL}(2,\mathbb{R})$, the tempered spectrum may be identified under a natural mapping with $[1/4,\infty)$ which is often referred to as the tempered spectrum of the Laplacian on $\mathbb{H}$.

We now state the main theorem of Brumley-Matz [BM21], but we first recall that a sequence of lattices in G is called *uniformly discrete* if there is a universal lower bound on the injectivity radii of the associated locally symmetric spaces.

Theorem I.5 (Brumley-Matz [BM21])**.** *Let $d \geq 3$. Let $G = \mathrm{SL}(d,\mathbb{R})$ and $K = \mathrm{SO}(d)$. Suppose $\Gamma_n < G$ is a uniformly discrete sequence of torsionfree cocompact lattices. Let $Y_n = \Gamma_n\backslash G/K$. Suppose $\mathrm{vol}(Y_n) \to \infty$. Let a_n be a measurable function on Y_n such that $||a_n||_\infty \leq 1$. Let $\{\psi_j^{(n)}\}$ be an orthonormal basis for $L^2(Y_n)$ of eigenfunctions of $\mathcal{D}(G/K)$ with associated spectral parameters $\{\nu_j^{(n)}\}$. Then, there is $\rho > 1$ such that for sufficiently regular $\nu \in i\mathfrak{a}^*$, we have*

$$\lim_{n\to\infty} \frac{1}{N(B_0(\nu,\rho),Y_n)} \sum_{j:\nu_j^{(n)}\in B_0(\nu,\rho)} \left|\langle\psi_j^{(n)}, a_n\psi_j^{(n)}\rangle - \frac{1}{\mathrm{vol}(Y_n)}\int_{Y_n} a_n \ d\mathrm{vol}_{Y_n}\right|^2 = 0,$$

where $B_0(\nu,\rho) = \{\lambda \in i\mathfrak{a}^* : ||\lambda-\nu||_2 \leq \rho\}$ is the ball of radius ρ centered at ν in the tempered spectrum, and $N(B_0(\nu,\rho),Y_n) = \#\{j : \nu_j^{(n)} \in B_0(\nu,\rho)\}$.

The assumption that $\mathrm{vol}(Y_n) \to \infty$ is in fact known to be equivalent to Benjamini-Schramm convergence in rank at least 2 [ABB+17]. Furthermore, in contrast to Theorem

I.4 and the work of [LMS17, ABLM18] for rank one locally symmetric spaces, there is no uniform spectral gap assumption needed as it is in fact automatic in rank at least 2 by property (T) [BdlHV08].

I.4: Main result: quantum ergodicity in the Benjamini-Schramm limit for the Bruhat-Tits building associated to $\mathrm{PGL}(3, F)$

Bruhat-Tits buildings are infinite simplicial complexes constructed from reductive algebraic groups over non-archimedean local fields [BT72]. The simplest example is the Bruhat-Tits building associated to $\mathrm{SL}(2, \mathbb{Q}_p)$, which is the infinite $(p+1)$-regular tree. Bruhat-Tits buildings may be viewed as non-archimedean analogues of symmetric spaces of non-compact type. On the other hand, their quotients may be seen as "higher rank" generalizations of regular graphs. Such quotients have also been studied recently because, in certain cases, they provide examples of high-dimensional expanders known as Ramanujan complexes (see, e.g. [LSV05b, LSV05a, Lub14]).

Suppose F is a non-archimedean local field of characteristic zero, and $\mathcal{O}$ is its ring of integers. Suppose G is a reductive algebraic group over F, and K is a hyperspecial maximal compact subgroup. Let $\mathcal{B}$ be the associated building. Then K is the stabilizer of a unique special vertex $x_0 \in \mathcal{B}$. There is a correspondence between maximal F-split tori in G and so-called *apartments* in $\mathcal{B}$; these apartments are the analogues for Bruhat-Tits buildings of maximal flats in symmetric spaces. Let $T < G$ be a maximal F-split torus whose associated apartment contains x_0. Let $\Gamma < G$ be a lattice. Let $M = Z_K(T)$ and let $A = T/(T \cap K)$. Similarly to the case of (locally) symmetric spaces, there is an ergodic right A-action on $\Gamma\backslash G/M$. This double coset space may be viewed as the "bundle of oriented apartments with special vertex basepoint" over $\Gamma\backslash\mathcal{B}$. It is this ergodic action which in some sense we are "quantizing" in this work.

We shall be particularly concerned with the Bruhat-Tits building associated to $G = \mathrm{PGL}(d, F)$ (specifically the case of $d = 3$). Then $K = \mathrm{PGL}(d, \mathcal{O})$, and G/K may be identified with the vertices of $\mathcal{B}$. The *spherical Hecke algebra* $H(G, K)$ is the analogue of $\mathcal{D}(G, K)$ in this context, and it acts on functions on G/K (or quotients thereof). In the case of $d = 2$, the algebra $H(G, K)$ is generated by one element whose associated action on G/K is equivalent to the adjacency operator on the infinite $(q+1)$-regular tree [Ser80]. If $\Gamma < G$ is a cocompact torsionfree lattice, then $Y = \Gamma\backslash G/K$ is a finite simplicial complex. The space $L^2(Y)$ has an orthonormal basis of eigenfunctions $\{\psi_j\}$ of $H(G, K)$. By the Satake isomorphism, the spherical Hecke algebra is isomorphic to the coordinate ring of a variety Ω, and hence given an eigenfunction ψ_j, it makes sense to talk about its associated *spectral parameter* $\nu_j \in \Omega$. There is a distinguished sublocus of Ω called the *tempered spectrum*, denoted by Ω^+_{temp}. The tempered spectrum carries a natural probability measure called the *Plancherel measure*. In the case of $n = 2$, the tempered spectrum is transformed into $[-2\sqrt{q}, 2\sqrt{q}]$ under a natural mapping; this locus is often referred to as the tempered spectrum of the adjacency operator on the $(q + 1)$-regular tree, and it appeared in Theorem I.4.

Our main result is the following:

Theorem I.6. *Let $G = \mathrm{PGL}(3, F)$ and $K = \mathrm{PGL}(3, \mathcal{O})$, where F is a non-archimedean local field of characteristic zero and $\mathcal{O}$ is its ring of integers. Let $\Gamma_n < G$ be a sequence of torsionfree lattices. Let $Y_n = \Gamma_n\backslash G/K$. Suppose $\mathrm{card}(Y_n) \to \infty$. Let a_n be a function on Y_n such that $||a_n||_\infty \leq 1$. Let $\Theta \subset \Omega^+_{\text{temp}}$ be a compact subset with positive Plancherel measure and not meeting a certain codimension one exceptional locus Ξ. Let $\{\psi_j^{(n)}\}$ denote an orthonormal basis of eigenfunctions of $H(G, K)$ acting on $L^2(Y_n)$. Then*

$$\lim_{n\to\infty} \frac{1}{N(\Theta, Y_n)} \sum_{\psi_j^{(n)}:\nu_j^{(n)}\in\Theta} \left|\langle\psi_j^{(n)}, a_n\psi_j^{(n)}\rangle - \frac{1}{\mathrm{card}(Y_n)} \sum_{\text{vertices } v\in Y_n} a_n(v)\right|^2 = 0, \tag{I.4.1}$$

where

$$N(\Theta, Y_n) = \#\{j : \nu_j^{(n)} \in \Theta\}. \tag{I.4.2}$$

The notation $\mathrm{card}(\cdot)$ refers to the cardinality of a set. The codimension one exceptional locus Ξ is defined in Chapter IV.2.

Remark I.7. Theorem I.6 may be viewed simultaneously as a higher rank analogue of Theorem I.4 and a non-archimedean analogue of Theorem I.5. The assumption that $\mathrm{card}(Y_n) \to \infty$ is in fact known to be equivalent to Benjamini-Schramm convergence in this setting by Gelander-Levit [GL18]. No spectral gap assumption need be made as it is automatic by property (T) [BdlHV08]. Since we are assuming $\mathrm{char}(F) = 0$, no uniform discreteness assumption need be made [Mar91], and in fact all lattices are cocompact which implies that the underlying Y_n's are finite simplicial complexes [Ser80].

Remark I.8. The analogue of the exceptional locus (or, more precisely, of what is called Ξ_1 in Chapter IV.2) for the $(q+1)$-infinite regular tree is the set $\{-2\sqrt{q}, 2\sqrt{q}\}$, for the hyperbolic plane is $\{1/4\}$, and for $\mathrm{SL}(d, \mathbb{R})/\mathrm{SO}(d)$ is those points in $i\mathfrak{a}^*/W$ whose stabilizer is nontrivial (i.e. the non-regular points of a Weyl chamber). Notice that in Theorem I.4, subsets of the tempered spectrum meeting this exceptional locus are allowed. However, in all of the aforementioned works regarding symmetric spaces [LMS17, ABLM18, BM21], one does not allow subsets of the tempered spectrum intersecting the exceptional locus. We believe that we can in fact strengthen Theorem I.6 to not have the condition about avoiding Ξ, but such results are still in preparation. See also Chapter III.2 and (the proof of) Lemma III.1.

CHAPTER II
Outline of the Proof of Theorem I.6

II.1: Notation and summary of preliminaries

We now set the notation and summarize the preliminary material needed to outline the proof of Theorem I.6. This preliminary material is further discussed in Appendices A, B, and C. The relevant propositions and lemmas are combined to explain the proof of Theorem I.6 in Chapter II.3.

II.1.1: The group G and associated objects

We let $G = \mathrm{PGL}(3, F)$ and $K = \mathrm{PGL}(3, \mathcal{O})$ where F is a non-archimedean local field of characteristic zero and $\mathcal{O}$ is its ring of integers. We denote the Haar measure on G by $\mathrm{vol}(\cdot)$ and normalize it so that $\mathrm{vol}(K) = 1$. Let ϖ be a uniformizer of $\mathcal{O}$, and let q be the order of the residue field. We let $\pi : G \to G/K$ denote the obvious projection.

Let $\Gamma < G$ denote a torsionfree lattice, and (Γ_n) denote a sequence of torsionfree lattices whose covolume goes to infinity. Let ρ^Γ be the unitary G-representation corresponding to the right action on $L^2(\Gamma\backslash G)$. Let $E \subset G$ be a set with positive, finite Haar measure. Suppose $f \in L^2(\Gamma\backslash G)$. We define ρ^Γ_E to be the operator on $L^2(\Gamma\backslash G)$ such that

$$[\rho^\Gamma_E.f](\Gamma h) := \frac{1}{\mathrm{vol}(E)} \int_E f(\Gamma h g) dg. \tag{II.1.1}$$

We say that a function on G is (M_1, M_2)-invariant, with $M_1, M_2 < G$, if it is left-invariant under M_1 and right-invariant under M_2. The *spherical Hecke algebra* $H(G, K)$ is the algebra of compactly supported (K, K)-invariant functions on G with product structure given by convolution.

Let $T < G$ denote the maximal F-split torus consisting of all diagonal matrices, and let $A < T$ denote the subgroup of matrices all of whose diagonal entries are powers of ϖ. Let A^+

denote those elements in A whose diagonal powers of ϖ are weakly decreasing. By reading off the exponents of ϖ along the diagonal, we may associate to each element in A a tuple $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{Z}^3$ up to shifting all entries by the same integer. Elements in A^+ may be parametrized by partitions: $\lambda_1 \geq \lambda_2 \geq \lambda_3 = 0$. We define $\varpi^\lambda := \mathrm{diag}(\varpi^{\lambda_1}, \varpi^{\lambda_2}, \varpi^{\lambda_3})$.

II.1.2: Identification of A with a lattice in $\mathbb{R}^2$

Let $\mathfrak{a}$ denote the vector space

$$\mathfrak{a} := \{(x_1, x_2, x_3) : x_i \in \mathbb{R} \text{ and } x_1 + x_2 + x_3 = 0\}.$$

Then $\mathfrak{a}$ naturally sits inside of $\mathbb{R}^3$. Let $p : \mathbb{R}^3 \to \mathfrak{a}$ denote the orthogonal projection using the standard inner product on $\mathbb{R}^3$. Under p, the image of $A \simeq \mathbb{Z}^3/\langle(1,1,1)\rangle$ is a lattice Λ inside of $\mathfrak{a}$. The identity matrix in A gets mapped to the origin, and elements in A^+ correspond to the intersection of Λ with the *standard Weyl chamber* $\mathfrak{a}^+$, namely the image under p of the region $x_1 \geq x_2 \geq x_3$ in $\mathbb{R}^3$; we call this intersection Λ^+. We let $(\cdot, \cdot)$ denote the standard pairing between $\mathfrak{a}^*$ and $\mathfrak{a}$. We let $\delta = (1, 0, -1)$. If we think of $\mathfrak{a}$ as the collection of diagonal trace zero matrices in $\mathfrak{sl}(3)$ and we identify $\mathfrak{a} \simeq \mathfrak{a}^*$ using the coordinates on $\mathfrak{a}$ (and the standard dot product in these coordinates), then δ corresponds to half the sum of positive roots.

II.1.3: The Bruhat-Tits building $\mathcal{B}$

Let $\mathcal{B}$ denote the Bruhat-Tits building (as a simplicial complex) associated to G, and G/K denote the set of vertices of $\mathcal{B}$ induced by identifying some particular vertex in $\mathcal{B}$ with the coset $1K$. If $B \subset G/K$ is a set of vertices of $\mathcal{B}$, we let $\mathrm{card}(B)$ denote the cardinality of this set; this is equivalent to computing the Haar measure of $\pi^{-1}(B) \subset G$.

Let $Y = \Gamma\backslash G/K$ (and $Y_n = \Gamma_n\backslash G/K$, resp.) denote the vertices of the simplicial complex $\hat{Y} = \Gamma\backslash\mathcal{B}$ (and $\hat{Y}_n = \Gamma_n\backslash\mathcal{B}$, resp.). Let $\tau_\Gamma : G/K \to Y$ and $\hat{\tau}_\Gamma : \mathcal{B} \to \hat{Y}$ denote the obvious projections. Let $a \in L^\infty(Y)$ (and $a_n \in L^\infty(Y_n)$, resp.) be a test function with $||a||_\infty \leq 1$ (and $||a_n||_\infty \leq 1$, resp.). Let D (and D_n, resp.) be a fundamental domain for the action of Γ (and Γ_n, resp.) on G/K. Notice that functions on Y are the same as (Γ, K)-invariant functions on G.

The building $\mathcal{B}$ has a natural metric $d(\cdot, \cdot)$ which is invariant under the G-action. This metric descends to any $\hat{Y}$ (see Appendix B.6.7). This metric is normalized so that the distance between adjacent vertices is 1.

See also Appendix B.6.1.

II.1.4: Benjamini-Schramm convergence and injectivity radius

Recall that Y corresponds to the vertices of the simplicial complex $\hat{Y}$, which is a metric space. Suppose $y \in Y$. Let $\tilde{y} \in \mathcal{B}$ be any lift of y under $\hat{\tau}_\Gamma$. We define the injectivity radius of y, denoted $\mathrm{InjRad}_Y(y)$, to the supremum of all r such that the ball of radius r in $\mathcal{B}$ centered at $\tilde{y}$ (with respect to the metric $d(\cdot,\cdot)$) maps injectively under $\hat{\tau}_\Gamma$ to $\hat{Y}$. We define the *injectivity radius* of Y, denoted $\mathrm{InjRad}(Y)$, to be the supremum over all $y \in Y$ of $\mathrm{InjRad}_Y(y)$. We say that (Y_n) *Benjamini-Schramm converges to* G/K if, for every $R > 0$, we have

$$\frac{\mathrm{card}(\{y \in Y_n : \mathrm{InjRad}_{Y_n}(y) \leq R\})}{\mathrm{card}(Y_n)} \to 0 \tag{II.1.2}$$

as $n \to \infty$. The assumption that $\mathrm{card}(Y_n) \to \infty$ implies that (Y_n) Benjamini-Schramm converges to G/K [GL18].

II.1.5: The Weyl chamber-valued metric and polytopal balls

The Cartan decomposition (Proposition A.1) allows us to define a *Weyl chamber-valued metric* $d_{A^+}(\cdot,\cdot)$ on G/K taking values in A^+ (or equivalently, in Λ^+). Suppose $Q \subset \mathfrak{a}$ is a polytope. Let Q_m denote the mth dilate of Q, and let $Q_m^\Lambda := Q_m \cap \Lambda$ (when $m = 1$, we omit the subscript). We may naturally view any Q_m^Λ as subset of A, and hence also of G, in which case it makes sense to consider $KQ_m^\Lambda K \subset G$.

Given Q, we may define associated *polytopal balls*: given a vertex $v \in G/K$, we define the *Q-shaped ball at v* to be:

$$B_Q(v) := \{w \in G/K : d_{A^+}(v,w) \in Q^\Lambda\}.$$

We may also define a *polytopal norm* on $\mathfrak{a}^+$ induced by Q as follows:

$$|\lambda|_Q := \inf\{m \in \mathbb{R}_{\geq 0} : \lambda \in Q_m\}.$$

We shall often be interested in the ceiling of the polytopal norm which we denote

$$|\lambda|_Q^{\mathrm{ceil}} := \lceil |\lambda|_Q \rceil.$$

There are three polytopes which are of particular interest to us in this paper which we call P, P^*, and H; their definitions are given in Appendix C.6.1. The polytope P has a distinguished vertex

$$p^\dagger := (4/3, -2/3, -2/3).$$

See also Appendices B.6.6 and C.6.

II.1.6: Conventions and notations for Hilbert spaces

We denote inner products on Hilbert spaces by $\langle \cdot, \cdot \rangle$, and we follow the convention that inner products are $\mathbb{C}$-linear in the second entry and sesquilinear in the first entry.

Suppose U is an operator on a Hilbert space $\mathcal{H}$. Recall that

$$||U||_{\mathrm{HS}} := \big(\sum_j ||Ue_j||^2\big)^{1/2}$$

is the *Hilbert-Schmidt norm* of U, where $\{e_j\}$ is any orthonormal basis of $\mathcal{H}$.

Suppose $\mathcal{H} = L^2(X, \mu)$ where (X, μ) is a measure space. An operator U acting on $\mathcal{H}$ has a *kernel function* $\mathcal{K} : X \times X \to \mathbb{C}$ if

$$[U.f](x) = \int_X \mathcal{K}(x, y) f(y) d\mu(y)$$

for all $f \in \mathcal{H}$.

II.1.7: Tempered spectrum and Plancherel measure

We denote the tempered spectrum by Ω^+_{temp}, and let μ denote the Plancherel measure normalized so that $\mu(\Omega^+_{\mathrm{temp}}) = 1$. We let Ξ denote the exceptional locus.

See Appendices A.16, A.10 and A.11, as well as Chapters IV.1 and IV.2.

II.1.8: A word on the use of "$\lesssim$"

We use the notation $f \lesssim g$ to mean that an inequality is true up to a positive multiplicative constant. If both f and g depend on some parameter b, then the notation $f \lesssim_b g$ means that the inequality is true up to a positive multiplicative constant (not depending on b) for all b sufficiently large. We also use the notation $f \lesssim_{b,c} g$, where b and c are two different parameters (in the sequel, often $b, c = n, M$), to mean that we both have $f \lesssim_b g$ and $f \lesssim_c g$

independently, namely there exist b_0 and c_0 such that for all (b, c) such that $b \geq b_0$ and $c \geq c_0$, we have $f(b, c) \lesssim g(b, c)$.

II.2: Ancillary propositions and lemmas

In this subsection, we collect a sequence of Propositions and Lemmas which will allow us to prove Theorem I.6. In cases where these results have straightforward or non-technical proofs, we include them here; otherwise we defer the proof to a later chapter.

II.2.1: Reduction to the case of mean-zero test function

Given a test function a_n, we may subtract off its mean to get a mean-zero test function:

$$\bar{a}_n := a_n - \frac{1}{\operatorname{card}(Y_n)} \sum_{v \in Y_n} a_n(v).$$

This simplifies the summands in (I.4.1) to $|\langle \psi_j^{(n)}, \bar{a}_n \psi_j^{(n)} \rangle|^2$. Notice also that if $||a_n||_\infty \leq 1$, then $||\bar{a}_n||_\infty \leq 2$. Hence from now on we shall assume that a_n is a mean-zero test function such that $||a_n||_\infty \leq 2$.

II.2.2: Polytopal ball averaging operators

Let E_m denote the P_m-shaped ball centered at $1K$:

$$E_m := B_{P_m}(1K) = K P_m^\Lambda K.$$

We shall abuse notation and let

$$xE_m := B_{P_m}(x),$$

with $x \in G/K$. Notice that E_m is a K-invariant set in G/K and hence

$$\tilde{E}_m := \pi^{-1}(E_m)$$

is a (K, K)-invariant set in G. The function $\mathbb{1}_{\tilde{E}_m}$ is in $H(G, K)$ and hence acts on functions on G by right convolution (see Appendix B.7). Furthermore it preserves (Γ, K)-invariant functions, so it acts on functions on both G/K and on Y. We let U_m denote the operator corresponding to right convolution with $\mathbb{1}_{\tilde{E}_m}$. As explained in Appendix B.7.1, when U_m is applied to a function on G/K (or equivalently, a $(1, K)$-invariant function on G) we obtain:

$$[U_m.f](gK) = \sum_{hK \in g\tilde{E}_m^{-1}} f(hK).$$

Also note that

$$g\tilde{E}_m = \pi^{-1}(B_{P_m}(gK)).$$

The operator U_m has kernel function $\mathcal{K}_m : G \times G \to \mathbb{C}$ defined as

$$\mathcal{K}_m(g, h) = \begin{cases} 1 & g^{-1}h \in \tilde{E}_m^{-1} \\ 0 & \text{otherwise.} \end{cases} \tag{II.2.1}$$

As explained in Appendix C.6.1, the polytope P^* satisfies $\pi(\tilde{E}_m^{-1}) = B_{P_m^*}(1K) = K(P_m^*)^\Lambda K$. Hence U_m acting on $L^2(G/K)$ corresponds to summing up over the P_m^*-shaped ball centered at each vertex. From (II.2.1), it is clear that the formula for kernel function of the adjoint U_m^* is given by

$$\mathcal{K}_m^*(g, h) := \mathcal{K}_m(h, g) = \begin{cases} 1 & g^{-1}h \in \tilde{E}_m \\ 0 & \text{otherwise.} \end{cases} \tag{II.2.2}$$

Recall that a denotes a test function on Y. We identify a with the operator corresponding to multiplication by a. Suppose M is a positive integer. Consider the following operator on $L^2(Y)$:

$$A_M := \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\operatorname{card}(E_m)} U_m^* \circ a \circ U_m. \tag{II.2.3}$$

When we use a_n instead of a, we shall refer to the corresponding operator as A_M^n. Note that we may also consider A_M as an operator on functions on G.

If ψ_ν is an eigenfunction of the $H(G, K)$-action on $L^2(Y)$ with spectral parameter ν (see Appendix A.7), then $U_m\psi_\nu = h_m(\nu)\psi_\nu$ for some complex number $h_m(\nu)$. We then get that

$$\Big|\langle \psi_\nu, A_M\psi_\nu\rangle\Big|^2 = \Big(\frac{1}{M} \sum_{m=1}^{M} \frac{|h_m(\nu)|^2}{\operatorname{card}(E_m)}\Big)^2 \Big|\langle \psi_\nu, a\psi_\nu\rangle\Big|^2. \tag{II.2.4}$$

II.2.3: Spectral bound

Proposition II.1 (Spectral Bound; proof in Chapter IV.6). *Let $\Theta \subset \Omega^+_{\mathrm{temp}}$ be a compact subset of the tempered spectrum with positive Plancherel measure and not intersecting the exceptional locus Ξ. Then for all $\nu \in \Theta$,*

$$\frac{1}{M}\sum_{m=1}^{M}\frac{|h_m(\nu)|^2}{\operatorname{card}(E_m)} \gtrsim_M 1,$$

with the implied constant only depending on Θ and not on ν or M.

The proof of this proposition requires analysis of the spherical functions associated to each ν. The strategy is as follows: we first rearrange the relevant expression so that it essentially becomes the sum of an exponential function on the lattice points in a polytope. We then apply Brion's formula (see Appendix C.3) which in particular allows us to identify the dominating term. Finally we bound this dominating term asymptotically using a "linear independence of characters" argument. See Chapter IV.

II.2.4: Benjamini-Schramm convergence implies Plancherel convergence

Recall the definition of $N(\Theta, Y_n)$ in (I.4.2).

Proposition II.2 (BS Convergence Implies Plancherel Convergence; proof in Chapter V.3). *The distribution of spectral parameters associated to the $H(G,K)$-action on $L^2(Y_n)$ weak-* converges to the Plancherel measure μ as $n \to \infty$ if (Y_n) Benjamini-Schramm converges to G/K. In particular for any compact $\Theta \subseteq \Omega^+_{\mathrm{temp}}$, we have*

$$\lim_{n\to\infty}\frac{N(\Theta, Y_n)}{\operatorname{card}(Y_n)} = \mu(\Theta).$$

This proposition follows almost immediately from results of Deitmar [Dei18] regarding the relationship between Benjamini-Schramm convergence and convergence to the Plancherel measure. Similar results in the case of semisimple algebraic groups over local fields of characteristic zero appears in the work of Gelander-Levit [GL18].

The Plancherel measure is in fact a measure with respect to a certain topology on Ω^+_{temp} called the Fell topology (see Appendix A.5.4). This topology is closely related to a more obvious "Euclidean" topology on Ω^+_{temp}. The work of proving Proposition II.2 using the theorems of Deitmar and Gelander-Levit lies in relating these two topologies. In particular we show that if a set is compact in the Euclidean topology, then it is also compact and μ-regular in the Fell topology. See Chapter V.

II.2.5: Reduction to bounding the Hilbert-Schmidt norm

The expression $|\langle\psi_j^{(n)}, a_n\psi_j^{(n)}\rangle|^2$ is often referred to as the *quantum variance* (of a_n with respect to the wave function $\psi_j^{(n)}$).

Proposition II.3 (Quantum Variance Bounded by Hilbert-Schmidt Norm)**.** *Suppose $\Theta \subset \Omega^+_{\text{temp}}$ is compact and $\mu(\Theta) > 0$. Suppose (Y_n) Benjamini-Schramm converge to G/K. Then*

$$\frac{1}{N(\Theta, Y_n)} \sum_{\psi_j^{(n)}:\nu_j^{(n)}\in\Theta} \left|\langle\psi_j^{(n)}, a_n\psi_j^{(n)}\rangle\right|^2 \lesssim_{n,M} \frac{1}{\text{card}(Y_n)}||A_M^n||^2_{\text{HS}},$$

where the implied constant depends on Θ, but does not depend on n, M as long as both n and M are sufficiently large (recall the notation in Chapter II.1.8).

Proof. Note that by Cauchy-Schwarz $|\langle e_j, Ue_j\rangle|^2 \leq ||Ue_j||^2||e_j||^2$ for any operator U acting on some Hilbert space, and hence, if $\{e_j\}$ is an orthonormal basis of the underlying Hilbert space,

$$\sum_j |\langle e_j, Ue_j\rangle|^2 \leq ||U||^2_{\text{HS}}. \tag{II.2.5}$$

We thus have:

$$\begin{aligned}
\frac{1}{N(\Theta, Y_n)} \sum_{\psi_j^{(n)}:\nu_j^{(n)}\in\Theta} \left|\langle\psi_j^{(n)}, a_n\psi_j^{(n)}\rangle\right|^2 &\lesssim_M \frac{1}{N(\Theta, Y_n)} \sum_{\psi_j^{(n)}:\nu_j^{(n)}\in\Theta} \left|\langle\psi_j^{(n)}, A_M^n\psi_j^{(n)}\rangle\right|^2 \\
&\qquad\qquad \text{(by (II.2.4) and Prop. II.1)} \\
&\lesssim_n \frac{1}{\text{card}(Y_n)} \sum_{\psi_j^{(n)}:\nu_j^{(n)}\in\Theta} \left|\langle\psi_j^{(n)}, A_M^n\psi_j^{(n)}\rangle\right|^2 \\
&\qquad\qquad \text{(by Prop. II.2)} \\
&\leq \frac{1}{\text{card}(Y_n)} \sum_{\text{all } \psi_j^{(n)}} \left|\langle\psi_j^{(n)}, A_M^n\psi_j^{(n)}\rangle\right|^2 \\
&\leq \frac{1}{\text{card}(Y_n)}||A_M^n||^2_{\text{HS}} \\
&\qquad\qquad \text{(by (II.2.5))}.
\end{aligned}$$

$\square$

II.2.6: Bounding the Hilbert-Schmidt norm using the kernel function on G/K

Recall that when an operator has an associated kernel function, then its Hilbert-Schmidt norm may be computed by computing the L^2-norm of the kernel function. An operator acting on $L^2(Y)$ would then naturally have its associated kernel function as a function on $Y \times Y$. However, in many ways it's easier to work with functions on $G/K \times G/K$ because this space is homogeneous. If we have a function on $G/K \times G/K$ which is invariant under the diagonal Γ-action and is only supported near the diagonal, then by summing up over Γ, we may define an operator on $L^2(Y)$. The following lemma allows us to relate the Hilbert-Schmidt norm of this descended operator to the L^2-norm of the kernel function on $D \times G/K$ where D is a fundamental domain for the Γ-action on G/K. We also obtain an error term in terms of number of points in Y with small injectivity radius.

Lemma II.4 (Lifting the Kernel to G/K; proof in Chapter VI.2)**.** *Let $\mathcal{K} : G/K \times G/K \to \mathbb{C}$ be a function which is invariant under the diagonal Γ-action. Suppose $R \geq 0$ is such that $\mathcal{K}(z,w) = 0$ whenever $d(z,w) \geq R$. Let $\bar{\mathcal{K}}_Y^{\mathrm{Op}}$ denote the operator on $L^2(Y)$ defined by this kernel. Then there exist $C_1, C_2 > 0$, independent of R and Γ, such that*

$$||\bar{\mathcal{K}}_Y^{\mathrm{Op}}||_{\mathrm{HS}}^2 \leq \sum_{z \in D} \sum_{w \in G/K} |\mathcal{K}(z,w)|^2 \tag{II.2.6}$$

$$+ \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} ||\mathcal{K}||_\infty^2 \mathrm{card}(\{y \in Y : \mathrm{InjRad}_Y(y) \leq R\}). \tag{II.2.7}$$

Note that the term of the form $q^{C_2 R}$ comes from the "volume" of a ball of radius R in $\mathcal{B}$ (with respect to $d(\cdot,\cdot)$), and the term $\mathrm{InjRad}(Y)^2$ in the denominator comes from the area of a Euclidean ball of radius $\mathrm{InjRad}(Y)$ in $\mathbb{R}^2$. See Chapter VI.

II.2.7: Explicit formula for the kernel function of A_M

We ultimately wish to apply Lemma II.4 to the appropriate function $\mathcal{K} : G/K \times G/K \to \mathbb{C}$ so that $A_M = \bar{\mathcal{K}}_Y^{\mathrm{Op}}$.

Proposition II.5 (Formula for the Kernel Function)**.** *Let $\mathcal{L}_M : G \times G \to \mathbb{C}$ be defined by:*

$$\mathcal{L}_M(x,y) := \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\mathrm{card}(E_m)} \int_{x\tilde{E}_m \cap y\tilde{E}_m} a(z)dz. \tag{II.2.8}$$

Then $\mathcal{L}_M$ is invariant under the right $(K \times K)$-action and the left diagonal Γ-action and hence defines a function $\mathcal{L}'_M : G/K \times G/K \to \mathbb{C}$ as in Lemma II.4. The associated operator on $L^2(Y)$ is exactly A_M.

Proof. We first find the kernel function for $U_m^* \circ a \circ U_m$. Suppose f is a function on G. Then by (II.2.1) and (II.2.2),

$$[(a \circ U_m).f](z) = a(z) \int_{z\tilde{E}_m^{-1}} f(y)dy,$$
$$[U_m^*.g](x) = \int_{x\tilde{E}_m} g(z)dz,$$
$$[(U_m^* \circ a \circ U_m).f](x) = \int_{x\tilde{E}_m} a(z) \int_{z\tilde{E}_m^{-1}} f(y)dydz. \tag{II.2.9}$$

We now wish to change the order of integration. Suppose y is fixed. We then must consider those z such that $y \in z\tilde{E}_m^{-1}$ and $z \in x\tilde{E}_m$. But $y \in z\tilde{E}_m^{-1} \iff z \in y\tilde{E}_m$. Hence we need $z \in x\tilde{E}_m \cap y\tilde{E}_m$. Therefore we may rewrite (II.2.9) as

$$\int_G \left(\int_{x\tilde{E}_m \cap y\tilde{E}_m} a(z)dz \right) f(y)dy,$$

from which it is clear that the kernel function for $U_m^* \circ a \circ U_m$ is exactly

$$\mathcal{H}_m(x,y) := \int_{x\tilde{E}_m \cap y\tilde{E}_m} a(z)dz.$$

Thus by (II.2.3), it is clear that the kernel function for A_M is exactly

$$\mathcal{L}_M(x,y) := \frac{1}{M} \sum_{m=1}^{M} \frac{\mathcal{H}_m(x,y)}{\mathrm{card}(E_m)}.$$

If we replace (x,y) by (xk_1, yk_2) in (II.2.7), with $k_1, k_2 \in K$, then we instead integrate over $xk_1\tilde{E}_m \cap yk_2\tilde{E}_m$, but because $\tilde{E}_m$ is $(K,1)$-invariant, this is the same as $x\tilde{E}_m \cap y\tilde{E}_m$. Similarly, if we replace (x,y) in (II.2.7) by $(\gamma x, \gamma y)$ with $\gamma \in \Gamma$, we now integrate $a(z)$ over z in $\gamma x\tilde{E}_m \cap \gamma y\tilde{E}_m = \gamma(x\tilde{E}_m \cap y\tilde{E}_m)$, but this is the same as integrating $a(\gamma u) = a(u)$ over u in $x\tilde{E}_m \cap y\tilde{E}_m$. Hence $\mathcal{L}_M$ is invariant under the right $(K \times K)$-action and the left diagonal Γ-action. $\square$

If we replace a with a_n in the formula for $\mathcal{L}_M$, we instead write the kernel function as $\mathcal{L}_M^n$, and similarly we write $(\mathcal{L}_M^n)'$ for the associated function on $G/K \times G/K$.

II.2.8: Determining when two polytopal balls intersect

In addition to P, another polytope that is of interest is the polytope H which is defined in Appendix C.6.1. The following proposition explains its importance.

Proposition II.6 (When Polytopal Balls Intersect; proof in Chapter IX.5)**.** *Suppose $x, y \in G/K$. The polytopal balls xE_m and yE_m intersect if and only if $d_{A^+}(x,y) \in H_m^\Lambda$. Phrased another way, we have $xE_m \cap yE_m \neq \emptyset$ if and only if $|d_{A^+}(x,y)|_H^{\text{ceil}} \leq m$.*

The polytope H_M is contained in the restriction to $\mathfrak{a}^+$ of the ball of radius $(2\sqrt{3})M$.

Corollary II.7 (Kernel Function is Supported Near the Diagonal; proof in Chapter IX.5)**.** *We have $\mathcal{L}_M(z,w) = 0$ if $d(zK, wK) > (2\sqrt{3})M$.*

II.2.9: Bounding the error term from lifting the kernel to G/K

We wish to use Lemma II.4 (Lifting the Kernel Function to G/K) and Proposition II.5 (Explicit Formula for the Kernel Function) to reduce to analyzing a function on $G \times G$. However, there is an error term showing up in (II.2.7) which is handled by the following lemma.

Lemma II.8 (Bounding the Error Term from Lifting the Kernel Function)**.** *Suppose (Y_n) Benjamini-Schramm converges to G/K. Then, for every positive integer M and $\varepsilon > 0$, there exists an $n_{M,\varepsilon}$ such that for all $n \geq n_{M,\varepsilon}$, we have*

$$\frac{C_1 q^{C_2(2\sqrt{3})M}}{\mathrm{InjRad}(Y_n)^2} ||\mathcal{L}_M^n||_\infty^2 \frac{\mathrm{card}(\{y \in Y_n : \mathrm{InjRad}_{Y_n}(y) \leq (2\sqrt{3})M\})}{\mathrm{card}(Y_n)} \leq \varepsilon. \tag{II.2.10}$$

Proof. As mentioned in Remark I.7, we know that there exists a positive lower bound on $\mathrm{InjRad}(Y_n)$ for all n (i.e. uniform discreteness is automatic). Furthermore, by the explicit formula for $\mathcal{L}_M^n$ in Proposition II.5, it is clear that $||\mathcal{L}_M^n||_\infty \leq ||a_n||_\infty$ and we have $||a_n||_\infty \leq 2$ by assumption. Furthermore, by the definition of Benjamini-Schramm convergence (II.1.2), we know that for any fixed M,

$$\lim_{n\to\infty} \frac{\mathrm{card}(\{y \in Y_n : \mathrm{InjRad}_{Y_n}(y) \leq (2\sqrt{3})M\})}{\mathrm{card}(Y_n)} = 0.$$

By combining these observations we conclude that such an $n_{M,\varepsilon}$ exists. $\square$

II.2.10: Changing variables in the integral of the kernel function

We wish to rewrite the expression on the right hand side of (II.2.6). Rather than sum over pairs $(z, w) \in D \times G/K$, we instead group elements based on their relative positions, namely $d_{A^+}(z, w)$. For any given point $z \in G/K$ and $\lambda \in A^+$, there are exactly N_λ many w's such that $d_{A^+}(z, w) = \lambda$; N_λ is defined in (II.2.11). The shape of any given $xE_m \cap yE_m$ with $x, y \in G/K$ only depends on m and $\lambda = d_{A^+}(x, y)$. These intersections are exactly translations of the set $E_m^\lambda = E_m \cap \varpi^\lambda E_m$. Translating E_m^λ to be based at all the different vertices of Y and then integrating the test function a over these translated sets amounts to convolving the indicator function of the lift of E_m^λ to G (namely $\tilde{E}_m^\lambda$ defined in (II.2.12)) with the lift of a to a K-invariant function on $\Gamma\backslash G$ (recall the definition of ρ_E^Γ in (II.1.1)). Unraveling all of these observations ultimately results in the following proposition.

Proposition II.9 (Changing Variables in the Kernel Integral; proof in Chapter VII.2)**.** *We have*

$$\sum_{z\in D}\sum_{w\in G/K} |\mathcal{L}'_M(z,w)|^2 = \frac{1}{M^2}\sum_{\lambda\in A^+} N_\lambda \int_{\Gamma\backslash G} \left| \sum_{m=1}^{M} \frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)} [\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g)\right|^2 dg,$$

where

$$\begin{aligned} N_\lambda &:= \mathrm{vol}(K\varpi^\lambda K), && \text{(II.2.11)}\\ E_m^\lambda &:= E_m \cap \varpi^\lambda E_m, \\ \tilde{E}_m^\lambda &:= \{g\in G : d_{A^+}(1K, gK)\in P_m^\Lambda \text{ and } d_{A^+}(\varpi^\lambda K, gK)\in P_m^\Lambda\} && \text{(II.2.12)}\\ &= \pi^{-1}(E_m^\lambda). \end{aligned}$$

II.2.11: Changing the order of integration

A straightforward application of the Minkowski integral inequality allows us to change the order of integration in the integral for the L^2-norm of the kernel function.

Proposition II.10 (Changing the Order of Integration in the Kernel Integral)**.** *We have*

$$\sum_{\lambda\in A^+} N_\lambda \int_{\Gamma\backslash G}\left|\sum_{m=1}^{M}\frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)}[\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g)\right|^2 dg \le \sum_{\lambda\in H_M^\Lambda} N_\lambda \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M}\frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)}||\rho^\Gamma_{\tilde{E}_m^\lambda}.a||_{L^2(\Gamma\backslash G)}\Big)^2. \tag{II.2.13}$$

Proof. Because of Proposition II.6 (When Polytopal Balls Intersect) $\mathrm{card}(E_m^\lambda) \neq 0$ only if $m \geq |\lambda|_H$, so we may replace the sum on the right hand side of (II.2.13) with a sum starting from $m = |\lambda|_H^{\mathrm{ceil}}$. Furthermore we have $m \leq M$; thus in order to have $|\lambda|_H \leq M$, we must also have $\lambda \in H_M^\Lambda$. Therefore we can replace the sum over A^+ with simply the sum over H_M^Λ.

Recall that the Minkowski integral inequality says the following: suppose S_1 and S_2 are two measure spaces and $F(x,y)$ is a jointly measurable function. Then

$$\int_{S_2} \Big| \int_{S_1} F(x,y)dx \Big|^2 dy \leq \left(\int_{S_1} \Big| \int_{S_2} |F(x,y)|^2 dy \Big|^{1/2} dx \right)^2.$$

If we apply this to the right side of (II.2.13) with $S_1 = \{|\lambda|_H^{\mathrm{ceil}}, \ldots, M\}$ and $S_2 = \Gamma\backslash G$, we obtain

$$\int_{\Gamma\backslash G} \Big| \sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_\lambda^m)}{\mathrm{card}(E_m)} [\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g) \Big|^2 dg \leq \left(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)} \Big| \int_{\Gamma\backslash G} \big| [\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g) \big|^2 dg \Big|^{1/2} \right)^2$$
$$= \left(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)} \big|\big| \rho^\Gamma_{\tilde{E}_m^\lambda}.a \big|\big|_{L^2(\Gamma\backslash G)} \right)^2.$$

□

II.2.12: A Nevo-style ergodic theorem for G

On the right hand side of (II.2.13) we have an expression containing $||\rho^\Gamma_{\tilde{E}_m^\lambda}.a||$. The following proposition shows that we can bound this just in terms of the volume of $\tilde{E}_m^\lambda$ and the L^2-norm of a. This is quite remarkable as many drastically different sets in G have the same volume. A result of this form for semisimple Lie groups is due to Nevo [Nev98]. One of the ingredients in Nevo's proof is the Kunze-Stein phenomenon for semisimple Lie groups [KS60, Cow78]. The Kunze-Stein phenomenon was later shown by Veca [Vec02] to also hold for simply connected simple algebraic groups over non-archimedean local fields (such as $\mathrm{SL}(d,F)$). In Chapter VIII, we explain how this in turn implies that $\mathrm{PGL}(d,F)$ also has the Kunze-Stein phenomenon. We may then retrace Nevo's proof [Nev98] to obtain the following proposition. See also [GN10, GN15].

Proposition II.11 (Nevo-Style Ergodic Theorem; proof in Chapter VIII.4)**.** *Suppose a is a mean-zero function in $L^2(\Gamma\backslash G)$. There exist constants $\theta > 0$ and $C > 0$, not depending on* *a or Γ, such that for any $E \subset G$ with finite positive Haar measure,*

$$||\rho_E^\Gamma . a||_{L^2(\Gamma\backslash G)} \leq \frac{C}{\mathrm{vol}(E)^\theta}\ ||a||_{L^2(\Gamma\backslash G)}.$$

Corollary II.12 (Applying the Nevo-Style Ergodic Theorem)**.** *There exists a $\theta > 0$ such that*

$$\sum_{\lambda \in H_M^\Lambda} N_\lambda \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)} ||\rho^\Gamma_{\tilde{E}_m^\lambda} . a||_{L^2(\Gamma\backslash G)} \Big)^2 \lesssim ||a||_2^2 \sum_{\lambda \in H_M^\Lambda} N_\lambda \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)^{1-\theta}}{\mathrm{card}(E_m)} \Big)^2,$$

where the implied constant does not depend on Γ or M.

Proof. This follows immediately from Proposition II.11 (Nevo-Style Ergodic Theorem) using the fact that $\mathrm{vol}(\tilde{E}_m^\lambda) = \mathrm{card}(E_m^\lambda)$. □

II.2.13: Bounding the size of intersections of polytopal balls

Corollary II.12 (Applying the Nevo-Style Ergodic Theorem) allows us to now focus on bounding the size of N_λ and the cardinality of the sets E_m and $\tilde{E}_m^\lambda$. There is an explicit formula for N_λ due to MacDonald [Mac95] from which the following is an easy consequence.

Proposition II.13 (Upper Bound on N_λ; proof in Chapter IV.5)**.** *We have*

$$N_\lambda = \mathrm{vol}(K\varpi^\lambda K) \lesssim (q^2)^{(\delta,\lambda)},$$

where the implied constant does not depend on λ.

Recall that the polytope P has a distinguished vertex $p^\dagger = (4/3, -2/3, -2/3)$ and that $\delta = (1, 0, -1)$ is half the sum of positive roots. Note that $(\delta, p^\dagger) > 0$. Using the explicit formula for N_λ together with Brion's formula (Theorem C.3), one may show the following:

Proposition II.14 (Lower Bound on $\mathrm{card}(E_m)$; proof in Chapter IV.5)**.** *We have*

$$\mathrm{card}(E_m) \gtrsim (q^2)^{(\delta, m\cdot p^\dagger)},$$

where the implied constant does not depend on m.

The following proposition is the most difficult in the entire paper. It provides an upper bound on the size of E_m^λ.

Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$; proof in Chapter X.4)**.** *We have that*

$$\mathrm{card}(E_m^\lambda) \lesssim (q^2)^{(\delta, m\cdot p^\dagger - \frac{\lambda}{2})},$$

where the implied constant does not depend on m or λ.

The method of proof of Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$) goes roughly as follows:

(1) (Roughly) classify relative positions of triples of points in $\mathcal{B}$. See Chapter IX.4.

(2) Use the classification in Step (1) to "coordinatize" those possible z which appear in $xE_m \cap yE_m$ with $x, y \in G/K$. For each fixed m and $\lambda = d_{A^+}(x, y)$, the allowable collection of coordinates will correspond to lattice points in some higher-dimensional polytope $\mathcal{P}(m, \lambda)$. See Chapter X.2 and Proposition X.4.

(3) Count the number of z with a given set of coordinates as in Step (2). The number will be roughly equal to an exponential function (with base q^2) whose power is a linear functional in these coordinates. See Lemma X.3.

(4) Identify that the greatest value that the linear functional in Step (3) can take is exactly equal to $(\delta, m \cdot p^\dagger - \frac{\lambda}{2})$. See Lemma X.5

(5) Use Brion's formula (Theorem C.3) to bound the size of $xE_m \cap yE_m$ by summing the exponential function in Step (3) over all lattice points in $\mathcal{P}(m, \lambda)$ from Step (2). In fact Brion's formula as originally stated does not work because certain "degeneracies" can occur. Hence we instead derive in Appendix C.5.2 a so-called "degenerate Brion's formula."

(6) Notice that as m and $\lambda = d_{A^+}(x, y)$ change, the collection of polytopes arising in Step (2) belong to finitely many "families" of polytopes of the same "type" parametrized by (m, λ) (see Appendix C.2.1). Analyze the exact formula given by (degenerate) Brion's formula from Step (5) to uniformly bound the size of $xE_m \cap yE_m$ over all polytopes in a given family. See Chapter X.4.

We may now combine the bounds on N_λ (Proposition II.13), the size of E_m (Proposition II.14), and the size of E_m^λ (Proposition II.15) to obtain the following.

Corollary II.16 (Combining Bounds on N_λ, $\mathrm{card}(E_m)$, and $\mathrm{card}(E_m^\lambda)$)**.** *We have*

$$\sum_{\lambda \in H_M^\Lambda} N_\lambda \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)^{1-\theta}}{\mathrm{card}(E_m)} \Big)^2 \lesssim \sum_{\lambda \in H_M^\Lambda} (q^2)^{\theta(\delta, \lambda - 2|\lambda|_H \cdot p^\dagger)}, \tag{II.2.14}$$

where the implied constant does not depend on M.

Proof. By Propositions II.14 (Lower Bound on $\mathrm{card}(E_m)$) and II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$), we get that

$$\begin{aligned} \frac{\mathrm{card}(E_m^\lambda)^{1-\theta}}{\mathrm{card}(E_m)} &\lesssim \frac{(q^2)^{(\delta, m\cdot p^\dagger - \frac{\lambda}{2})(1-\theta)}}{(q^2)^{(\delta, m \cdot p^\dagger)}} \\ &= (q^2)^{-\theta(\delta, m\cdot p^\dagger)} (q^2)^{-(1-\theta)(\delta, \frac{\lambda}{2})}. \end{aligned}$$

So then

$$\Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{vol}(\tilde{E}_m^\lambda)^{1-\theta}}{\mathrm{card}(E_m)} \Big)^2 \lesssim (q^2)^{-(1-\theta)(\delta,\lambda)} \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} (q^2)^{-\theta(\delta, m\cdot p^\dagger)} \Big)^2. \tag{II.2.15}$$

We have

$$\begin{aligned} \sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} (q^2)^{-\theta(\delta, m \cdot p^\dagger)} &= \frac{(q^2)^{-|\lambda|_H^{\mathrm{ceil}}\theta(\delta, p^\dagger)} - (q^2)^{-(M+1)\theta(\delta,p^\dagger)}}{1 - (q^2)^{-\theta(\delta, p^\dagger)}} \\ &\lesssim (q^2)^{-|\lambda|_H^{\mathrm{ceil}}\theta(\delta,p^\dagger)} \\ &\lesssim (q^2)^{-|\lambda|_H \theta(\delta, p^\dagger)}, \end{aligned} \tag{II.2.16}$$

because $(\delta, p^\dagger) > 0$ and $|\lambda|_H^{\mathrm{ceil}} \geq |\lambda|_H$.

Combining Proposition II.13 (Upper Bound on N_λ) with (II.2.15) and (II.2.16), we obtain

$$\begin{aligned} \sum_{\lambda \in H_M^\Lambda} N_\lambda \Big(\sum_{m=|\lambda|_H^{\mathrm{ceil}}}^{M} \frac{\mathrm{card}(E_m^\lambda)^{1-\theta}}{\mathrm{card}(E_m)} \Big)^2 &\lesssim \sum_{\lambda \in H_M^\Lambda} (q^2)^{(\delta,\lambda)} \cdot (q^2)^{-(1-\theta)(\delta,\lambda)} \cdot (q^2)^{-2|\lambda|_H \theta(\delta, p^\dagger)} \\ &= \sum_{\lambda \in H_M^\Lambda} (q^2)^{\theta(\delta, \lambda - 2|\lambda|_H \cdot p^\dagger)}. \end{aligned}$$

$\square$

II.2.14: Bounding the sum over H_M^Λ

We now wish to apply Brion's formula yet again. However, the expression on the right hand side of (II.2.14) is not exactly an exponential function whose power is a linear functional in the coordinates of Λ. However, we may derive an explicit formula for $|\lambda|_H$. From this we observe that we may partition $\mathfrak{a}^+$ into two sub-polytopes such that on each one $|\lambda_H|$ is a linear functional. This allows us to split the sum on the right hand side of in (II.2.14) into two pieces and and apply (degenerate) Brion's formula on each piece to arrive at the following proposition.

Proposition II.17 (Bounding the Sum over H_M^Λ; proof in Chapter XI)**.** *We have*

$$\sum_{\lambda \in H_M^\Lambda} (q^2)^{\theta(\delta, \lambda - 2|\lambda|_H \cdot p^\dagger)} \lesssim M,$$

where the implied constant does not depend on M.

II.3: Proof of Theorem I.6

Proof of Theorem I.6 (Quantum Ergodicity in the BS Limit for $\mathrm{PGL}(3, F)$*).* We now combine all of the preceding propositions and lemmas. Recall from Theorem I.6 that we must show that

$$\frac{1}{N(\Theta, Y_n)} \sum_{\psi_j^{(n)} : \nu_j^{(n)} \in \Theta} \left| \langle \psi_j^{(n)}, a_n \psi_j^{(n)} \rangle \right|^2 \to 0 \tag{II.3.1}$$

as $n \to \infty$, where a_n is a mean-zero test function such that $||a_n||_\infty \leq 2$ (see Chapter II.2.1). By Proposition II.3 (Quantum Variance Bounded by Hilbert-Schmidt Norm), we know that to show (II.3.1) it suffices to show that along some sequence (n, M_n) with M_n eventually always larger than some specific M_0, we have

$$\frac{1}{\mathrm{card}(Y_n)} ||A_{M_n}^n||_{\mathrm{HS}} \to 0$$

as $n \to \infty$.

By Lemma II.4 (Lifting the Kernel to G/K), Proposition II.5 (Formula for the Kernel Function), and Corollary II.7 (Kernel Function is Supported Near the Diagonal) it suffices to show that for some such sequence (n, M_n) as above, we have both

$$\frac{1}{\operatorname{card}(Y_n)} \sum_{z \in D_n} \sum_{w \in G/K} |(\mathcal{L}^n_{M_n})'(z,w)|^2 \to 0, \tag{II.3.2}$$

$$\frac{C_1 q^{C_2 R}}{\operatorname{InjRad}(Y_n)^2} ||(\mathcal{L}^n_{M_n})'(z,w)||^2_\infty \frac{\operatorname{card}(\{y \in Y_n : \operatorname{InjRad}_{Y_n}(y) \leq 2\sqrt{3} M_n\})}{\operatorname{card}(Y_n)} \to 0. \tag{II.3.3}$$

We handle each piece separately. We first handle (II.3.2). Combining Propositions II.9 (Changing Variables in the Kernel Integral) and Proposition II.10 (Changing the Order of Integration in the Kernel Integral) with Corollary II.12 (Applying the Nevo-Style Ergodic Theorem), we get

$$\begin{aligned}
&\frac{1}{\operatorname{card}(Y_n)} \sum_{z \in D_n} \sum_{w \in G/K} |(\mathcal{L}^n_{M_n})'(z,w)|^2 \\
&\qquad = \frac{1}{\operatorname{card}(Y_n)} \frac{1}{M^2} \sum_{\lambda \in A^+} N_\lambda \int_{\Gamma \backslash G} \left| \sum_{m=1}^{M} \frac{\operatorname{card}(E^\lambda_m)}{\operatorname{card}(E_m)} [\rho^\Gamma_{\tilde{E}^\lambda_m}.a](\Gamma g) \right|^2 dg \\
&\qquad \leq \frac{1}{\operatorname{card}(Y_n)} \frac{1}{M^2} \sum_{\lambda \in H^\Lambda_M} N_\lambda \Big(\sum_{m=|\lambda|^{\text{ceil}}_H}^{M} \frac{\operatorname{card}(E^\lambda_m)}{\operatorname{card}(E_m)} ||\rho^\Gamma_{\tilde{E}^\lambda_m}.a||_{L^2(\Gamma\backslash G)} \Big)^2 \\
&\qquad \lesssim \frac{||a_n||^2_2}{\operatorname{card}(Y_n)} \frac{1}{M^2} \sum_{\lambda \in H^\Lambda_M} N_\lambda \Big(\sum_{m=|\lambda|^{\text{ceil}}_H}^{M} \frac{\operatorname{card}(E^\lambda_m)^{1-\theta}}{\operatorname{card}(E_m)} \Big)^2.
\end{aligned}$$

Notice that

$$\frac{||a_n||^2_2}{\operatorname{card}(Y_n)} \leq ||a_n||^2_\infty.$$

Hence by Corollary II.16 (Combining Bounds on N_λ, $\operatorname{card}(E_m)$, and $\operatorname{card}(E^\lambda_m)$) and Proposition II.17 (Bounding the Sum over H^Λ_M) we have

$$\begin{aligned}
\frac{||a_n||^2_2}{\operatorname{card}(Y_n)} \frac{1}{M^2} \sum_{\lambda \in H^\Lambda_M} N_\lambda \Big(\sum_{m=|\lambda|^{\text{ceil}}_H}^{N} \frac{\operatorname{card}(E^\lambda_m)^{1-\theta}}{\operatorname{card}(E_m)} \Big)^2 &\lesssim \frac{1}{M^2} \sum_{\lambda \in H^\Lambda_M} (q^2)^{\theta(\delta, \lambda - 2|\lambda|^{\text{ceil}}_H \cdot p^\dagger)} \\
&\lesssim \frac{1}{M}.
\end{aligned}$$

Therefore (II.3.2) is true as long as $M \to \infty$.

Lastly we may bound (II.3.3) using Lemma II.8 (Bounding the Error Term from Lifting the Kernel Function): let $M^{(k)}$ be some sequence going to infinity. By Lemma II.8, for every k we can find an $n_{M^{(k)}}$ such that for every $n \geq n_{M^{(k)}}$, the expression in (II.2.10) holds with $\varepsilon = \frac{1}{k}$. Let n_k be some increasing sequence such that $n_k \to \infty$ and $n_k \geq n_{M^{(k)}}$. Let $\lambda(n) = \sup\{k : n \geq n_k\}$. Notice then that $\lambda(n) \to \infty$ as $n \to \infty$. Let $M_n = M^{(\lambda(n))}$. Then, since $n \geq n_{\lambda(n)} \geq n_{M^{(\lambda(n))}}$, we have that

$$\frac{C_1 q^{C_2 R}}{\operatorname{InjRad}(Y_n)^2} ||(\mathcal{L}^n_{M_n})'(z,w)||^2_\infty \frac{\operatorname{vol}(\{y \in Y_n : \operatorname{InjRad}_{Y_n}(y) \leq 2\sqrt{3}M_n\})}{\operatorname{vol}(Y_n)} \leq \frac{1}{\lambda(n)}.$$

Hence for this choice of M_n, we have this quantity going to zero as $n \to \infty$. □

II.4: Some analogies between the style of proof of Theorem I.6 and Theorem I.1

II.4.1: Sketch of the proof of Quantum Ergodicity in the Large Eigenvalue Limit

We now sketch the proof of the original quantum ergodicity theorem (Theorem I.1) so that we may point out its similarities to the proof of quantum ergodicity in the Benjamini-Schramm limit. First suppose $a \in C^\infty(M)$ is a mean-zero test function. Let Op_h denote a quantization scheme parametrized by a small parameter h which we may interpret as "Planck's constant". Let $\{\psi_n\}$ be the eigenfunctions of Δ with eigenvalues $\{\lambda_n\}$ ordered by magnitude. Let h_n be such that $\lambda_n h_n^2 = 1$. Let $p(x,\xi) = |\xi|^2$, with $x \in M$ and $\xi \in T_x M$, be the symbol of the Laplacian/Hamiltonian function generating the geodesic flow.

(1) Use the approximate functional calculus to construct a compactly supported symbol $\hat{a}$, i.e. a function in $C_0^\infty(T^*M)$, such that, for each eigenfunction ψ_j with eigenvalue at most λ_n, we have

$$\langle \psi_j, a\psi_j \rangle = \langle \psi_j, \operatorname{Op}_{h_n}(\hat{a})\psi_j \rangle + O(h_n).$$

Then use the unitarity of the operators $e^{it\Delta}$ to obtain that

$$\langle \psi_j, \operatorname{Op}_{h_n}(\hat{a})\psi_j \rangle = \langle \psi_j, \frac{1}{T}\int_0^T e^{-ith_n\Delta}\operatorname{Op}_{h_n}(\hat{a})e^{ith_n\Delta} dt\ \psi_j \rangle,$$

for all $T > 0$. This allows us to reduce to analyzing a two-parameter family of operators depending on n and T.

(2) Use the Egorov theorem to obtain that

$$\langle \psi_j, \frac{1}{T}\int_0^T e^{-ith_n\Delta}\mathrm{Op}_{h_n}(\hat{a})e^{ith_n\Delta}dt\ \psi_j\rangle = \langle \psi_j, \mathrm{Op}_{h_n}([\hat{a}]_T)\psi_j\rangle + O_T(h_n),$$

where

$$[\hat{a}]_T := \frac{1}{T}\int_0^T \Phi_t.\hat{a}\ dt$$

with Φ_t denoting the geodesic flow on symbols. This allows us to incorporate the underlying $\mathbb{R}$-action into the analysis.

(3) Use the approximate homomorphism between the (Poisson) algebra of symbols and the algebra of operators on $L^2(M)$, together with Cauchy-Schwarz, to obtain

$$|\langle \psi_j, \mathrm{Op}_{h_n}([\hat{a}]_T)\psi_j\rangle|^2 \leq \langle \psi_j, \mathrm{Op}_{h_n}(|[\hat{a}]_T|^2)\psi_j\rangle + O_T(h_n).$$

This allows us to focus on bounding the operator $\mathrm{Op}_{h_n}(|[\hat{a}]_T|^2)$.

(4) Use the local Weyl law to obtain that, for fixed T:

$$\lim_{n\to\infty}\frac{1}{N(\lambda_n)}\sum_{h_n^2\lambda_j\leq 1}|\langle \psi_j, \mathrm{Op}_{h_n}(|[\hat{a}]_T|^2)\psi_j\rangle|^2 = C\int_{p^{-1}([0,1])}|[\hat{a}]_T|^2 dx d\xi,$$

for an appropriate constant C.

(5) Use the ergodicity of the geodesic flow (and the fact that a and thus $\hat{a}$ is mean-zero) to obtain that, as $T\to\infty$,

$$|[\hat{a}]_T|^2 \to 0,$$

on every level set of p.

II.4.2: Analogous steps in Quantum Ergodicity in the Benjamini-Schramm Limit

All of the steps listed above have analogues in this proof. The role of h is replaced by the "distance" of each Y_n from the universal cover; namely we will have approximate equalities and inequalities up to an error that goes to zero under the assumption of Benjamini-Schramm convergence.

(1) The analogue of Step (1) is Proposition II.1 (Spectral Bound). Instead of considering the $\mathbb{R}$-parametrized Schrödinger flow (namely conjugation by $e^{ith_n\Delta}$), we instead consider the $\mathbb{N}$-parametrized "wave propagation" corresponding to polytopal ball averaging operators of different radii. The spectral bound tells us that A_M^n roughly preserves the mass of the eigenfunctions with spectral parameter in Θ; this may be seen as a substitute of the unitarity of the Schrödinger propagator.

(2) In Lemma II.4 (Lifting the Kernel to G/K) and Proposition II.9 (Changing Variables in the Kernel Integral), we lift to the universal cover up to an error which goes to 0 under Benjamini-Schramm convergence, then perform a change of variables and a change in the order of integration. This allows us to incorporate the G-action on $\Gamma\backslash G$ into the analysis and then ultimately to reduce to bounding the operator norms of certain elements in $L^1(G)$ acting on $L^2(\Gamma\backslash G)$. This may be seen as a rough substitute for the Egorov theorem in Step (2) above which also relates the operator from Step (1) to the underlying group action (the geodesic flow).

(3) In Proposition II.3 (Quantum Variance Bounded by Hilbert-Schmidt Norm), we reduce to bounding the Hilbert-Schmidt norm of an appropriate operator which is somewhat analogous to Step (3) above.

(4) A special case of the local Weyl law is the classical Weyl law telling us that the number of eigenvalues satisfying $h_n^2\lambda_j \leq 1$ is asymptotically equal to the volume of $p^{-1}([0,1]) \subset T^*M$ with respect to the Liouville measure. This aspect of the local Weyl law is analogous to Proposition II.2 (BS Convergence Implies Plancherel Convergence) which tells us that the distribution of spectral parameters asymptotically agrees with the distribution of spectral parameters on the universal cover (which is controlled by the Plancherel measure).

(5) In Corollary II.12 (Applying the Nevo-Style Ergodic Theorem), we utilize the Nevo ergodic theorem which allows us to relate the spectral gap of the underlying G-action on $L^2(\Gamma\backslash G)$ with the decay of operator norms of certain operators coming from functions in $L^1(G)$. This may be seen as the analogue of the use of the ergodicity of the geodesic flow in Step (5) above.

II.5: The polytope P vs. the polytope used in Brumley-Matz

We use the polytope P to define our polytopal ball averaging operators. Brumley-Matz [BM21] define analogous polytopal ball averaging operators but using a polytope which, somewhat conincidentally, is essentially identical to the polytope that we call H. The properties that P has which are used critically at various steps in the proof of Theorem I.6 (Main Theorem) are:

(1) The vertices of P are lattice points in Λ. This is necessary to be able to apply Brion's formula such as is done in the Proposition II.1 (Spectral Bound) and Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$).

(2) There is a unique vertex $p^\dagger$ maximizing the dot product with δ. This in particular implies that the volume of our polytopal ball averaging operators grows like $(q^2)^{m(\delta,p^\dagger)}$ as opposed to some polynomial times such an expression. This is used in the spectral bound and the geometric bound. This is in contrast to metric balls which do not have this property. See also (IV.4.1) and the discussion in Section 2.5 of Brumley-Matz [BM21].

(3) The vertex $p^\dagger$ is on the boundary of $\mathfrak{a}^+$. This is the property that distinguishes our polytope from the one used by Brumley-Matz. This property is used critically in the proof of Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$). See also Remark X.5.

(4) The orbit of P under $\mathfrak{S}_3$ results in a convex polytope in $\mathfrak{a}$. This is used critically in the proof of Proposition II.6 (When Polytopal Balls Intersect) which shows that whether or not two polytopal balls intersect is controlled by yet another convex polytope, namely H.

CHAPTER III
The Case of Regular Graphs

III.1: Previous results regarding quantum ergodicity on regular graphs

III.1.1: Work prior to Anantharaman-Le Masson [ALM15]

Notions of quantum ergodicity (and delocalization of eigenfunctions) on graphs appeared prior to the work of Anantharaman-Le Masson [ALM15]. In the physics literature, it was noted in [KS97] that regular graphs provide a good toy model for exploring quantum ergodicity. A form of quantum ergodicity for certain families of "quantum graphs" was shown in [BKS07]. On the other hand, Brooks-Lindenstrauss [BL13] proved delocalization of all eigenfunctions on large regular graphs using harmonic analysis on regular trees/regular graphs. That paper has overlap with the tools used in Brooks-Lindenstrauss-Le Masson [BLML16] to reprove Theorem I.4; this latter paper contains the original inspiration for many of the tools used in the sequel.

III.1.2: The work of Anantharaman-Le Masson [ALM15]: spherical function quantum ergodicity

The work of Anantharaman-Le Masson [ALM15] in fact proves a stronger result than Theorem I.4. Every eigenvalue of the adjacency operator on any $(q+1)$-regular graph lies in the interval $[-(q+1), q+1]$. Given a value $\lambda \in [-(q+1), q+1]$ and a choice of distinguished vertex v_0 on the infinite regular tree $\mathcal{T}_{q+1}$, there is a distinguished eigenfunction of the adjacency operator on $\mathcal{T}_{q+1}$ known as the spherical function associated to λ and denoted Φ_λ (note that if we write $\lambda = q^{\frac{1}{2}+s} + q^{\frac{1}{2}-s}$, then $\Phi_\lambda(x) = \omega_s(\varpi^{(d(v_0,x),0)})$ using the notation of (A.3.3) with $G = \mathrm{PGL}(2, F)$). This function has the property that $\Phi_\lambda(v_0) = 1$ and its value at a vertex x only depends on $d(v_0, x)$; in fact these properties uniquely define Φ_λ. Because of the radial invariance, we may define

$$\hat{\Phi}_\lambda : \mathbb{N} \to \mathbb{C}, \qquad k \mapsto \Phi(v_0, x) \text{ with } d(v_0, x) = k.$$

Now suppose (G_n) is a sequence of regular graphs satisfying the hypotheses of Theorem I.4. Suppose for each n we have a function $\mathcal{K}_n : V_n \times V_n \to \mathbb{C}$ such that there exists a D, independent of n, such that $\mathcal{K}_n(x, y) = 0$ if $d(x, y) > D$, and $|\mathcal{K}_n(x, y)| \leq 1$ for all x, y. Let $\mathcal{K}_n^{\mathrm{Op}}$ denote the operator on $L^2(V_n)$ whose kernel function is given by $\mathcal{K}_n$. Suppose $\{\psi_j^{(n)}\}$ are the eigenfunctions of the adjacency operator acting on $L^2(V_n)$ with associated eigenvalues $\{\lambda_j^{(n)}\}$. We then have that

$$\langle \psi_j^{(n)}, \mathcal{K}_n^{\mathrm{Op}} \psi_j^{(n)} \rangle = \sum_{x,y \in V_n} \mathcal{K}_n(x, y) \overline{\psi_j^{(n)}(x)} \psi_j^{(n)}(y).$$

Hence we can think of $\mathcal{K}_n^{\mathrm{Op}}$ as providing a means to test the correlation of eigenfunctions at different vertices whose distance is at most D apart.

On the other hand, we can compare this correlation with the correlation of the spherical function $\Phi_{\lambda_j^{(n)}}$ at different vertices. Consider the following expression:

$$\begin{aligned} \langle \mathcal{K}_n \rangle_\lambda :&= \frac{1}{\mathrm{card}(G_n)} \sum_{x,y \in V_n} \mathcal{K}_n(x, y) \hat{\Phi}_\lambda(d(x, y)) \\ &= \frac{1}{\mathrm{card}(G_n)} \sum_{x \in V_n} \sum_{y \in V_n \text{ s.t. } d(x,y) \leq D} \mathcal{K}_n(x, y) \hat{\Phi}_\lambda(d(x, y)). \end{aligned} \tag{III.1.1}$$

Let $x \in V_n$ and suppose for simplicity that the injectivity radius of x is at least D. Let $\tau : \mathcal{T}_{q+1} \to G_n$ be a covering map sending v_0 to x. Then τ is invertible on a ball of radius D centered at v_0. The inner sum in (III.1.1) thus expresses what the correlation is of the function $\Phi_\lambda \circ \tau^{-1}$ between x and all other vertices in the ball of radius D centered at x as "tested" via $\mathcal{K}_n(x, y)$ (treating Φ_λ as being centered at x). Then the entire expression in (III.1.1) is the average of this procedure over all vertices in G_n.

Anantharaman and Le Masson [ALM15] showed that

$$\lim_{n \to \infty} \frac{1}{N(I, G_n)} \sum_{\psi_j^{(n)} : \lambda_j^{(n)} \in I} \left| \langle \psi_j^{(n)}, \mathcal{K}_n^{\mathrm{Op}} \psi_j^{(n)} \rangle - \langle \mathcal{K}_n \rangle_{\lambda_j^{(n)}} \right|^2 = 0, \tag{III.1.2}$$

assuming the hypotheses in Theorem I.4 hold with the conditions on a_n now replaced with the conditions given above on $\mathcal{K}_n$. This expression going to zero implies that, for large graphs and in a weak sense, the eigenfunctions $\psi_j^{(n)}$ "look like" the spherical function with the same eigenvalue, which is some sense in the "canonical" choice of eigenfunction with that eigenvalue on $\mathcal{T}_{q+1}$. If we restrict to the case of $D = 0$, then $\mathcal{K}_n$ is only non-zero along its diagonal, in which case $\mathcal{K}_n^{\mathrm{Op}}$ is equivalent to multiplication by a test function a_n on the vertices. Hence we can recover (I.2.1) from (III.1.2).

As explained in [Ana17], there's a sense in which we should think of these operators $\mathcal{K}_n$ as coming from functions on non-backtracking paths of length at most D in G_n. The quantity (III.1.1) then provides a "microlocal lift" of the measure $|\psi_j^{(n)}|^2$ from G_n to this "path space." We can think of such path spaces as approximations to the "unit tangent bundle" of G_n, namely the space of (rooted and directed) infinite paths in G_n, which has on it an ergodic "geodesic flow". Hence (III.1.2) is more in line with the full version of the Quantum Ergodicity Theorem I.1 which concerns not only the eigenfunction measures on the manifold M, but also their microlocal lifts to the unit tangent bundle S^*M (see Remark I.2). In fact the proof of (III.1.2) originally given in [ALM15] utilized a "microlocal calculus" on trees developed in [LM14].

III.1.3: The work of Brooks-Le Masson-Lindenstrauss

Brooks-Le Masson-Lindenstrauss [BLML16] reproved Theorem I.4 using a new technique. Their techniques did not recover the full strength of (III.1.2), but the core ideas of their technique have subsequently been applied to many other contexts, namely hyperbolic surfaces [LMS17], rank one locally symmetric spaces [ABLM18], locally symmetric spaces associated to $\mathrm{SL}(d, \mathbb{R})/\mathrm{SO}(d)$ [BM21], and, in this work, compact quotients of the Bruhat-Tits building associated to $\mathrm{SL}(3, F)$. One of the main ideas is to use a sort of "wave propagator" on regular graphs which roughly preserves the mass of eigenfunctions. This property allows one to instead analyze the kernel function of certain operators derived from this wave propagator. See also Chapter III.2.

III.1.4: The work of Nelson [Nel18]

An approach to quantum (unique) ergodicity on regular graphs via p-adic representation theory can be found in Nelson [Nel18]. He consider a fixed finite regular graph Y arising as $\Gamma\backslash\mathrm{GL}(2, \mathbb{Q}_p)/\mathrm{GL}(2, \mathbb{Z}_p)$ for a (necessarily cocompact) arithmetic lattice Γ. Similarly to the discussion in Chapter III.1.2, for each N he considers the space Y_N of non-backtracking paths of length $2N$ on Y. For each prime $\ell \neq p$, there is an associated Hecke correspondence

T_ℓ coming from the arithmetic structure which we may use to put a graph structure on this path space. These operators are analogous to the Hecke correspondences on arithmetic hyperbolic surfaces which were used crucially in Lindenstrauss' proof of quantum unique ergodicity [Lin06]. We may think of these spaces Y_N as providing better and better approximations to the space $\Gamma\backslash\mathrm{GL}(2,\mathbb{Q}_p)$ which essentially corresponds to the collection of bi-infinite non-backtracking paths on Y and which carries an action of the diagonal subgroup of $\mathrm{GL}(2,\mathbb{Q}_p)$; geometrically this action essentially corresponds to the geodesic flow on Y. This space is the inverse limit as $N \to \infty$ of the Y_N's. Given an eigenfunction ϕ of all the Hecke correspondences on Y_N (more precisely a "newvector", see Definition 1 in [Nel18]), its pullback to $\Gamma\backslash\mathrm{GL}(2,\mathbb{Q}_p)$ generates an irreducible subrepresentation under the $\mathrm{GL}(2,\mathbb{Q}_p)$-action. Furthermore, we may pullback the measure $|\phi|^2$ on Y_N to $\mathrm{GL}(2,\mathbb{Q}_p)$. Nelson proves that, given a sequence of such eigenfunctions on Y_N as $N \to \infty$ whose pullbacks each generate subrepresentations belonging to the "principal series", then the only "quantum limit" of the associated pullback measures is the Haar measure on $\Gamma\backslash\mathrm{GL}(2,\mathbb{Q}_p)$. Along the way he constructs a representation theoretic "microlocal lift" which, asymptotically as $N \to \infty$, agrees with the above-described lifts of eigenfunction measures on Y_N to $\Gamma\backslash\mathrm{GL}(2,\mathbb{Q}_p)$. This is analogous to representation theoretic microlocal lifts that were considered in [Zel87, Wol01, Lin06, SV07].

III.2: "Wave propagators" vs. (polytopal) ball averaging operators

Many of the techniques of this paper may be seen to have their origins in [BLML16] which concerned regular graphs of any degree. However, for the sake of simplifying the following discussion, we shall only focus on the case of $(q+1)$-regular graphs where q is a power of a prime. In [BLML16] the authors do not work with operators which correspond to summing up over a ball of radius m but instead consider the operators obtained by plugging the adjacency operator into the mth Chebyshev polynomial of the first kind. Suppose C_m is the mth Chebyshev polynomial of the first kind and $\mathcal{A}$ is the adjacency operator of some finite $(q+1)$-regular graph. Let ψ be an eigenfunction of $\mathcal{A}$, and hence also an eigenfunction of the entire spherical Hecke algebra when we identify the $(q+1)$-regular tree with $\mathrm{PGL}(2,F)/\mathrm{PGL}(2,\mathcal{O})$ and identify ψ with a function on $\mathrm{PGL}(2,F)$ in the usual way. Suppose ψ has Satake parameters (q^s, q^{-s}) with $s \in (i\mathbb{R})/(\frac{2\pi i}{\ln(q)}\mathbb{Z})$; this implies that

$\mathcal{A}\psi = \sqrt{q}(q^s + q^{-s})$. The defining property of C_m implies

$$C_m\Big(\frac{\mathcal{A}}{\sqrt{q}}\Big)\psi = (q^{ms} + q^{-ms})\psi. \tag{III.2.1}$$

This remarkable spectral property makes the subsequent analysis in the spectral bound of the graph case simpler because, when s is fixed and m varies, the eigenvalue associated to $C_m(\frac{\mathcal{A}}{\sqrt{q}})$ is a sum of exponentials each of whose exponents depend linearly on m (see Lemma 2.1 in [BLML16]).

An analogous spectral property occurs when one uses the generalized Chebyshev polynomials of the first kind. These polynomials C_λ are indexed by partitions $\lambda = (\lambda_1, \dots, \lambda_d = 0)$. When $G = \mathrm{PGL}(d, F)$ and $K = \mathrm{PGL}(d, \mathcal{O})$, there is a natural generating set for $H(G, K)$ (as an algebra) given by $\mathcal{A}_1, \dots, \mathcal{A}_{d-1}$ (see Appendix B.7.2). If ψ is an eigenfunction of $H(G, K)$ with Satake parameters $(q^{s_1}, \dots, q^{s_d})$ with $s_j \in (i\mathbb{R})/(\frac{2\pi i}{\ln(q)}\mathbb{Z})$ and $s_1 + \dots + s_d = 0$, then we have that (see Proposition 2.1 in [LSV05b]):

$$\mathcal{A}_k\psi = q^{k(d-k)/2}\tau_k(q^{s_1}, \dots, q^{s_d})$$

where τ_k is the kth elementary symmetric function. Then the defining property of the generalized Chebyshev polynomials of the first kind (see Section 6.1 in [Bee91]) tells us that:

$$C_\lambda\Big(\frac{\mathcal{A}_1}{q^{(d-1)/2}}, \frac{\mathcal{A}_2}{q^{2(d-2)/2}}, \dots, \frac{\mathcal{A}_{d-1}}{q^{(d-1)/2}}\Big)\psi = (\sum_{\sigma \in \mathfrak{S}_d} q^{(\sigma.\lambda, s)})\psi. \tag{III.2.2}$$

Hence, when s is fixed and λ varies, we obtain that the associated eigenvalue is a sum of exponentials whose exponent is a linear functional in λ.

In Lemma 3.1 of [BLML16] the authors give a geometric interpretation of the operator in (III.2.1); it involves a sort of weighted sum over a ball of radius m but with some weights positive and others negative. The analysis required in the geometric bound is dependent on the initial choice of "wave propagator". The authors are able to perform the requisite analysis by direct analysis of the geometry of the regular tree (they also do not use a Nevo-style ergodic theorem, relying instead on a direct spectral bound on the relevant operators by passing to an analysis of the non-backtracking random walk).

In the work of [LMS17, ABLM18, BM21], the "wave propagator" used corresponds to summing up over some sort of (potentially "polytopal") ball. This is the approach that we have also used. However, it would be interesting to try to carry out the argument instead using the operators defined in (III.2.2). This would simplify the proof of the spectral bound, but would result in new complications in the geometric bound; in fact it is not clear what the geometric interpretation of these operators is. Given that the operators defined in (III.2.1) have been interpreted as the "right" analogue of wave propagation on regular graphs, and their utility in the proof of quantum unique ergodicity [Lin06, BL14], further analyzing the operators in (III.2.2) may be of interest in their own right and may connect to "wave propagation" on buildings.

III.3: The metric ball in a tree as a "polytopal ball"

Consider the Bruhat-Tits tree $\mathcal{T}_{q+1}$ associated to $G = \mathrm{PGL}(2, F)$ (see Appendix B.8). Let $K = \mathrm{PGL}(2, \mathcal{O})$. Recall that $A^+ < G$ is those matrices of the form $\mathrm{diag}(\varpi^{\lambda_1}, \varpi^{\lambda_2})$ with $\lambda_1 \geq \lambda_2$. Because we may shift both entries of (λ_1, λ_2) by the same integer amount and obtain the same element in G, we may parametrize A^+ via $\mathbb{N}$ with $\lambda \in \mathbb{N}$ corresponding to $\mathrm{diag}(\varpi^{\lambda}, 1) \in A^+$. The Weyl-chamber valued metric now takes values in $\mathbb{N}$ and in fact exactly agrees with the Euclidean metric on $\mathcal{T}_{q+1}$ (normalized so that adjacent vertices are distance 1 apart).

We may in turn identify A with lattice points $\mathbb{Z} \cdot (1/2, -1/2) \subset \mathfrak{a}$ with

$$\mathfrak{a} = \{(x_1, x_2) : x_1 + x_2 = 0\}.$$

We let $\mathfrak{a}^+ = \{(x_1, x_2) \in \mathfrak{a} : x_1 \geq x_2\}$. Then we may identify elements in A^+ with elements of the form $\mathbb{N} \cdot (1/2, -1/2)$.

An alternate way of coordinatizing $\mathfrak{a}$ is in terms of the basis $\{(1/2, -1/2)\}$ in which case the lattice in question is simply $\mathbb{Z}$; we call this coordinatization the *cone coordinates*, and the original coordinatization the $\mathfrak{a}$-*coordinates*. The interval $P_m = [0, m]$ in cone coordinates is a convex lattice polytope and we may take the P_m-shaped ball centered at $1K \in G/K$. This is exactly the same as the (vertices of the) metric ball in $\mathcal{T}$ of radius m centered at $1K$.

III.4: Spectral bound for averaging over a ball (including on the exceptional locus of the tempered spectrum)

Suppose we have a sequence of finite $(q + 1)$-regular graphs which may be realized as quotients of the Bruhat-Tits tree for $\mathrm{PGL}(2, F)$. We wish to present a reworking of some steps of the proof of Theorem I.4 (Quantum Ergodicity on Large Regular Graphs) but in such a way as to motivate some of the new ideas which are involved in the $\mathrm{PGL}(3, F)$ case.

Let

$$\mathcal{S} := (i\mathbb{R})/\Big(\frac{2\pi i}{\ln(q)}\mathbb{Z}\Big).$$

We can parametrize Ω^+_{temp} as:

$$\Omega^+_{\text{temp}} := \{(q^s, q^{-s}) : s \in \mathcal{S}\}/\mathfrak{S}_2,$$

where $\mathfrak{S}_2$ acts by permuting the coordinates. Topologically this corresponds to an interval (an alternate and perhaps more standard way to parametrize the tempered spectrum is by the interval $[-2\sqrt{q}, 2\sqrt{q}] = \{q^{1/2+s} + q^{1/2-s} : s \in \mathcal{S}\}$).

Let E_m be the ball of radius m centered at $1K$, and let U_m be the operator (on the tree) corresponding to summing up over a ball of radius m. Suppose ψ_s is an eigenfunction of $H(G, K)$ with associated spectral parameter (q^s, q^{-s}) as above. Let $h_m(s)$ be defined by the equation $U_m\psi_s = h_m(s)\psi_s$. Then, by following a completely identical procedure to Chapter II.2.2, we are naturally led to analyzing

$$\frac{1}{M}\sum_{m=1}^{M}\frac{|h_m(s)|^2}{\mathrm{card}(E_m)}.$$

Proposition III.1. *For all $(q^s, q^{-s}) \in \Omega^+_{\text{temp}}$ we have*

$$\frac{1}{M}\sum_{m=1}^{M}\frac{|h_m(s)|^2}{\mathrm{card}(E_m)} \gtrsim_M 1,$$

with the implied constant not depending on s or M.

Sketch of proof. We may compute $h_m(s)$ using the spherical function. We obtain

$$h_m(s) = \sum_{d=0}^{m}\frac{q^{\frac{d}{2}}}{\nu_{(d,0)}(q^{-1})}(q^{s\cdot d}c(s) + q^{-s\cdot d}c(-s)),$$

where

$$c(s) = \frac{q^s - q^{-1}q^{-s}}{q^s - q^{-s}},$$

and $\nu_{(d,0)}(q^{-1}) = 1$ if $d \neq 0$ and $\nu_{(d,0)}(q^{-1}) = 1 + q^{-1}$ is $d = 0$.

Using geometric series, we get that

$$\begin{aligned} h_m(s) &= c(s)\frac{q^{(m+1)(s+\frac{1}{2})} - 1}{q^{s+\frac{1}{2}} - 1} + c(-s)\frac{q^{(m+1)(-s+\frac{1}{2})} - 1}{q^{-s+\frac{1}{2}} - 1} \\ &+ \Big(\frac{1}{1+q^{-1}} - 1\Big)(c(s) + c(-s)) \\ &\approx q^{\frac{m}{2}}\Big(\frac{c(s)}{q^{s+\frac{1}{2}} - 1}\, q^{ms} + \frac{c(-s)}{q^{s-\frac{1}{2}} - 1}\, q^{-ms}\Big), \\ \mathrm{card}(E_m) &\approx q^m. \end{aligned}$$

Hence,

$$\frac{h_m(s)}{\sqrt{\mathrm{card}(E_m)}} \approx \kappa(s)q^{ms} - \kappa(-s)q^{-ms},$$

where

$$\kappa(s) := \frac{c(s)}{q^{s+\frac{1}{2}} - 1}.$$

We also define

$$\kappa'(s) := \frac{q^s - q^{-1}q^{-s}}{q^{s+\frac{1}{2}} - 1}$$

so that $\kappa(s) = \kappa'(s)/(q^s - q^{-s})$.

There are two "exceptional points" where the above formulas no longer work: $(1,1)$ and $(-1,-1)$, namely $s = 0$ and $s = i\pi/\ln(q)$. We shall separate the analysis into those points which are of distance of order at least $\frac{1}{M}$ away from these points, and those which are a distance of order at most $\frac{1}{M}$ away.

Let $A > 0$ be some constant. Suppose s is such that $|q^s - 1| \geq \frac{A}{M}$. We then have:

$$\begin{aligned} \frac{1}{M}\sum_{m=1}^{M}\frac{|h_m(s)|^2}{\mathrm{card}(E_m)} &\approx \frac{1}{M}\sum_{m=1}^{M}|\kappa(s)q^{ms} + \kappa(-s)q^{-ms}|^2 \\ &\approx |\kappa(s)|^2 + |\kappa(-s)|^2 + \kappa(s)\overline{\kappa(-s)}\frac{q^{Ms} - 1}{M(q^s - 1)} + \overline{\kappa(s)}\kappa(-s)\frac{q^{-Ms} - 1}{M(q^{-s} - 1)} \end{aligned}$$

$$\geq |\kappa(s)|^2 + |\kappa(-s)|^2 - |\kappa(s)||\kappa(-s)|\left(\Big|\frac{q^{Ms}-1}{M(q^s-1)}\Big| + \Big|\frac{q^{-Ms}-1}{M(q^{-s}-1)}\Big|\right)$$
$$\geq \frac{1}{|q^s - q^{-s}|^2}(|\kappa'(s)|^2 + |\kappa'(-s)|^2 - \frac{2}{A}(\sup_r |\kappa'(r)|)^2$$

Because $|\kappa'(r)|$ is bounded on Ω^+_{temp}, we get that we can choose A large enough that

$$\frac{1}{M}\sum_{m=1}^{M} \frac{|h_m(s)|^2}{\text{card}(E_m)} \geq \frac{1}{2}(|\kappa(s)|^2 + |\kappa(-s)|^2). \tag{III.4.1}$$

The right hand side of (III.4.1) has a universal positive lower bound.

Notice that, for M large, $|q^s - 1| \gtrsim \frac{1}{M}$ implies $|s| \gtrsim \frac{1}{M}$. Hence the above shows that we can find a universal lower bound on the quantity of interest (which depends on M) for points which whose distance from $s = 0$ is of order $\frac{1}{M}$.

We now consider points whose distance from $s = 0$ is of size at most $\frac{1}{M}$. For these points we employ a different strategy. First note that by Cauchy-Schwarz we have:

$$\Big|\frac{1}{M}\sum_{m=1}^{M} \frac{h_m(s)}{\sqrt{\text{card}(E_m)}}\Big|^2 \leq \frac{1}{M}\sum_{m=1}^{M} \frac{|h_m(s)|^2}{\text{card}(E_m)}.$$

On the other hand we have

$$\frac{1}{M}\sum_{m=1}^{M} \frac{h_m(s)}{\sqrt{\text{card}(E_m)}} \approx \kappa(s)\frac{q^{sM}-1}{M(q^s-1)} + \kappa(-s)\frac{q^{-sM}-1}{M(q^{-s}-1)}. \tag{III.4.2}$$

Let $s = \frac{t}{M}$. By a L'Hopital argument, one can show that

$$\lim_{M\to\infty} \frac{1}{M}\left(\kappa\Big(\frac{t}{M}\Big)\frac{q^t - 1}{M(q^{\frac{t}{M}}-1)} + \kappa\Big(-\frac{t}{M}\Big)\frac{q^{-t}-1}{M(q^{-\frac{t}{M}}-1)}\right) = \frac{1-q^{-1}}{q^{\frac{1}{2}}-1}\frac{(q^t-1)+(q^{-t}-1)}{2\ln(q)^2 t^2},$$

and the convergence is uniform on compact sets such as $|t| \leq B$. The term in the limit on the left hand side is exactly (III.4.2) divided by M. Also notice that the term on the right hand side has a positive limit as $t \to 0$ and is in fact positive for all t values. We therefore can find a positive lower bound of the form $C_2 \cdot M$ for all s such that $|s| \leq \frac{B}{M}$ for any B. Combining this with (III.4.1) (taking $B \geq A$), we obtain the result (a similar argument can be carried out near $s = i\pi/\ln(q)$). □

Remark III.2. Notice that in contrast to Proposition II.1 (the Spectral Bound for the

$\mathrm{PGL}(3, F)$ case), here we have obtained a spectral bound for the entire tempered spectrum including the "exceptional locus" $\{(1,1),(-1,-1)\}$. We strongly believe that the methods presented here for the spectral bound can be adapted to strengthen Proposition II.1 to also hold on the exceptional locus, and we have made a lot of progress in this direction; we are still in the process of writing up these results. See also Remark III.2.

III.5: The geometric bound on the regular tree

In order to motivate the method of proof for the geometric bound, we shall present a proof of the geometric bound on the tree using the same method that we employ in rank 2. As in Chapter III.4, we use the metric ball to define our "wave propagator." One may perform a similar analysis to that carried out in Chapters II.2.10, II.2.11, and II.2.12 to reduce to computing the size of sets of the form

$$E_m^r := E_m \cap \varpi^{(r,0)} E_m.$$

Notice that in $\mathfrak{a}$-coordinates $d_{A^+}(1K, \varpi^{(r,0)}K) = (r/2, -r/2)$.

The polytope P_m defining E_m is the interval $[0, m]$ in cone coordinates, or equivalently the convex hull of $(0,0)$ and $(\frac{m}{2}, -\frac{m}{2})$ in $\mathfrak{a}$-coordinates. We have that

$$\delta = \left(\frac{1}{2}, -\frac{1}{2}\right)$$

in $\mathfrak{a}$-coordinates, and the vertex of P_m maximizing $(\delta, \cdot)$ is

$$p^\dagger := \left(\frac{m}{2}, -\frac{m}{2}\right)$$

in $\mathfrak{a}$-coordinates. The analogue of Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$) is the following:

Proposition III.3. *We have*

$$\mathrm{card}(E_m^r) \lesssim q^{(2\delta, m \cdot p^\dagger - \frac{1}{2}(\frac{r}{2}, -\frac{r}{2}))} = q^{m - \frac{r}{2}},$$

where the implied constant does not depend on m or r.

We now describe the steps of the procedure from Chapter II.2.13 as applied to bounding $\mathrm{card}(E_m^r)$.

III.5.1: Classify triples of points

Given two points $x, y \in \mathcal{T}$, there is a unique geodesic segment connecting them which we denote by $\mathrm{geod}(x, y)$; this is the rank one analogue of the parallelograms used later in Chapter IX.4. We say that $(x, y; z)$ forms a *primitive triple* if $\mathrm{geod}(x, z) \cap \mathrm{geod}(y, z) = \{z\}$. We claim that that the only such triples are $(x, y; z)$ such that $z \in \mathrm{geod}(x, y)$. To see this, note that $\mathrm{geod}(x, z) \cap \mathrm{geod}(y, z)$ is a geodesic segment. One of the endpoints of this geodesic segment is z, and the other, call it w, must be closer to x and y than z is. The initial edge of the geodesic from w to x and the initial edge of the geodesic from w to y must be distinct (otherwise there would be additional points in $\mathrm{geod}(x, z) \cap \mathrm{geod}(y, z)$). Therefore in concatenating these geodesics, we must obtain the unique geodesic from x to y which is to say that z lies along this geodesic.

III.5.2: Coordinatization of triples of points

Now given any triple of points $(x, y; z)$, we may associate a *confluence-branch point* w, namely the unique vertex on $\mathrm{geod}(x, y)$ contained in $\mathrm{geod}(x, z) \cap \mathrm{geod}(y, z)$; this terminology is such because w is simultaneously playing the role of the confluence point and the branch line as in the discussion in Chapter X.2. Let $a = d(x, w)$ and let $b = d(w, z)$. Let $r = d(x, y)$. We thus have that $d(x, z) = a + b$ and $d(y, z) = (r - a) + b$.

Now consider the intersection of two balls of radius m, one centered at x and the other at y (with $d(x, y) = r$). For each point z in the intersection, we can associate to it a confluence-branch point w and a tuple (a, b). The collection of coordinates (a, b) which show up this way for some m and some r is exactly the integer lattice points satisfying:

$$\begin{aligned} a + b &\leq m \\ (r - a) + b &\leq m \\ a &\leq r \\ m, r, a, b &\geq 0. \end{aligned} \tag{III.5.1}$$

When m and r are fixed, the shape of the corresponding intersection of balls looks like Figure 1. The collection of allowable (a, b)'s corresponds to lattice points in a polytope whose "type" is as in Figure 2 (actually there are two possible types depending on whether or not $r \geq m$).

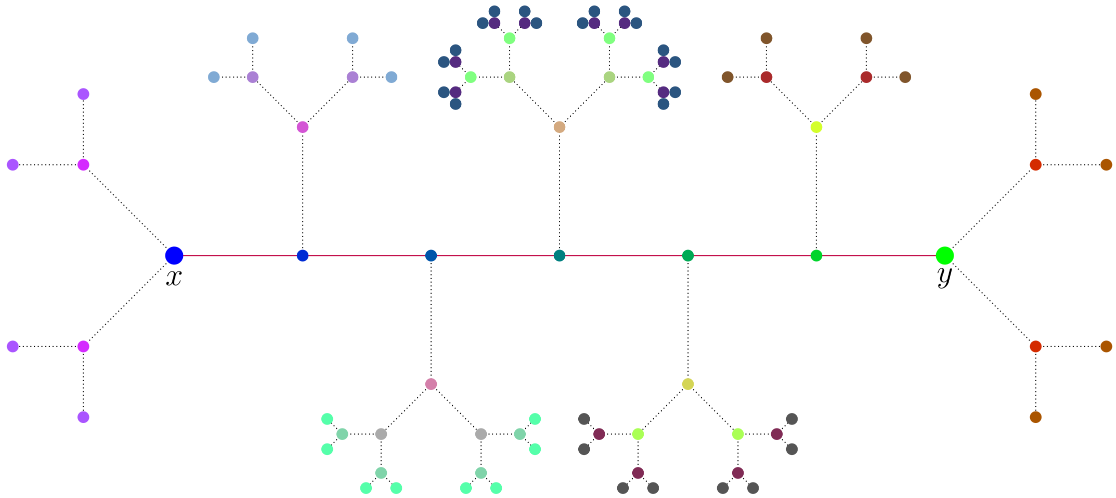


Figure 1: This figure shows the intersection in a 3-regular tree (hence $q = 2$) of two balls of radius 8 centered at two points x, y whose distance apart is 6 (hence $r = 6$ and $m = 8$). The red edges represent the geodesic segment connecting x and y. The possible confluence-branch points are the points along these red edges. We may associate coordinates (a, b) to each point. All points which receive the same coordinates are colored with the same color. Notice that the greatest contribution comes from those vertices whose confluence-branch point is the midpoint of the geodesic segment joining x and y.

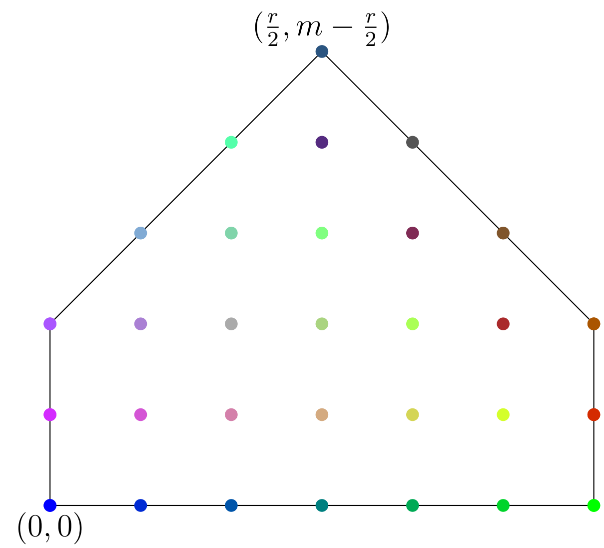


Figure 2: The collection of possible coordinates (a, b) of vertices in Figure 1 corresponds to lattice points in a convex polytope as below. We have colored each lattice point with the same color as all vertices in Figure 1 which receive those coordinates. As we vary r and m (but keep satisfying $r \leq m$), then all of the resulting polytopes have the same "type" (see Appendix C.2.1).

III.5.3: Counting the number of points with a given set of coordinates

Suppose m and r are fixed. The points z whose associated coordinates are (a, b) are exactly those vertices which are obtained by taking b steps away from the geodesic joining x and y starting from the point along this geodesic whose distance from x is a. At the first step we have $(q-1)$ choices (unless $a = 0$ or $a = r$ in which case we have q choices), and every subsequent step we have q choices. Therefore

$$\#\{z \text{ with coordinates } (a, b)\} \leq q^b. \tag{III.5.2}$$

III.5.4: Identifying the dominating term

We now wish to understand what the greatest value that b can take is on the lattice points in the polytopes in (III.5.1). If we branch at some vertex w with $d(x, w) = a$, then we may take at most $\min\{m - a, m - r - a\}$ steps. The greatest this can ever be is exactly when $a = \frac{r}{2}$, namely w is exactly the midpoint of the geodesic from x to y, in which case we may take exactly $m - \frac{r}{2}$ steps.

III.5.5: Bounding the size of the intersection of balls

We now wish to bound $\mathrm{card}(E_m^r)$. By (III.5.2), we know that we can bound this quantity by the sum of q^b over the lattice points (a, b) in the associated polytope. If we restrict to the case of $r \leq m$ (as in Figure 2), then all of the associated polytopes have the same type. We always have a vertex at $(r/2, m - r/2)$. The cone generated by this vertex in the polytope contains the rest of the polytope; as we vary over polytopes in this family, we always get the same cone but shifted; see Figure 3. It is straightforward to see that the sum of q^b over the lattice points in this cone is bounded by (see C.3.2):

$$\frac{(1+q^{-1})}{(1-q^{-1})(1-q^{-1})}\, q^{m-\frac{r}{2}}.$$

This hence provides an upper bound on $\mathrm{card}(E_m^r)$.

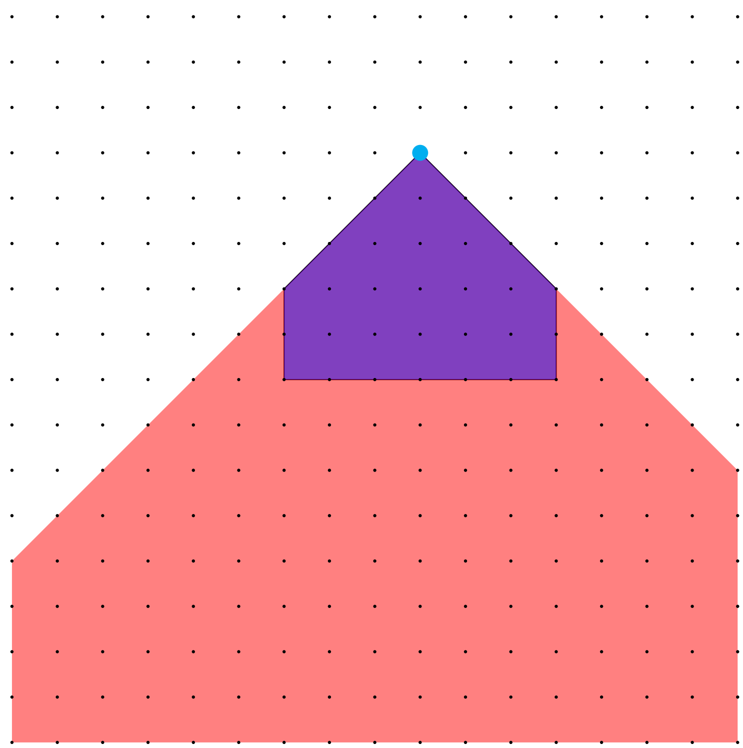


Figure 3: In this figure, the vertex of the cone has coordinates $(r/2, m - r/2)$, the purple polytope is the same as the polytope from Figure 2, and the red cone is the cone generated by the rays at the vertex in that polytope.

CHAPTER IV
Spectral Bound for Polytopal Ball Averaging Operators

IV.1: Parametrization of the tempered spectrum

Let $\mathcal{S}$ be defined as:

$$\mathcal{S} := \{(s_1, s_2, s_3) : s_j \in (i\mathbb{R})/\Big(\frac{2\pi i}{\ln(q)}\mathbb{Z}\Big) \text{ and } s_1 + s_2 + s_3 = 0 \ \Big(\text{mod } \frac{2\pi i}{\ln(q)}\mathbb{Z}\Big)\}.$$

Notice that we may naturally view $\mathcal{S}$ as a two-torus. Given $s = (s_1, s_2, s_3) \in \mathcal{S}$, we define

$$q^s := (q^{s_1}, q^{s_2}, q^{s_3}).$$

Then q^s parametrizes the tempered spectrum Ω^+_{temp} uniquely up to permuting coordinates; this essentially corresponds to the Satake parameters (see Appendix A.4.3). Hence Ω^+_{temp} corresponds to the two-torus corresponding to $\mathcal{S}$ quotiented by the $\mathfrak{S}_3$-action of permuting the coordinates.

In the sequel we shall let $\sigma.s$ with $\sigma \in \mathfrak{S}_3$ denote the associated permutation of the entries (not the indices), e.g. $(1\ 2\ 3).(s_1, s_2, s_3) = (s_3, s_1, s_2)$. Notice that entrywise multiplication by -1 is well-defined on $\mathcal{S}$. We shall, e.g., use the notation $-\sigma.s$ mean the result of applying σ to the entries of s, and then multiplying each entry by -1.

IV.2: The exceptional locus

The exceptional locus $\Xi \subset \Omega^+_{\text{temp}}$ is composed of two pieces:

$$\Xi_1 := \{q^{is} : \sigma.s = s \text{ for some } \sigma \neq 1 \text{ in } \mathfrak{S}_3\}/\mathfrak{S}_3,$$

$$\Xi_2 := \{q^{is} : q^{(p^\dagger, -\sigma_1.s+\sigma_2.s)} = 1 \text{ for some } \sigma_1, \sigma_2 \in \mathfrak{S}_3\}/\mathfrak{S}_3.$$

See Figure 4 for a visualization of Ω^+_{temp}.

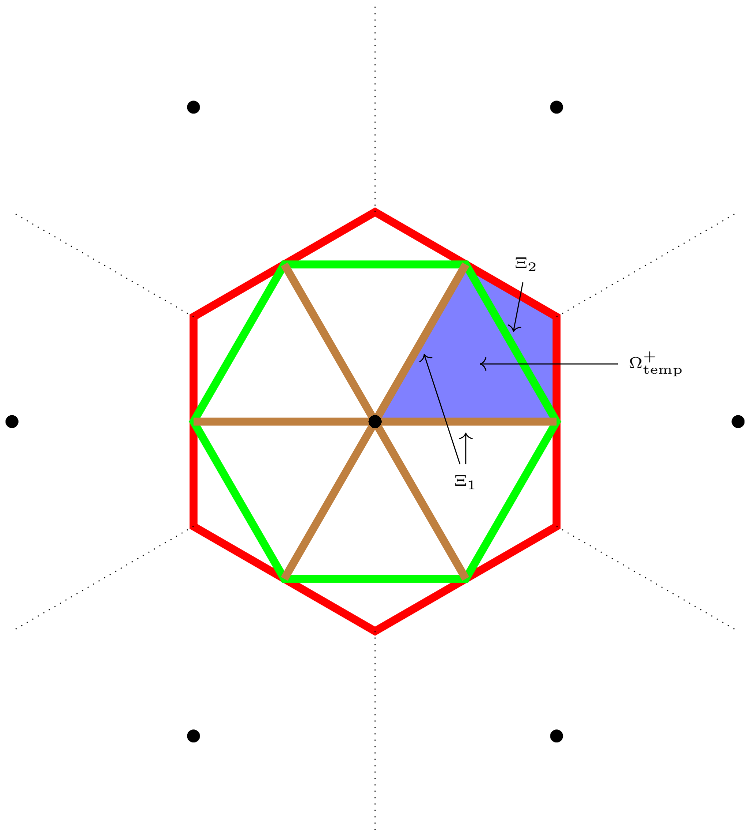


Figure 4: The red hexagon represents a fundamental domain for the hexagonal lattice in $\mathbb{R}^2$ (represented by the black dots). We may naturally identify this fundamental domain with $\mathcal{S}$ (or equivalently with $\{q^s : s \in \mathcal{S}\}$). The blue region is a fundamental domain for the $\mathfrak{S}_3$-action of permuting the coordinates and hence provides a realization of Ω^+_{temp}. The brown lines in the blue region correspond to the locus Ξ_1, and the green line in the blue region corresponds to Ξ_2.

IV.3: Explicit formula for $|h_m(s)|^2$

Suppose $Y = \Gamma\backslash G/K$. Let ψ_s be an eigenfunction of the $H(G,K)$-action on $L^2(Y)$ with spectral parameter $q^s \in \Omega^+_{\text{temp}}$. We may think of ψ_s as a (Γ, K)-invariant function on G. Let ω_s be the spherical function with spectral paramter q^s. For any function $\mathfrak{h} \in H(G,K)$, we know that $\psi_s * \mathfrak{h} = \hat{\mathfrak{h}}(s)\psi_s$ with $\hat{\mathfrak{h}}(s) := (\omega_s * \mathfrak{h})(1)$. On the other hand, we know that

$$g_\lambda(s) := (\omega_s * \mathbb{1}_{K\varpi^\lambda K})(1) = q^{(\delta,\lambda)} P_\lambda(q^{-s_1}, q^{-s_2}, q^{-s_3}; q^{-1})$$

where $\varpi^\lambda \in A^+$ and P_λ is the Hall-Littlewood polynomial associated to $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ (see Appendix A.2.3). Plugging in the explicit formula for P_λ, we obtain

$$\begin{aligned} g_\lambda(s) &= \frac{q^{(\lambda,\delta)}}{\nu_\lambda(q^{-1})} \sum_{\sigma \in \mathfrak{S}_3} \sigma.\Big(q^{(\lambda,-s)} \prod_{j<k} \frac{q^{-s_j} - q^{-1}q^{-s_k}}{q^{-s_j} - q^{-s_k}}\Big) \\ &= \frac{1}{\nu_\lambda(q^{-1})} \sum_{\sigma \in \mathfrak{S}_3} \Big(c\big(-\sigma.s\big)\ q^{(\lambda,\delta-\sigma.s)}\Big), \end{aligned}$$

where the c-function $c(s)$ is defined as

$$c(s) = \prod_{j<k} \frac{q^{s_j} - q^{-1}q^{s_k}}{q^{s_j} - q^{s_k}}, \tag{IV.3.1}$$

and

$$\begin{array}{ll} \nu_\lambda(q^{-1}) = 1 & \text{if } \lambda_1 > \lambda_2 > \lambda_3, \\ \nu_\lambda(q^{-1}) = 1 + q^{-1} & \text{if } \lambda_1 = \lambda_2 > \lambda_3 \text{ or } \lambda_1 > \lambda_2 = \lambda_3, \\ \nu_\lambda(q^{-1}) = (1+q^{-1})(1+q^{-1}+q^{-2}) & \text{if } \lambda_1 = \lambda_2 = \lambda_3. \end{array} \tag{IV.3.2}$$

We are in particular interested in computing $h_m(s)$, which is defined by $\psi_s * \mathbb{1}_{E_m} = h_m(s)\psi_s$. Hence $h_m(s) = (\omega_s * \mathbb{1}_{E_m})(1)$. On the other hand

$$\mathbb{1}_{E_m} = \sum_{\lambda \in P_m^\Lambda} \mathbb{1}_{K\varpi^\lambda K}.$$

We therefore arrive at the formula

$$h_m(s) = \sum_{\sigma \in \mathfrak{S}_3} \sum_{\lambda \in P_m^\Lambda} \frac{1}{\nu_\lambda(q^{-1})} \Big(c\big(-\sigma.s\big)\ q^{(\lambda,\delta-\sigma.s)}\Big).$$

If $s \in \mathcal{S}$ and $\sigma \in \mathfrak{S}_3$ are fixed and we ignore the factor $\frac{1}{\nu_\lambda(q^{-1})}$, then the inner sum reduces to a sum of an exponential function over P_m^Λ and hence may be handled by Brion's formula (see Appendix C.3). However, $\nu_\lambda(q^{-1})$ is not constant in λ: it is 1 on the interior of the Weyl chamber $\mathfrak{a}^+$, it is $1 + q^{-1}$ on the intersection of Λ with the interior of each extremal ray of

$\mathfrak{a}^+$, and it is $(1+q^{-1})(1+q^{-1}+q^{-2})$ at the base vertex of $\mathfrak{a}^+$ namely $(0,0,0) \in \mathfrak{a}$. Therefore, using inclusion-exclusion and the labels as in Figure 5, we can write

$$\begin{aligned} h_m(s) = \sum_{\sigma \in \mathfrak{S}_3} c(-\sigma.s) \Bigg(& \sum_{\lambda \in P_m^\Lambda} q^{(\lambda, \delta - \sigma.s)} \\ & + \Big(\frac{1}{1+q^{-1}} - 1 \Big) \sum_{\lambda \in (e_{1,2})_m^\Lambda} q^{(\lambda, \delta - \sigma.s)} \\ & + \Big(\frac{1}{1+q^{-1}} - 1 \Big) \sum_{\lambda \in (e_{1,3})_m^\Lambda} q^{(\lambda, \delta - \sigma.s)} \\ & + \Big(\frac{1}{(1+q^{-1})(1+q^{-1}+q^{-2})} - 1 - 2\Big(\frac{1}{1+q^{-1}} - 1 \Big) \Big) \Bigg). \end{aligned}$$

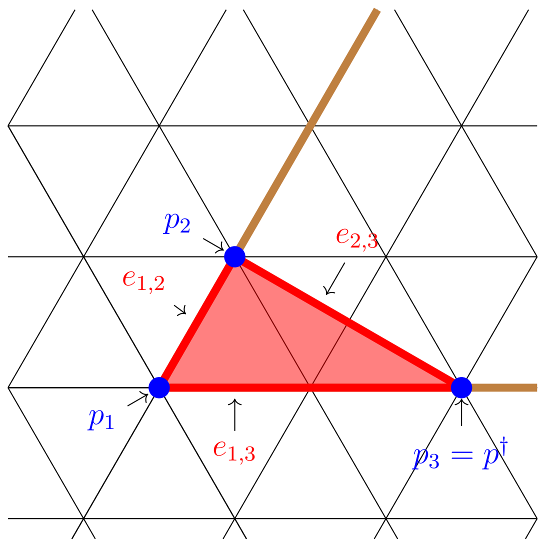


Figure 5: This shows the vertices and edges of P.

We can use Brion's formula to compute each term:

$$\begin{aligned} h_m(s) = \sum_{\sigma \in \mathfrak{S}_3} c(-\sigma.s) \Bigg(& \sum_{j \in \{1,2,3\}} \sigma(\mathrm{Cone}_P(p_j); \delta - \sigma.s) q^{m(p_j, \delta - \sigma.s)} \\ & + \Big(\frac{1}{1+q^{-1}} - 1 \Big) \sum_{j \in \{1,2\}} \sigma(\mathrm{Cone}_{e_{1,2}}(p_j); \delta - \sigma.s) q^{m(p_j, \delta - \sigma.s)} \\ & + \Big(\frac{1}{1+q^{-1}} - 1 \Big) \sum_{j \in \{1,3\}} \sigma(\mathrm{Cone}_{e_{1,3}}(p_j); \delta - \sigma.s) q^{m(p_j, \delta - \sigma.s)} \end{aligned}$$

$$+\Big(\frac{1}{(1+q^{-1})(1+q^{-1}+q^{-2})}-1-2\big(\frac{1}{1+q^{-1}}-1\big)\Big)\Bigg).$$

IV.4: The cardinality of E_m

Now let's compute card(E_m). We have that (see (II.2.11)):

$$N_\lambda := \mathrm{vol}(K\varpi^\lambda K) = q^{2(\lambda,\delta)}\frac{\nu_3(q^{-1})}{\nu_\lambda(q^{-1})},$$

where $\nu_3(q^{-1})$ is given by (A.2.2); notice that it is non-zero and does not depend on λ. Using the same style of analysis as before, we get that

$$\begin{aligned}
\mathrm{card}(E_m) &= \sum_{\lambda\in P_m^\Lambda} N_\lambda \\
&= \nu_3(q^{-1})\Bigg(\sum_{j\in\{1,2,3\}}\sigma(\mathrm{Cone}_P(p_j);2\delta)q^{2m(p_j,\delta)} \\
&+ \Big(\frac{1}{1+q^{-1}}-1\Big)\sum_{j\in\{1,2\}}\sigma(\mathrm{Cone}_{e_{1,2}}(p_j);2\delta)q^{2m(p_j,\delta)} \\
&+ \Big(\frac{1}{1+q^{-1}}-1\Big)\sum_{j\in\{1,3\}}\sigma(\mathrm{Cone}_{e_{1,3}}(p_j);2\delta)q^{2m(p_j,\delta)} \\
&+ \Big(\frac{1}{(1+q^{-1})(1+q^{-1}+q^{-2})}-1-2\big(\frac{1}{1+q^{-1}}-1\big)\Big)\Bigg)
\end{aligned} \tag{IV.4.1}$$

IV.5: Proof of Propositions II.13 and II.14

Proof of Proposition II.13 (Upper Bound on N_λ). From (A.3.2) and (IV.3.2), it is clear that

$$N_\lambda \le \nu_3(q^{-1})q^{2(\delta,\lambda)}.$$

□

Proof of Proposition II.14 (Lower Bound on card(E_m)*).* Recall that card(E_m) is computed by summing up N_λ over P_m^Λ. Furthermore N_λ is positive for every λ. We always have

$m \cdot p^\dagger \in P_m^\Lambda$. If $m \geq 1$, then

$$N_{m\cdot p^\dagger} = \frac{\nu_3(q^{-1})}{1+q^{-1}}\, q^{2(\delta, m\cdot p^\dagger)} \leq \mathrm{card}(E_m).$$

If $m = 0$ we get

$$N_{m\cdot p^\dagger} = 1 = q^{2(\delta, m\cdot p^\dagger)} = \mathrm{card}(E_m).$$

$\square$

IV.6: Proof of Proposition II.1

Proof of Proposition II.1 (Spectral Bound). We wish to find a lower bound for

$$\frac{1}{M}\sum_{m=1}^{M} \frac{|h_m(s)|^2}{\mathrm{card}(E_m)}.$$

We first seek to bound $\frac{|h_m(s)|^2}{\mathrm{card}(E_m)}$. Analyzing Figure 5, we see that $p^\dagger$ is the unique vertex of P maximizing $|q^{(\lambda, \delta - \sigma.s)}|$ (recall that $q^{(\lambda,s)}$ is on the unit circle), and that $p^\dagger = p_3$ is part of P and $e_{1,3}$. Furthermore, observe that the c-function (IV.3.1) is holomorphic away from Ξ_1. Hence because we are assuming that Θ is a compact subset not meeting Ξ_1, we conclude that the norm of the c-function has a universal upper bound on Θ. Also note that all of the terms of the form $\sigma(C; \delta - w.s)$, where C is some cone, are independent of m and also universally upper bounded in norm for all $q^s \in \Theta$ (in fact for all $q^s \in \Omega^+_{\mathrm{temp}}$). Therefore, we obtain

$$h_m(s) = \Big(\sum_{\sigma\in\mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger, -\sigma.s)}\Big) q^{m(\delta, p^\dagger)} + R_1(m,s) q^{m((p^\dagger,\delta)-\eta_1)}, \tag{IV.6.1}$$

where

$$\kappa(s) = c(-s)\Big[\sigma(\mathrm{Cone}_P(p^\dagger); \delta - s) - \frac{\sigma(\mathrm{Cone}_{e_{1,3}}(p^\dagger); \delta - s)}{q+1}\Big], \tag{IV.6.2}$$

and $R_1(m,s)$ is some function which is universally bounded for all m and all $q^s \in \Theta$, and η_1 is some positive constant.

Since we also have that $\kappa(\sigma.s)$ is bounded for all $\sigma \in \mathfrak{S}_3$ and $q^s \in \Theta$, we get that

$$|h_m(s)|^2 = \Big|\sum_{\sigma\in\mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger, -\sigma.s)}\Big|^2 q^{2m(\delta, p^\dagger)} + R_2(m,s) q^{2m((p^\dagger,\delta)-\eta_2)},$$

where $R_2(m,s)$ is some function which is universally bounded for all m and all $q^s \in \Theta$, and η_2 is some positive constant.

We now analyze $\mathrm{card}(E_m)$. Again, because $p^\dagger$ is the unique vertex in P maximizing $q^{2(\lambda,\delta)}$, we see that

$$\mathrm{card}(E_m) = (D + R_3(m)) q^{2m(\delta, p^\dagger)},$$

where

$$D = \nu_3(q^{-1})\left(\sigma(\mathrm{Cone}_P(p^\dagger); 2\delta) + (\frac{1}{1+q^{-1}} - 1)\sigma(\mathrm{Cone}_{e_{1,3}}(p^\dagger); 2\delta)\right), \tag{IV.6.3}$$

and $R_3(m) \to 0$ as $m \to \infty$. A straightforward calculation shows that D is positive.

Therefore, using the identity

$$\frac{C_1}{C_2 + R} = \frac{C_1}{C_2} - \frac{C_1 R}{C_2(C_2+R)},$$

we obtain

$$\begin{aligned}
\frac{|h_m(s)|^2}{\mathrm{card}(E_m)} &= \frac{|\sum_{\sigma\in\mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger,-\sigma.s)}|^2 q^{2m(\delta,p^\dagger)} + R_2(m,s) q^{2m((p^\dagger,\delta)-\eta_2)}}{(D+R_3(m)) q^{2m(\delta,p^\dagger)}} \\
&= \frac{|\sum_{\sigma\in\mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger,-\sigma.s)}|^2}{D} \\
&\quad - \frac{|\sum_{\sigma\in\mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger,-\sigma.s)}|^2 R_3(m)}{D(D+R_3(m))} + \frac{R_2(m,s)}{(D+R_3(m)) q^{2m\cdot\eta_2}}.
\end{aligned} \tag{IV.6.4}$$

Because $R_3(m) \to 0$ as $m \to \infty$, it is clear that given $\varepsilon > 0$, there exists an M_0 such that for all $m \geq M_0$ and for all $q^s \in \Theta$, the first term in (IV.6.4) is bounded in absolute value by ε. Furthermore, since $\eta_2 > 0$, it is also clear that there exists an M_1 such that for all $m \geq M_1$ and for all $s \in \Theta$, the second term in (IV.6.4) is bounded in absolute value by ε.

Let $R_4(m,s)$ equal (IV.6.4). We define $R_5(m)$:

$$R_5(m) := \sup_{\{s : q^s \in \Theta\}} |R_4(m,s)|.$$

It is clear that $R_5(m) \to 0$ as $m \to \infty$. We therefore have that for all $q^s \in \Theta$,

$$\Big| \frac{1}{M} \sum_{m=1}^{M} R_4(m,s) \Big| \leq \frac{1}{M} \sum_{m=1}^{M} R_5(m) \to 0$$

as $m \to \infty$.

In the sequel we shall prove the following lemma:

Lemma IV.1. *There exist an $M_2 \in \mathbb{N}$ and $C > 0$ such that for all $M \geq M_2$ and for all $q^s \in \Theta$,*

$$\frac{1}{M} \sum_{m=1}^{M} \Big| \sum_{\sigma \in \mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger, -\sigma.s)} \Big|^2 \geq C.$$

Given this lemma we can finish the proof of Proposition II.1. Let M_3 be such that for all $M \geq M_3$, we have

$$\frac{1}{M} \sum_{m=1}^{M} R_5(m) \leq \frac{C}{2D},$$

with C given by Lemma IV.1 and D given by (IV.6.3). We then have that if $M \geq M_2$ and $M \geq M_3$, then

$$\begin{aligned}
\frac{1}{M} \sum_{m=1}^{M} \frac{|h_m(s)|^2}{\operatorname{card}(E_m)} &= \frac{1}{M} \sum_{m=1}^{M} \frac{|\sum_{\sigma \in \mathfrak{S}_3} \kappa(\sigma.s) q^{m(p^\dagger, -\sigma.s)}|^2}{D} + \frac{1}{M} \sum_{m=1}^{M} R_4(m,s) \\
&\geq \frac{C}{D} - \frac{C}{2D} \\
&= \frac{C}{2D}.
\end{aligned}$$

□

Proof of Lemma IV.1. Recall that $p^\dagger = (4/3, -2/3, -2/3)$, which is fixed by the permutation (23) (and the trivial permutation). Let $L = \{1, (23)\}$. Notice that if $\sigma_1 \sigma_2^{-1} \in L$ (i.e. σ_1 and σ_2 send the same element to 1), then $q^{(p^\dagger, -\sigma_1.s + \sigma_2.s)} = 1$. Notice also that $\overline{\kappa(s)} = \kappa(-s)$.

Therefore we shall expand and group terms as follows

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M}\Big|\sum_{\sigma\in\mathfrak{S}_3}\kappa(\sigma.s)q^{m(p^\dagger,-\sigma.s)}\Big|^2 &= \frac{1}{M}\sum_{m=1}^{M}\sum_{\sigma_1,\sigma_2\in\mathfrak{S}_3}\kappa(\sigma_1.s)\kappa(-\sigma_2.s)q^{m(p^\dagger,-\sigma_1.s+\sigma_2.s)}\\ &= \sum_{\sigma\in\mathfrak{S}_3}|\kappa(\sigma.s)|^2+\sum_{\sigma\in\mathfrak{S}_3}\kappa(\sigma.s)\kappa\big(-(23)\sigma.s\big)+\text{ r.t.}\\ &= \sum_{\text{right cosets }L.\sigma}\big|\kappa(\sigma.s)+\kappa\big((23)\sigma.s\big)\big|^2+\text{ r.t.,}\end{aligned}\tag{IV.6.5}$$

where r.t. stands for remainder terms. We wish to show that the sum in (IV.6.5) is non-zero. Evidently this amounts to showing that at least one $\kappa(\sigma.s)+\kappa\big((23)\sigma.s\big)\neq 0$ for every s.

Recall that $s=(s_1,s_2,s_3)$ with $s_1+s_2+s_3=0$ and each s_i purely imaginary and only defined up to $2\pi i\mathbb{Z}/\ln(q)$. For the polytope P, the coprime generators for the rays generating the cone at $p^\dagger$ are $\{(-\frac{2}{3},\frac{1}{3},\frac{1}{3}),(-1,1,0)\}$. For the polytope $e_{1,3}$, the coprime generator for the ray generating the cone at $p^\dagger$ is $(-\frac{2}{3},\frac{1}{3},\frac{1}{3})$. Therefore, using (C.3.1), we obtain

$$\begin{aligned}\sigma(\mathrm{Cone}_P(p^\dagger);\delta-s) &= \frac{1}{1-q^{(\delta-s,(-\frac{2}{3},\frac{1}{3},\frac{1}{3}))}}\cdot\frac{1}{1-q^{(\delta-s,(-1,1,0))}}\\ &= \frac{1}{1-q^{-1+s_1}}\cdot\frac{1}{1-q^{-1+s_1-s_2}},\end{aligned}\tag{IV.6.6}$$

$$\sigma(\mathrm{Cone}_{e_{1,3}}(p^\dagger);\delta-s) = \frac{1}{1-q^{(\delta-s,(-\frac{2}{3},\frac{1}{3},\frac{1}{3}))}} = \frac{1}{1-q^{-1+s_1}}.\tag{IV.6.7}$$

Using (IV.3.1), (IV.6.2), (IV.6.6), and (IV.6.7), we can perform a change a variables $x_1=q^{-s_1}$, $x_2=q^{-s_2}$, $x_3=q^{-s_3}$, $t=q^{-1}$ and write each term as a rational function in these variables. We also have the condition $x_1x_2x_3=1$ coming from $q^{-s_2-s_2-s_3}=1$. Using a computer algebra system (such as Mathematica), we find for example that $\kappa(s)+\kappa\big((23).s\big)$ becomes

$$\frac{x_1(x_1^2-t^3x_2x_3)}{(-t+x_1)(x_1-x_2)(x_1-x_3)}.$$

If this were zero, then $x_1(x_1^2-t^3x_2x_3)=0$. Then $x_1^3=t^3x_1x_2x_3=t^3$. However, x_1 is on the unit circle, and $t^3=\frac{1}{q^3}<1$. Hence this expression is never zero on the locus of interest to us.

Now we focus on the remainder terms which are all of the form:

$$\frac{1}{M}\sum_{m=1}^{M}\Big(\kappa(\sigma_1.s)+\kappa\big((23)\sigma_1.s\big)\Big)\overline{\Big(\kappa(\sigma_2.s)+\kappa\big((23)\sigma_2.s\big)\Big)}q^{m(p^\dagger,-\sigma_1.s+\sigma_2.s)},$$

where σ_1, σ_2 satisfy $L.\sigma_1 \neq L.\sigma_2$. Since Θ does not intersect Ξ_1, and hence the c-function is bounded on Θ, we get the following bound for all $q^s \in \Theta$:

$$\left| \Big(\kappa(\sigma_1.s) + \kappa\big((23)\sigma_1.s\big)\Big) \overline{\Big(\kappa(\sigma_2.s) + \kappa\big((23)\sigma_2.s\big)\Big)} \right| \leq F$$

for all choices of σ_1 and σ_2, and some F. Furthermore, since Θ does not intersect Ξ_2, we get that

$$|q^{(p^\dagger, -\sigma_1.s+\sigma_2.s)} - 1| \geq \eta > 0,$$

for all $q^s \in \Theta$, all choices of σ_1 and σ_2, and some $\eta > 0$. Therefore, using geometric series, we get that each remainder term (of which there are finitely many) can be uniformly bounded by an expression of the form

$$F \frac{q^{M(p^\dagger, -\sigma_1.s+\sigma_2.s)} - q^{(p^\dagger, -\sigma_1.s+\sigma_2.s)}}{M(q^{(p^\dagger, -\sigma_1.s+\sigma_2.s)} - 1)} \leq \frac{2F}{M\eta}.$$

This bound goes to zero as $M \to \infty$. Hence Lemma IV.1 is proven. $\square$

CHAPTER V
Benjamini-Schramm Convergence and Plancherel Convergence

V.1: Benjamini-Schramm convergence implies Plancherel convergence

V.1.1: Plancherel sequences

Suppose M is a topological group. Let $\hat{M}$ denote the collection of irreducible unitary representations of M; it is a topological space with respect to the Fell topology (see Appendix A.5.4). Suppose Γ is a cocompact lattice in M. We then have

$$L^2(\Gamma\backslash M) = \bigoplus_{\rho\in\hat{M}} N_\Gamma(\rho)\rho$$

with each $N_\Gamma(\rho)$ finite, and only countably many of them non-zero. Let $C_c^\infty(M)$ denote space of test functions on M (see, e.g., Definition 1.2 of [Dei18]); in case $M = \mathrm{PGL}(d, F)$, then $C_c^\infty(M)$ is the set of compactly supported locally constant functions on M. Each $f \in C_c^\infty(M)$ defines an operator on each ρ (in Appendix A.10.3 this is denoted by $\hat{f}(\rho)$), and hence also on $L^2(\Gamma\backslash M)$; it turns out this operator is in fact trace class ([DE09], Theorem 9.3.2).

We shall define a measure on $\hat{M}$, called the spectral measure associated to Γ, by

$$\nu_\Gamma = \sum_{\rho\in\hat{M}} N_\Gamma(\rho)\delta_\rho,$$

where δ_ρ is the indicator function for $\rho \in \hat{M}$. Suppose (Γ_n) is a sequence of cocompact lattices and assume that M is type I (see Appendix A.10.2). We say that (Γ_n) is a *Plancherel* *sequence* if, for every $f \in C_c^\infty(M)$,

$$\frac{1}{\mathrm{vol}(\Gamma_n\backslash M)} \int_{\hat{M}} \mathrm{Tr}[\hat{f}(\rho)]\, d\nu_{\Gamma_n}(\rho) \to \int_{\hat{M}} \mathrm{Tr}[\hat{f}(\rho)]\, d\nu(\rho),$$

where ν is the Plancherel measure. We refer to this type of convergence as *Plancherel convergence.*

Remark V.1. When M is a product of groups, each one a linear reductive group over a local field of characteristic zero, then Plancherel convergence implies that

$$\frac{1}{\mathrm{vol}(\Gamma_n\backslash M)} \nu_{\Gamma_n}(E) \to \nu(E) \tag{V.1.1}$$

for every E which is relatively compact and ν-regular with respect to the Fell topology (see Remark 1.6 of [Dei18] and Theorem 10.2 of [GL18] which both cite the Sauvageot density principle [Sau97]). See also Remark V.4

V.1.2: Benjamini-Schramm convergence

We say that (Γ_n) *Benjamini-Schramm converges to* $\{1\}$ if, for every compact subset $C \subset M$,

$$\frac{\mathrm{vol}(\{x \in \Gamma_n\backslash M : x^{-1}(\Gamma_n \setminus \{1\})x \cap C \neq \emptyset\})}{\mathrm{vol}(\Gamma_n\backslash M)} \to 0.$$

Here in the numerator we are using the volume on $\Gamma_n\backslash M$ since the set in question is clearly $(\Gamma_n, 1)$-invariant.

On the other hand, suppose $(\Gamma_n\backslash G/K)$ is a sequence of compact quotients of the Bruhat-Tits building. As discussed previously, we say this sequence Benjamini-Schramm converges to G/K if, for every $R > 0$,

$$\frac{\mathrm{card}(\{x \in \Gamma_n\backslash G/K : \mathrm{InjRad}_{\Gamma_n\backslash G/K}(x) \leq R\})}{\mathrm{card}(\Gamma_n\backslash G/K)} \to 0.$$

Proposition V.2. *The sequence of lattices* (Γ_n) *Benjamini-Schramm converges to* $\{1\}$ *if and only if the sequence of spaces* $(\Gamma_n\backslash G/K)$ *Benjamini-Schramm converges to* G/K*.*

Proof. The proof of Proposition 2.4 in [Dei18] also works in this case. □

V.1.3: Uniform discreteness

We say that (Γ_n) is *uniformly discrete* if there exists a neighborhood U of the identity such that $x^{-1}\Gamma_n x \cap U = \{1\}$ for every n and for every $x \in M$. In the case that M is a semisimple algebraic group over a non-archimedean local field of zero characteristic, any sequence of lattices is uniformly discrete (see Remark 1.6 of [GL18] and the Remark on p. 1 of [PT21] which both draw from [Mar91]).

V.1.4: Statement of theorem

Theorem V.3 ([Dei18] Theorem 2.6; see also Theorem 1.3 of [GL18])**.** *Suppose (Γ_n) is a sequence of cocompact, uniformly discrete lattices in a locally compact group M. Then the following are equivalent*

(1) (Γ_n) is a Plancherel sequence.

(2) (Γ_n) Benjamini-Schramm converges to $\{1\}$.

V.2: The Fell topology vs. the Euclidean topology on Ω^+_{temp}

Now let $G = \mathrm{PGL}(3, F)$ as before. Let Ω^+ denote the class 1 representations of G, or equivalently the elements in $\hat{G}$ containing a K-fixed vector (see Appendix A.6). Recall that $\Omega^+_{\text{temp}} \subset \Omega^+$ denotes the tempered representations. Elements of Ω^+ all arise as subquotients of representations I_χ as in Proposition A.7 which are induced from (not necessarily unitary) characters χ of a torus. Further, up to the $\mathfrak{S}_3$-action of permuting coordinates, each element in Ω^+ is associated to a unique χ. By Theorem A.17, the Plancherel measure of G restricted to Ω^+ is supported on Ω^+_{temp}, and the χ associated to these latter representations are all unitary characters. We shall let $\hat{G}_P$ denote the collection of all unitarizable irreducible subquotients of any I_χ (the subscript P appears here because $\hat{G}_P$ relates to the minimal parabolic subgroup which consists of all upper triangular matrices in G).

On Ω^+ we have two natural topologies: the Fell topology and the topology of (pointwise) convergence of the underlying χ. We shall call this second topology the Euclidean topology. In this paper we shall work with Euclidean compact subsets of Ω^+_{temp}. In light of Remark V.1, we wish to show that such sets are relatively compact and ν-regular with respect to the Fell topology. Along the way of showing this we shall partially elucidate the relationship between these two topologies.

Note that the Fell topology in general is not Hausdorff. See [Fol95] (specifically Figure 7.3 on p. 247) for an illustrative picture of the Fell topology for $\mathrm{SL}(2, \mathbb{R})$ (of particular note in this picture is that as we move towards one end of the complementary series, we converge to three different points: the trivial representation and two mock discrete series representations). These non-Hausdorff points can be understood as arising from the fact that, e.g., the I_χ may have several distinct irreducible unitarizable subquotients (at most $d!$ of them for $\mathrm{PGL}(d, F)$; see, e.g., [Cas95] Corollary 7.2.3).

The following result of Tadic [Tad83b] explains much of the connection between the two topologies. It essentially tells us that convergence in the Fell topology implies convergence in the Euclidean topology.

Proposition V.4 ([Tad83b], Theorem 5.6)**.** *Suppose (ρ_n, V_n) is a sequence in $\hat{G}$ converging in the Fell topology to some irreducible subquotient of I_χ (i.e. a point in $\hat{G}_P$). Then past some point all of the (ρ_n, V_n) are irreducible subquotients of some I_{χ_n} (i.e. lie in $\hat{G}_P$) and $\chi_n \to \chi$.*

It is known that $\hat{G}$ is first countable ([Tad87], p. 388), and that in first countable spaces, the closure of a set is the set of sequential limit points.

Lemma V.5. *The set $\Omega^+ \subset \hat{G}$ is open in the Fell topology.*

Proof. We shall show that $(\Omega^+)^c$ is closed. Suppose $\rho_n \to \rho$ with $\rho_n \in (\Omega^+)^c$. This means that $\rho \prec \oplus\rho_n$ by Proposition A.10. Suppose for the sake of contradiction that ρ contains a K-fixed vector (i.e. $\rho \in \Omega^+$). We know that $\rho|_K \prec \oplus\rho_n|_K$, and since K is compact, weak containment implies strong containment ([BdlHV08], Appendix F). Since $\rho|_K$ contains a trivial representation, so must one of the $\rho_n|_K$, i.e. one of them contains a K-fixed vector, which would contradict the fact that $\rho_n \notin \Omega^+$. □

Lemma V.6. *Suppose $U \subset \Omega^+$ is open in the Euclidean topology. Then it is also open in the Fell topology.*

Proof. Let $U' = \Omega^+ \setminus U$. This set is clearly closed in the Euclidean topology (on Ω^+). Let's compute its closure in the Fell topology. If $\rho_n \to \rho$ in the Fell topology, then we have $\chi_n \to \chi$ for the associated torus characters by Proposition V.4. But this is the same as convergence in the Euclidean topology, i.e. the class 1 subquotient of I_χ must be in U'. Therefore U' is closed in the Fell topology, so $(U')^c$ is open in the Fell topology. Since Ω^+ is also open, we get that $(U')^c \cap \Omega^+ = U$ is open in the Fell topology. □

Lemma V.7. *If $C \subseteq \Omega^+$ is compact in the Euclidean topology, then it is also compact in the Fell topology.*

Proof. The space Ω^+ is known to be compact in the Fell topology (since K is open; see [Mac71] p. 12). If C is compact in the Euclidean topology, then it is closed in the Euclidean topology and hence closed in the Fell topology. Therefore, as a closed subset of a compact space, C is compact in the Fell topology. □

Lemma V.8. *Suppose $C \subseteq \Omega^+$ is compact in the Euclidean topology. Then C is relatively compact and ν-regular.*

Proof. By Lemma V.7, C is clearly relatively compact (since it is compact). We know that C is regular with respect to Lebesgue measure and the Euclidean topology. Since ν is absolutely continuous with respect to Lebesgue measure by (A.9.1), C must also be ν-regular with respect to the Euclidean topology. Since the Fell topology is finer than the Euclidean topology, we get that C is ν-regular with respect to the Fell topology. □

V.3: Proof of Proposition II.2

Proof of Proposition II.2 (BS Convergence Implies Plancherel Convergence). Suppose $\Theta \subset \Omega^+_{\text{temp}}$ is compact in the Euclidean topology. By Lemma V.8 this implies that Θ is relatively compact and ν-regular. We clearly have $N(\Theta, \Gamma_n) = \nu_{\Gamma_n}(\Theta)$. Therefore by Theorem V.3, we have

$$\frac{N(\Theta, \Gamma_n)}{\text{card}(Y_n)} \to \nu(\Theta).$$

□

V.4: Effective rate of convergence

By unpacking the contents of the proof of Theorem V.3 and the Sauvageot density principle, we may obtain an effective bound on the rate of convergence in Proposition II.2.

More specifically, the Sauvageot density principle tells us the following:

Proposition V.9. *Let E be a ν-measurable, relatively compact subset of $\hat{G}$, with G a reductive algebraic group over a local field of characteristic zero. Let $\varepsilon > 0$. Then there exist functions $\phi, \psi \in C_c^\infty(G)$ such that*

(1) $\left|\mathbb{1}_E(\rho) - \mathrm{Tr}[\hat{\phi}(\rho)]\right| \leq \mathrm{Tr}[\hat{\psi}(\rho)]$ *for all* $\rho \in \hat{G}$.

(2) $\int_{\hat{G}} \mathrm{Tr}[\hat{\psi}]d\nu \leq \varepsilon$.

We therefore get:

$$\begin{aligned}
\Big|\frac{N(E,Y_n)}{\mathrm{vol}(\Gamma_n\backslash G)} - \nu(E)\Big| &= \Big|\frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\int_{\hat{G}} \mathbb{1}_E(\rho)d\nu_{\Gamma_n} - \int_{\hat{G}} \mathbb{1}_E(\rho)d\nu\Big| \\
&= \frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\Big|\int_{\hat{G}}(\mathbb{1}_E(\rho) - \mathrm{Tr}[\hat{\phi}(\rho)]) + \mathrm{Tr}[\hat{\phi}(\rho)]\ d\nu_{\Gamma_n} \\
&\qquad - \Big(\int_{\hat{G}}(\mathbb{1}_E(\rho) - \mathrm{Tr}[\hat{\phi}(\rho)]) + \mathrm{Tr}[\hat{\phi}(\rho)]\ d\nu\Big)\Big| \\
&\leq \frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\int_{\hat{G}} \mathrm{Tr}[\hat{\psi}(\rho)]\ d\nu_{\Gamma_n} + \varepsilon \\
&\qquad + \frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\Big|\int_{\hat{G}} \mathrm{Tr}[\hat{\phi}(\rho)]d\nu_{\Gamma_n} - \int_{\hat{G}} \mathrm{Tr}[\hat{\phi}(\rho)]d\nu\Big| \\
&\leq \Big|\frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\int_{\hat{G}} \mathrm{Tr}[\hat{\psi}(\rho)]\ d\nu_{\Gamma_n} - \int_{\hat{G}} \mathrm{Tr}[\hat{\phi}(\rho)]d\nu\Big| + 2\varepsilon \\
&\qquad + \Big|\frac{1}{\mathrm{vol}(\Gamma_n\backslash G)}\int_{\hat{G}} \mathrm{Tr}[\hat{\phi}(\rho)]d\nu_{\Gamma_n} - \int_{\hat{G}} \mathrm{Tr}[\hat{\phi}(\rho)]d\nu\Big|
\end{aligned}$$

We now follow the argument in [Dei18] p. 9-10 (which relies on the trace formula). Continuing the above inequality, we obtain that there exists an $r \geq 0$ and a compact subset $C \subset G$ such that

$$\Big|\frac{N(E,Y_n)}{\mathrm{vol}(\Gamma_n\backslash G)} - \nu(E)\Big| \leq r(||\phi||_\infty + ||\psi||_\infty)\frac{\mathrm{vol}(\{x : x^{-1}(\Gamma_n \setminus \{1\})x \cap C \neq \emptyset\})}{\mathrm{vol}(\Gamma_n\backslash G)} + 2\varepsilon.$$

CHAPTER VI
Lifting the Kernel Function to G/K

VI.1: The volume of balls in $\mathcal{B}$

Recall that elements in G/K correspond to vertices on the building $\mathcal{B}$. Let $d(\cdot,\cdot)$ be the Euclidean metric on $\mathcal{B}$ normalized so that adjacent vertices are distance 1 apart. With respect to this metric, $\mathcal{B}$ is a CAT(0) space, and hence in particular there is a unique geodesic joining any two points. The group G acts on $\mathcal{B}$ by isometries with respect to this metric on $\mathcal{B}$.

There is a natural measure on $\mathcal{B}$ coming from the Lebesgue measure on each apartment; we shall denote this measure by by $\mathrm{vol}_{d(\cdot,\cdot)}$. Under the above normalization, each chamber has measure $\frac{\sqrt{3}}{4}$ (the area of an equilateral triangle with side lengths 1). Since G acts transitively on vertices, the volume of a ball of radius R centered at any vertex is independent of that vertex. We let $B(x,R)$ denote the ball of radius R centered at vertex x, and we let $\mathrm{vol}_{d(\cdot,\cdot)}(B_R)$ denote the volume of the ball of radius R centered at any vertex.

There is another natural measure on $\mathcal{B}$ coming from the Haar measure on G which is obtained by simply counting the number of vertices in a given set. We shall refer to this as $\mathrm{vol}_{G/K}$ (this is essentially the same as what has previously been referred to as $\mathrm{card}(\cdot)$).

Proposition VI.1 (Lemma 2 in Leuzinger [Leu06])**.** *There exist constants $C, D > 0$ such that for all $R \geq 0$,*

$$\mathrm{vol}_{d(\cdot,\cdot)}(B_R) \leq C \cdot \mathrm{vol}_{G/K}(B_R) \leq \mathrm{vol}_{d(\cdot,\cdot)}(B_{R+D}).$$

On the other hand, Leuzinger [Leu06] also shows that

Proposition VI.2 ([Leu06], p. 487)**.** *There exist constants $C_1, C_2 > 0$ and $\ell \in \mathbb{N}$ such that* *for all* $R \geq 0$*,*

$$C_1 q^{2R} \leq \mathrm{vol}_{G/K}(B_R) \leq C_2 R^\ell q^{2R}.$$

Remark VI.3. In fact Proposition VI.2 can also be derived by using the techniques of Chapter IV. See Figure 6.

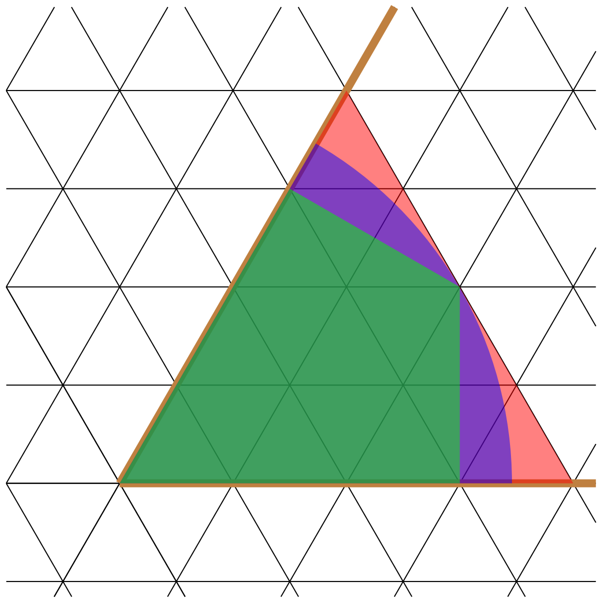


Figure 6: A ball of radius R in the Euclidean metric in the building centered at, say, $1K$ is obtained by taking the K-orbit of the restriction to $\mathfrak{a}^+$ of the ball of radius R centered at $0 \in \mathfrak{a}$; this is represented by the purple circular slice in the figure. This set in contained in the red polytopal region and contains the green polytopal region. By using similar techniques to the computation of the cardinality of E_m (see Chapter IV.4), we obtain that the cardinality of the K-orbit of the lattice points in the Rth dilate of green polytope is of order q^{2R}, and the cardinality of the K-orbit of the lattice points in the Rth dilate of the red polytope is of order Rq^{2R} (one must use degenerate Brion's formula to obtain this bound). These provide upper and lower bounds for $\mathrm{vol}_{G/K}(B_R)$.

Combining these facts, we get the following:

Proposition VI.4. *There exists* $C_1, C_2 > 0$ *and* $\ell \in \mathbb{N}$ *such that for all* R*,*

$$\mathrm{vol}_{d(\cdot,\cdot)}(B_R) \leq C_1 R^\ell q^{C_2 R}.$$

Proposition VI.5. *There exist universal constants* $C_1, C_2 > 0$ *such that for every* R*, for* *every lattice* Γ*, and for every* $z, w \in G/K$*,*

$$N_\Gamma(z, w; R) := \#\{\gamma \in \Gamma : d(z, \gamma.w) \leq R\} \leq \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(\Gamma \backslash G/K)^2}. \tag{VI.1.1}$$

Proof. This essentially just the combination of (the proof of) Lemma 6.18 in [ABB+17] and (the proof of) Lemma 5.1 in [BM21]. For completeness we include the argument here. First we show that it is true when $z = w$. Let $Y = \Gamma \backslash G/K$. First note that the inequality is clearly true when $R < \mathrm{InjRad}_Y(z)$ as then the left hand side is one and, changing the constant C_1 if necessary, the right hand side is bounded from below by one. Hence we may assume $R \geq \mathrm{InjRad}_Y(z)$. We then obtain

$$\begin{aligned} N_\Gamma(z, z; R) \cdot \mathrm{vol}_{d(\cdot,\cdot)}(z, \mathrm{InjRad}_Y(z)) &\leq \mathrm{vol}_{d(\cdot,\cdot)}(B(z, R + \mathrm{InjRad}_Y(z))) \\ &\leq \mathrm{vol}_{d(\cdot,\cdot)}(B(z, 2R)). \end{aligned}$$

By Proposition VI.4, we have that

$$\mathrm{vol}_{d(\cdot,\cdot)}(B(z, 2R)) \leq C_1 q^{C_2 R}.$$

On the other hand, by the definition of $\mathrm{vol}_{d(\cdot,\cdot)}$, we clearly have that the volume of a ball of radius R based at z in $\mathcal{B}$ is greater than the volume of a ball in any apartment containing z, namely πR^2. Therefore,

$$\mathrm{vol}_{d(\cdot,\cdot)}(z, \mathrm{InjRad}_Y(z)) \geq \pi \mathrm{InjRad}_Y(z)^2.$$

Therefore,

$$N(z, z; R) \leq \frac{C_1 q^{C_2 R}}{\pi \mathrm{InjRad}_Y(z)^2},$$

for all $z \in G/K$.

Now suppose $z, w \in G/K$. If $d(z, w) > R$, then the left hand side of (VI.1.1) is one, and the right hand side is bounded from below by a constant (which can be adjusted to be equal to at least one), so we need only handle the case when $d(z, w) \leq R$. Then for any $\gamma \in \Gamma$ such that $d(z, \gamma.w) \leq R$, we also have $d(z, \gamma z) \leq d(z, h) + d(z, \gamma.h) \leq 2R$. Therefore,

$$N(z, w; R) \leq N(z, z; 2R) \leq \frac{C_1 q^{2C_2 R}}{\pi \mathrm{InjRad}_Y(z)^2} \leq \frac{C_1 q^{2C_2 R}}{\pi \mathrm{InjRad}(Y)^2}.$$

$\square$

VI.2: Proof of Lemma II.4

Proof of Lemma II.4 (Lifting the Kernel to G/K). Let D be a fundamental domain in G/K for the action of Γ and let $Y = \Gamma\backslash G/K$. Given $x \in Y$, we let $\tilde{x}$ denote its unique lift to D, and given $w \in D$, we let $\bar{w}$ denote its projection to Y.

Suppose $\mathcal{K} : G/K \times G/K \to \mathbb{C}$ is invariant under the diagonal Γ-action and satisfies $\mathcal{K}(z, w) = 0$ if $d(z, w) > R$ for some $R \geq 0$. Let $\mathcal{K}^{\mathrm{Op}}$ denote the associated operator on $L^2(G/K)$. Suppose $f \in L^2(G/K)$ is $(\Gamma, 1)$-invariant. We claim that its image under this operator is also $(\Gamma, 1)$-invariant:

$$\begin{aligned}
[\mathcal{K}^{\mathrm{Op}}.f](\gamma.z) &= \sum_{w \in G/K} \mathcal{K}(\gamma.z, w) f(w) \\
&= \sum_{w \in G/K} \mathcal{K}(z, \gamma^{-1}.w) f(w) \quad (\mathcal{K} \text{ is invariant under diagonal } \Gamma\text{-action}) \\
&= \sum_{w \in G/K} \mathcal{K}(z, \gamma^{-1}.w) f(\gamma^{-1}.w) \quad (\Gamma\text{-invariance of } f) \\
&= \sum_{u \in G/K} \mathcal{K}(z, u) f(u) \qquad (u = \gamma^{-1}.w) \\
&= [\mathcal{K}^{\mathrm{Op}}.f](z).
\end{aligned}$$

Hence this descends to a well-defined operator on $L^2(Y)$ which we denote $\bar{\mathcal{K}}^{\mathrm{Op}}$. We claim that on this space this operator is represented by the kernel

$$\bar{\mathcal{K}}(x, y) = \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.\tilde{y}),$$

with $x, y \in Y$. To see this:

$$\begin{aligned}
[\mathcal{K}^{\mathrm{Op}}.f](\tilde{x}) &= \sum_{w \in G/K} \mathcal{K}(\tilde{x}, w) f(w) \\
&= \sum_{\tilde{y} \in D} \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.\tilde{y}) f(\gamma.\tilde{y}) \\
&= \sum_{y \in Y} \bar{\mathcal{K}}(x, y) f(y).
\end{aligned}$$

Thus the Hilbert-Schmidt norm of $\bar{\mathcal{K}}^{\mathrm{Op}}$ is

$$||\bar{\mathcal{K}}^{\mathrm{Op}}||_{\mathrm{HS}}^2 = \sum_{y \in Y} \sum_{x \in Y} \Big| \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.\tilde{y}) \Big|^2 = \sum_{\tilde{y} \in D} \sum_{\tilde{x} \in D} \Big| \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.\tilde{y}) \Big|^2.$$

Now we use the assumption that $\mathcal{K}(z, w) = 0$ whenever $d(z, w) > R$ to bound this Hilbert-Schmidt norm. Let $I(R)$ be the points in Y with injectivity radius greater than R. Let $I(R)^C$ be the remaining points in Y. We claim that if $y \in I(R)$, then the injectivity radius at y is greater than R, and thus $d(z, \gamma.\tilde{y}) \leq R$ for at most one $\gamma \in \Gamma$ for a given $z \in G/K$.

Note that $y \in I(R)$ implies that $d(\tilde{y}, \gamma.\tilde{y}) > 2R$ for $\gamma \neq 1$. Suppose $d(z, \tilde{y}) \leq R$. Suppose for the sake of contradiction that $d(z, \gamma.\tilde{y}) \leq R$ for some $\gamma \neq 1$. Then by joining geodesic segments from $\tilde{y}$ to z, then from z to $\gamma.\tilde{y}$, we'd get that $d(\tilde{y}, \gamma.\tilde{y}) \leq 2R$, a contradiction. Hence there is at most one γ such that $d(z, \gamma.\tilde{y}) \leq R$. This implies that for $y \in I(R)$, we have

$$\Big| \sum_{\gamma \in \Gamma} \mathcal{K}(z, \gamma.\tilde{y}) \Big|^2 = \sum_{\gamma \in \Gamma} |K(z, \gamma.\tilde{y})|^2$$

since both sides have at most one term.

We clearly have $|\mathcal{K}(\tilde{x}, \gamma.w)| \leq ||\mathcal{K}||_\infty$, and by Cauchy-Schwarz we have:

$$\begin{aligned} \Big| \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.w) \cdot 1 \Big|^2 &\leq \#\{\gamma \in \Gamma : d(\tilde{x}, \gamma.w) \leq R\} \cdot \sum_{\gamma \in \Gamma} |\mathcal{K}(\tilde{x}, \gamma.w)|^2 \\ &= N_\Gamma(\tilde{x}, w; R) \sum_{\gamma \in \Gamma} |\mathcal{K}(\tilde{x}, \gamma.w)|^2. \end{aligned}$$

Therefore, combining this with Proposition VI.5, we get

$$\begin{aligned} \sum_{\bar{w} \in I(R)^C} \sum_{\tilde{x} \in D} \Big| \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.w) \Big|^2 &\leq \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \sum_{\bar{w} \in I(R)^C} \sum_{\tilde{x} \in D} \sum_{\gamma \in \Gamma} |\mathcal{K}(\tilde{x}, \gamma.w)|^2 \\ &= \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \sum_{\bar{w} \in I(R)^C} \sum_{\tilde{x} \in D} \sum_{\gamma \in \Gamma} |\mathcal{K}(\gamma^{-1}.\tilde{x}, w)|^2 \\ &= \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \sum_{y \in I(R)^C} \sum_{u \in G/K} |\mathcal{K}(u, \tilde{y})|^2 \\ &= \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \sum_{y \in I(R)^C} \sum_{\{u \in G/K : d(u, \tilde{y}) \leq R\}} |\mathcal{K}(u, \tilde{y})|^2 \end{aligned}$$

$$\leq \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \sum_{y \in I(R)^C} ||\mathcal{K}||_\infty^2 \mathrm{vol}_{G/K}(B_R)$$
$$= \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \mathrm{vol}_{G/K}(B_R) ||K||_\infty^2 \mathrm{card}(I(R)^C).$$

Putting everything together and using Proposition VI.4 we get that there exist $C_3, C_4 \geq 0$ such that:

$$\begin{aligned} ||\bar{\mathcal{K}}^{\mathrm{Op}}||_{\mathrm{HS}}^2 &= \sum_{\bar{w} \in I(R)} \sum_{\tilde{x} \in D} \Big| \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.w) \Big|^2 + \sum_{\bar{w} \in I(R)^c} \sum_{\tilde{x} \in D} | \sum_{\gamma \in \Gamma} \mathcal{K}(\tilde{x}, \gamma.w)|^2 \\ &\leq \sum_{\bar{w} \in I(R)} \sum_{\tilde{x} \in D} \sum_{\gamma \in \Gamma} |\mathcal{K}(\tilde{x}, \gamma.w)|^2 + \frac{C_1 q^{C_2 R}}{\mathrm{InjRad}(Y)^2} \mathrm{vol}_{G/K}(B_R) ||\mathcal{K}||_\infty^2 \mathrm{card}(I(R)^C) \\ &\leq \sum_{z \in D} \sum_{w \in G/K} |\mathcal{K}(z, w)|^2 + \frac{C_3 q^{C_4 R}}{\mathrm{InjRad}(Y)^2} ||\mathcal{K}||_\infty^2 \mathrm{card}(\{y \in Y : \mathrm{InjRad}_Y(y) \leq R\}). \end{aligned}$$

$\square$

CHAPTER VII

Changing Variables in the Kernel Function Integral

VII.1: Geometric interpretation of $\Gamma\backslash G/M_\lambda$

Let o denote the point $1K$ in G/K. Given $\lambda \in A^+$, let $z_\lambda := \varpi^\lambda K$. Then $d_{A^+}(o, z_\lambda) = \lambda$. Let M_λ be the joint (pointwise) stabilizer of o and z_λ, namely $M_\lambda = K \cap (\varpi^\lambda K \varpi^{-\lambda})$. Let D be a fundamental domain for the Γ-action on G/K. Recall that we always assume that Γ is torsionfree.

Proposition VII.1. *Assuming that the Haar measure on G is normalized such that* $\mathrm{vol}(K) = 1$*, we have*

$$\mathrm{vol}(M_\lambda) = \frac{1}{\mathrm{vol}(K\varpi^\lambda K)} = \frac{1}{N_\lambda}.$$

Proof. We clearly have $M_\lambda < K$. Consider the right action of K on right cosets of M_λ in K. Given a coset $M_\lambda k$, we may associate to it the point $z_k := k^{-1} z_\lambda$. This point is clearly independent of the choice of coset representative (since M_λ stabilizes z_λ). Furthermore, if $k' \in K$ satisfies $(k')^{-1}.z_\lambda = k^{-1}.z_\lambda$, then $k'k^{-1} \in M_\lambda$, and thus $M_\lambda k = M_\lambda k'$. Therefore we have an injection from the cosets $M_\lambda k$ to the points in the K-orbit of z_λ. However, by the Cartan decomposition we know that the K-orbit of z_λ is exactly those points w satisfying $d_{A^+}(o, w) = \lambda$. Therefore we have a surjection, and hence a bijection, between the cosets of M_λ in K and those points in the K orbit of z_λ, namely $K\varpi^\lambda K$. Therefore, the number of M_λ cosets is exactly equal to $\mathrm{card}(K z_\lambda) = \mathrm{vol}(K\varpi^\lambda K)$. This in turn is equal to the index of M_λ in K. $\square$

Proposition VII.2. *Let D_λ be defined as*

$$D_\lambda := \{(x, y) \in D \times G/K : d_{A^+}(x, y) = \lambda\}.$$

Consider the map

$$\mathfrak{f}_\lambda : D_\lambda \to \Gamma\backslash G/M_\lambda$$

defined as follows: given $(x, y) \in D_\lambda$*, choose any* $g \in G$ *such that* $g.(o, z_\lambda) = (x, y)$*; set*

$$\mathfrak{f}_\lambda(x, y) = \Gamma\backslash g/M_\lambda.$$

Then $\mathfrak{f}_\lambda$ *is well-defined and defines a bijection between* D_λ *and* $\Gamma\backslash G/M_\lambda$*.*

Proof. First we show that a g as in the statement of the proposition exists. We can clearly find some g_1 such that $g_1.x = o$. Then $d_{A^+}(o, g_1.y) = \lambda$. Suppose $g_1.y = hK$. Writing $h = k_1 a k_2$ with $a \in A^+$ (the Cartan decomposition), we must have that $a = \varpi^\lambda$. Thus $k_1^{-1} hK = \varpi^\lambda K$, so we may take $g = g_1^{-1} k_1$.

Next we show that the map $\mathfrak{f}_\lambda$ is well-defined, i.e. does not depend on the choice of g. If g and g' both map (o, z_α) to (x, y), then $g^{-1}g' \in M_\lambda$, so they generate the same double coset.

We now show injectivity of $\mathfrak{f}_\lambda$. Consider the set $G_D := \{g \in G : g.o \in D\}$. Then clearly the collection of g's which arise in the definition of the map $\mathfrak{f}_\lambda$ (for all choices of $(x, y) \in D_\lambda$) is exactly G_D. Suppose $g_1, g_2 \in G_D$ are in the same (Γ, M_λ)-double coset. Then $g_1 = \gamma g_2 m$ with $\gamma \in \Gamma$ and $m \in M_\lambda$. Therefore $g_1.o = \gamma g_2 m.o = \gamma g_2.o$. Since $g_1 \in G_D$, we know that $g_1.o \in D$. Since we also have $g_2 \in G_D$, we must have $\gamma = 1$. Therefore $g_1.o = g_2.o$. We clearly then also have that $g_1.z_\lambda = \gamma g_2 m.z_\lambda = g_2.z_\lambda$. Therefore the double coset $\Gamma g_1 M_\lambda = \Gamma g_2 M_\lambda$ has exactly one preimage, namely $(g_1.o, g_1.z_\lambda)$. Therefore $\mathfrak{f}_\lambda$ is injective.

Lastly we show that $\mathfrak{f}_\lambda$ is surjective. Let g be some double coset representative. Then for a unique $\gamma \in \Gamma$, $\gamma g.o \in D$. Thus this double coset is the image under $\mathfrak{f}_\lambda$ of $(\gamma g.o, \gamma g.z_\lambda) \in D_\lambda$. □

Proposition VII.3. *Suppose* $\{g_i\}$ *is a complete set of double coset reprensenatives of* $\Gamma\backslash G/M_\lambda$*. Then the map:*

$$\mathfrak{j}_\lambda : \Gamma\backslash G/M_\lambda \times M_\lambda \to \Gamma\backslash G, \qquad (g_i, m) \mapsto \Gamma g_i m,$$

is a bijection. Furthermore this identification is measure-preserving when we take M_λ *to have measure* $1/N_\lambda$*.*

Proof. Because Γ is torsionfree (by assumption), no conjugate of Γ can intersect M_λ non-trivially because the intersection would be a compact discrete group which necessarily has torsion. We claim that this implies that every $g \in G$ may be represented uniquely as $g = \gamma g_j m$ with $\gamma \in \Gamma$ and $m \in M_\lambda$. To see this, first note that g_j clearly must be the unique element in $\{g_j\}$ generating the same (Γ, M_λ)-double coset as g. Suppose we have some equation of the following form:

$$\gamma_1 g_j m_1 = \gamma_2 g_j m_2 \iff g_j^{-1}(\gamma_2^{-1}\gamma_1)g_j = m_2 m_1^{-1}.$$

Because we know that $g_j^{-1}\Gamma g_j \cap M_\lambda = \{1\}$, we must have $\gamma_1 = \gamma_2$ and $m_1 = m_2$.

From this we can conclude that $\mathfrak{j}_\lambda$ is injective: if $\Gamma g_j m_1 = \Gamma g_k m_2$, then $\gamma_1 g_j m_1 = \gamma_2 g_k m_2$, and thus $g_j = g_k$ and $m_1 = m_2$.

Lastly we show that $\mathfrak{j}_\lambda$ is surjective. Given a coset Γg, we write $g = \gamma g_j m$. Therefore $\mathfrak{j}_\lambda(g_j, m) = \Gamma g$.

Clearly this map is measure-preserving as $\Gamma\backslash G/M_\lambda$ is a discrete set. □

VII.2: Proof of Proposition II.9

Proof of Proposition II.9 (Changing Variables in the Kernel Integral). Suppose $x \in G/K$. We now wish to rewrite the following integral:

$$\sum_{x\in D}\sum_{y\in G/K}\left|\frac{1}{M}\sum_{m=1}^{M}\frac{1}{\operatorname{card}(E_m)}\sum_{z\in xE_m\cap yE_m} a(z)\right|^2. \tag{VII.2.1}$$

We first focus on rewriting

$$\sum_{z\in xE_m\cap yE_m} a(z).$$

Recall the definition of the following sets:

$$E_m^\lambda := E_m \cap \varpi^\lambda E_m,$$
$$\tilde{E}_m^\lambda := \pi^{-1}(E_m^\lambda).$$

Notice that both sets are setwise invariant under left multiplication by elements in M_λ. Also notice that we have a (non-canonical) identification:

$$\tilde{E}_m^\lambda \simeq E_m^\lambda \times K$$

by choosing coset representatives $\tilde{w}$ for each $w \in E_m^\lambda$ (recall $w \in G/K$). This identification is measure-preserving because E_m^λ is a discrete set.

Notice that if $\mathfrak{f}_\lambda(x,y) = \Gamma g M_\lambda$ as in Proposition VII.2, then $gE_m^\lambda = xE_m \cap yE_m$. We therefore have (recalling that a may be viewed as a $(1,K)$-invariant function on $\Gamma\backslash G$):

$$\begin{aligned}
[\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g) &= \frac{1}{\mathrm{vol}(\tilde{E}_m^\lambda)}\int_{\tilde{E}_m^\lambda} a(\Gamma g h)dh \\
&= \frac{1}{\mathrm{vol}(\tilde{E}_m^\lambda)}\sum_{w\in E_m^\lambda}\int_K a(\Gamma g\tilde{w}k)dk \\
&= \frac{1}{\mathrm{vol}(\tilde{E}_m^\lambda)}\sum_{w\in E_m^\lambda} a(\Gamma g\tilde{w}) \\
&= \frac{1}{\mathrm{vol}(\tilde{E}_m^\lambda)}\sum_{z\in gE_m^\lambda} a(\Gamma \tilde{z}) \\
&= \frac{1}{\mathrm{vol}(\tilde{E}_m^\lambda)}\sum_{z\in xE_m\cap yE_m} a(z).
\end{aligned}$$

Let $\{g_i\}$ be a complete set of coset representatives for $\Gamma\backslash G/M_\lambda$. Thus our original integral from (VII.2.1) can now be written as

$$\begin{aligned}
&\sum_{x\in D}\sum_{y\in G/K}\Big|\frac{1}{M}\sum_{m=1}^{M}\frac{1}{\mathrm{card}(E_m)}\sum_{z\in xE_m\cap yE_m} a(z)\Big|^2 \\
=&\frac{1}{M^2}\sum_{\lambda\in A^+}\sum_{g_i\in\Gamma\backslash G/M_\lambda}\Big|\sum_{m=1}^{M}\frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)}\rho^\Gamma_{\tilde{E}_m^\lambda}a(\Gamma g_i)\Big|^2.
\end{aligned}$$

We now utilize Proposition VII.3:

$$\begin{aligned}
\int_{\Gamma\backslash G}\Big|\sum_{m=1}^{M}\frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)}[\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g)\Big|^2 dg &= \sum_{g_i\in\Gamma\backslash G/M_\lambda}\sum_{n\in M_\lambda}\Big|\sum_{m=1}^{M}\frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)}[\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g_i n)\Big|^2 \\
&= \sum_{g_i\in\Gamma\backslash G/M_\lambda}\sum_{n\in M_\lambda}\Big|\sum_{m=1}^{M}\frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)}[\rho^\Gamma_{n.\tilde{E}_m^\lambda}.a](\Gamma g_i)\Big|^2 \\
&= \sum_{g_i\in\Gamma\backslash G/M_\lambda}\sum_{n\in M_\lambda}\Big|\sum_{m=1}^{M}\frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)}[\rho^\Gamma_{\tilde{E}_m^\lambda}.a](\Gamma g_i)\Big|^2
\end{aligned}$$

$$= \sum_{g_i \in \Gamma\backslash G/M_\lambda} \mathrm{vol}(M_\lambda) \Big| \sum_{m=1}^{M} \frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)} [\rho^{\Gamma}_{\tilde{E}_m^\lambda}.a](\Gamma g_i) \Big|^2 .$$

Here we have used that $\tilde{E}_m^\lambda$ is invariant (as a set) under M_λ.

We therefore obtain that

$$\begin{aligned} &\frac{1}{M^2} \sum_{\lambda \in A^+} \sum_{g_i \in \Gamma\backslash G/M_\lambda} \Big| \sum_{m=1}^{M} \frac{\mathrm{vol}(\tilde{E}_m^\lambda)}{\mathrm{card}(E_m)} [\rho^{\Gamma}_{\tilde{E}_m^\lambda}.a](\Gamma g_i) \Big|^2 \\ =&\frac{1}{M^2} \sum_{\lambda \in A^+} N_\lambda \int_{\Gamma\backslash G} \Big| \sum_{m=1}^{M} \frac{\mathrm{card}(E_m^\lambda)}{\mathrm{card}(E_m)} [\rho^{\Gamma}_{\tilde{E}_m^\lambda}.a](\Gamma g) \Big|^2 dg. \end{aligned}$$

□

CHAPTER VIII

The Kunze-Stein Phenomenon and an Ergodic Theorem in the Style of Nevo

VIII.1: The Kunze-Stein phenomenon

Suppose M is a locally compact topological group. Let $\mathcal{M}(M)$ denote the collection of measurable functions on M. We say that M *satisfies the Kunze-Stein phenomenon* (or is a *KS group*, for short), if for every $1 \leq p < 2$, the map $L^p(M) \times L^2(M) \to \mathcal{M}(M)$ defined by convolution is continuous and has image contained in $L^2(M)$. This means that, for a fixed p, there exists a C_p such that for all $f \in L^p(M)$ and $g \in L^2(M)$ we have

$$||f * g||_2 \leq C_p ||f||_p ||g||_2. \tag{VIII.1.1}$$

Another way of expressing this is that for all $f \in L^p(M)$ we have $||\hat{f}(\lambda_M)|| \leq C_p ||f||_p$, where λ_M is the (left) regular representation and $\hat{f}(\lambda_M)$ is the operator on $L^2(M)$ given by (left) convolution with f.

VIII.2: The Kunze-Stein phenomenon and finite group extensions

Lemma VIII.1. *If M is a KS group, and $N \lhd M$ is a finite normal subgroup, then M/N is a KS group.*

Proof. Let $\pi : M \to M/N$ be the projection map. Let the Haar measures vol_M and $\mathrm{vol}_{M/N}$ on M and M/N be normalized such that for all $U \subset M/N$ measurable, $\mathrm{vol}_M(\pi^{-1}(U)) = \mathrm{vol}_{M/N}(U)$ (notice then that $\mathrm{vol}_{M/N}(\pi(V)) = \frac{\mathrm{vol}_M(V \cdot N)}{|N|}$ for $V \subset G$ measurable). Given $a \in \mathcal{M}(M/N)$, let $\tilde{a} \in \mathcal{M}(M)$ be the lift of a to a $(1, N)$-invariant function on M. Because

N is normal, $\tilde{a}$ must in fact be (N,N)-invariant:

$$\tilde{a}(nm) = \tilde{a}\Big(m(m^{-1}nm)\Big) = \tilde{a}(m),$$

with $n \in N$. The normalization of Haar measures makes it so that

$$\int_M \tilde{a}(m)dm = \int_{M/N} a(\ell N)d\ell.$$

We claim that we have the following commutative diagram:

$$\begin{array}{ccc} L^p(M/N) \times L^2(M/N) & \xrightarrow{\pi^*} & L^p(M) \times L^2(N) \\ \downarrow{*_{M/N}} & & \downarrow{*_M} \\ \mathcal{M}(M/N) & \xrightarrow{\pi^*} & \mathcal{M}(M). \end{array}$$

Let $a, b \in \mathcal{M}(M/N)$. We have:

$$\begin{aligned} [\tilde{a} *_M \tilde{b}](x) &= \int_M \tilde{a}(m)\tilde{b}(m^{-1}x)dm \\ &= \int_{M/N}\int_N \tilde{a}(yn)\tilde{b}(n^{-1}y^{-1}x)dndy \\ &= \int_{M/N} a(yN)b(y^{-1}xN)dy \\ &= [a *_{M/N} b](xN) \\ &= [a \widetilde{*_{M/N}} b](x). \end{aligned}$$

Therefore $\tilde{a} *_M \tilde{b} = a \widetilde{*_{M/N}} b$.

Now suppose $a \in L^p(M/N)$ and $b \in L^2(M/N)$; then $\tilde{a} \in \widetilde{L^p(M)}$ and $\tilde{b} \in \widetilde{L^2(N)}$. Since M is a KS group, $\tilde{a} *_M \tilde{b} \in L^2(M)$. On the other hand $\tilde{a} *_M \tilde{b} = a \widetilde{*_{M/N}} b$. Since $a \widetilde{*_{M/N}} b \in L^2(M)$, we must have $a *_{M/N} b \in L^2(M/N)$. Therefore M/N is a KS group. Because we may directly relate the L^p-norms of a, b and $\tilde{a}, \tilde{b}$, it is clear that we can find a C_p as in (VIII.1.1) using the KS property for M. □

Lemma VIII.2. *If N is a KS group, and $N \lhd M$ is a normal finite index subgroup, then M is a KS group.*

Proof. Let the Haar measure on M be such that if $U \subseteq N$ measurable, then $\mu_N(U) = \mu_M(U)$. This convention implies that integration on M of a function supported on a single coset of N is the same as integration of the analogous function on N. Suppose $a \in L^p(M)$ and $b \in L^2(M)$. Let $x_1N, \ldots, x_\ell N$ be the cosets of N. Write $a = a_1 + \cdots + a_\ell$ and $b = b_1 + \cdots + b_\ell$ for the decomposition of a and b into their components supported on each coset. Clearly $a \in L^p(M)$ if and only if $a_i \in L^p(N)$ for all i.

Let's consider $a_i * b_j$. We claim that this is supported on the coset $x_i x_j N$. Let $n \in N$. We have

$$a_i * b_j(x_k n) = \int_M a_i(m) b_j(m^{-1} x_k n) dm.$$

For this integral to be non-zero, we need $m = x_i n_1$ and $m^{-1} x_k n = n_1^{-1} x_i^{-1} x_k n = x_j n_2$ for some $n_1, n_2 \in N$. Since N is normal, this is the same as $x_i^{-1} x_k N = x_j N$, which is the same as $x_k N = x_i x_j N$. Hence, if $x_k N = x_i x_j N$, we get that the above integral reduces to just the integral over $m \in x_i N = N x_i$. Therefore,

$$\begin{aligned} a_i * b_j(n x_i x_j) &= \int_{N x_i} a_i(m) b_j(m^{-1} n x_i x_j) dm \\ &= \int_N a_i(\ell x_i) b_j(x_i^{-1} \ell^{-1} n x_i x_j) d\ell. \end{aligned}$$

Define the following functions on N:

$$\begin{aligned} \hat{a}_i(n) &:= a_i(n x_i), \\ \hat{b}_j(n) &:= b_j(x_i^{-1} n x_i x_j). \end{aligned}$$

Then,

$$\begin{aligned} \hat{a}_i * \hat{b}_j(n) &= \int_N \hat{a}_i(\ell) \hat{b}_j(\ell^{-1} n) d\ell \\ &= \int_N a_i(\ell x_i) b_j(x_i^{-1} \ell^{-1} n x_i x_j) dn. \end{aligned}$$

The above shows that convolution of a_i and b_j on M, which is supported on a single coset of N, is the equivalent to convolution of analogous functions on N. In particular, $\hat{a}_i * \hat{b}_j \in L^2(N)$ implies that $a_i * b_j \in L^2(M)$. By using the distributivity of convolution, we get that since N satisfies KS, M also satisfies KS. Because of the explicit relationship between convolution in N and M, it is clear that we can find a C_p as in (VIII.1.1) using the KS property for N. $\square$

VIII.3: The Kunze-Stein phenomenon for $\mathrm{PGL}(d, F)$

We now wish to show that $G = \mathrm{PGL}(d, F)$ is a KS group. We utilize the above lemmas and the following results.

Theorem VIII.3 (Veca [Vec02])**.** *Let F be a non-discrete, totally disconnected local field, and let M be the group of F-rational points of a simply connected algebraic group defined over F. Then M is a KS group.*

In particular this applies to the group $\mathrm{SL}(d, F)$.

Proposition VIII.4 ([Ser79], p. 214)**.** *If d is relatively prime to the characteristic of F, then $F^\times/(F^\times)^d$ has finite order.*

Corollary VIII.5. *The group $\mathrm{PGL}(d, F)$ is a KS group as long as d is relatively prime to the characteristic of F.*

Proof. We have the following long exact sequence:

$$1 \to \mu_d(F) \to \mathrm{SL}(d, F) \to \mathrm{PGL}(d, F) \to F^\times/(F^\times)^d \to 1,$$

where $\mu_d(F)$ is the group of dth roots of unity of F. Since $\mathrm{SL}(d, F)$ is a KS group by Theorem VIII.3, and since $\mu_d(F) \lhd \mathrm{SL}(d, F)$ is finite, $\mathrm{SL}(d, F)/\mu_d(F)$ is also a KS group by Lemma VIII.1. This group sits inside $\mathrm{PGL}(d, F)$ as a finite index normal subgroup, so by Lemma VIII.2, we get that $\mathrm{PGL}(d, F)$ is also a KS group as long as $F^\times/(F^\times)^d$ is finite (which occurs if, e.g., d is relatively prime to the characteristic of F by Proposition VIII.4). □

VIII.4: Proof of Proposition II.11

Let M be a semisimple algebraic group over a local field.

Lemma VIII.6. *Suppose $(\rho, \mathcal{H})$ is a unitary representation of M. Let $\psi \in L^1(M)$ be such that $\int_M \psi dg = 1$ and ψ is non-negative and real-valued. Then for any even positive integer $2m$,*

$$||\hat{\psi}(\rho)||^{2m} \leq ||\hat{\psi}(\rho^{\otimes 2m})||.$$

Proof. This essentially follows from the proof of Theorem 1 in Nevo [Nev98]. The main idea here is Jensen's inequality (and the convexity of the function t^{2m}). □

Recall the definition of integrability exponenent in Appendix A.11. It follows from Hölder's inequality that if ρ has integrability exponent t, then $\rho^{\otimes n}$ has integrability exponent t/n (the product of a function in $L^p(M)$ and a function in $L^q(M)$ is in $L^r(M)$ where $1/r = 1/p + 1/q$).

Corollary VIII.7. *Let $(\rho, \mathcal{H})$ be a unitary representation of M with integrability exponent $q(\rho) < \infty$. Let m be an integer such that $q(\rho) < 4m$. Then $\rho^{\otimes 2m}$ is weakly contained in the regular representation.*

Proof. The integrability exponenent of $\rho^{\otimes 2m}$ is $q(\rho)/2m \leq 2$, hence $\rho^{\otimes 2m} \prec \lambda_M$ by Proposition A.20. □

Lemma VIII.8. *Suppose M is a KS group. Let ρ be a unitary representation with almost integrability exponent $q(\rho)$. Let m be an integer such that $q(\rho) < 4m$. Let $1 \leq p < 2$. Suppose $\psi \in L^1(M) \cap L^p(M)$ with ψ non-negative, real-valued, and $\int_M \psi dg = 1$. Then there exists C_p such that*

$$||\hat{\psi}(\rho)|| \leq C_p ||\psi||_p^{1/2m}.$$

Proof. By Corollary VIII.7, we have that $\rho^{\otimes 2m}$ is weakly contained in the regular representation. From Lemma VIII.6 and Proposition A.9, we have that

$$||\hat{\psi}(\rho)||^m \leq ||\hat{\psi}(\rho^{\otimes 2m})|| \leq ||\hat{\psi}(\lambda_M)||.$$

By the KS property, we get that there exists a C_p depending only on p such that for all $\psi \in L^p(G)$,

$$||\hat{\psi}(\lambda_M)|| \leq C_p^{2m} ||\psi||_p.$$

Combining these inequalities, we get that

$$||\hat{\psi}(\rho)|| \leq C_p ||\psi||_p^{1/2m}.$$

□

Theorem VIII.9. *Suppose M is a KS group. Suppose M acts ergodically on the probability space (Y, ν) and with finite integrability exponent on $L_0^2(Y, \nu)$. Then there exist $C > 0$ depending only on M, and $\theta > 0$ depending only on the integrability exponent, such that for* *every measurable subset $E \subset M$ with finite positive measure and every $f \in L^2_0(Y, \nu)$,*

$$\Big|\Big|\frac{1}{\mathrm{vol}(E)} \int_E f(m^{-1}.x)dm\Big|\Big|_2 \leq C\mathrm{vol}(E)^{-\theta}||f||_2. \tag{VIII.4.1}$$

Proof. The function $\psi = \frac{1}{\mu(E)}\mathbb{1}_E$ is in $L^1(M) \cap L^p(M)$ for every p and is non-negative, real, and has L^1-norm equal to 1. Therefore, if $1 \leq p < 2$ we have

$$||\hat{\psi}(\rho)|| \leq C_p||\psi||_p^{1/2m}.$$

Notice that in fact $\hat{\psi}(\rho)$ is exactly the expression inside the $||\cdot||_2$ on the left hand side of (VIII.4.1).

We also have that

$$||\psi||_p = \mathrm{vol}(E)^{-1+1/p}.$$

Therefore,

$$||\hat{\psi}(\rho)|| \leq C_p\mathrm{vol}(E)^{-1/(2m)+1/(2pm)}.$$

We see that in fact we can choose θ to be of the form $-1/(4m) + \delta$ for any $\delta > 0$. □

Proof of Proposition II.11 (Nevo-Style Ergodic Theorem). This follows immediately from Theorem VIII.9. In particular, we take $(Y, \nu) = L^2(\Gamma\backslash G)$ with the underlying G-action given by ρ^Γ. Then if a is a mean-zero function, we have that $a \in L^2_0(\Gamma\backslash G)$. Then (VIII.4.1) exactly turns into

$$||\rho^\Gamma_E.a||_{L^2(\Gamma\backslash G)} \leq \frac{C}{\mathrm{vol}(E)^\theta}||a||_{L^2(\Gamma\backslash G)}.$$

Proposition A.21 tells us that we can in fact take the same θ for all Γ. □

CHAPTER IX
Classification of Primitive Triples of Vertices in $\mathcal{B}$

IX.1: Parallelograms in the Coxeter complex

Let $\mathfrak{a}$ denote the space $\{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}$, and let $\mathfrak{a}^+$ denote the standard Weyl chamber in $\mathfrak{a}$, namely $\{(x_1, x_2, x_3) \in \mathfrak{a} : x_1 \geq x_2 \geq x_3\}$. We let $\mathfrak{a}^{++}$ denote the regular elements in $\mathfrak{a}^+$, namely those elements for which $x_1 > x_2 > x_3$. There is a bijection between the Weyl chambers (centered at $0 \in \mathfrak{a}$) and elements of $\mathfrak{S}_3$; it arises from the $\mathfrak{S}_3$-action on coordinates (see Appendix B.3).

We may tesselate $\mathfrak{a}$ by equilateral triangles in such a way that, treating the resulting object as a simplicial complex $\mathcal{X}$, we obtain the Coxeter complex associated to $\tilde{A}_2$ (see Appendix B.6.4), and the vertices of this simplicial complex correspond to the lattice Λ. Let $\Lambda^+ = \Lambda \cap \mathfrak{a}^+$. As discussed previously, there is a natural identification of Λ with $A < G$, and Λ^+ with $A^+ \subset G$.

Given $x \in \Lambda$ and $\sigma \in \mathfrak{S}_3$, we define

$$S(x; \sigma) := \{y \in \Lambda : y - x \in \sigma.\mathfrak{a}^+\}.$$

This defines a sector based at x which is a translated copy of the sector $\sigma.\mathfrak{a}^+$ based at 0. We shall call two sectors $S(x_1; \sigma_1)$ and $S(x_2; \sigma_2)$ *parallel* if $\sigma_1 = \sigma_2$.

Let $\sigma^{\max} = (1\ 3)$ denote the element in $\mathfrak{S}_3$ with greatest Coxeter word length. Then, given $\sigma \in \mathfrak{S}_3$, the Weyl chambers based at 0 associated to σ and to $\sigma.\sigma^{\max}$ are *opposite* in the spherical Coxeter complex associated to $\mathfrak{S}_3$ (which we may identify with Weyl chambers based at 0).

Proposition IX.1. *Suppose $a, b \in \mathfrak{a}^+$. If $\sigma_1.a + \sigma_2.b = a + b$ with $\sigma_1, \sigma_2 \in \mathfrak{S}_3$, then $\sigma_1.a = a$ and $\sigma_2.b = b$. If $a, b \in \mathfrak{a}^{++}$, then $\sigma_1 = \sigma_2 = 1$.*

Proof. Suppose $a = (a_1, a_2, a_3)$ and $b = (b_1, b_2, b_3)$. Since $a_1 \geq a_2 \geq a_3$ and $b_1 \geq b_2 \geq b_3$, if $a_{\sigma_1^{-1}(1)} + b_{\sigma_2^{-1}(1)} = a_1 + b_1$, we must have $a_1 = a_{\sigma_1^{-1}(1)}$ and $b_1 = b_{\sigma_2^{-1}(1)}$ because $a_1 + b_1$ is the greatest possible value that $\alpha_1.a + \alpha_2.b$ could have for its first entry for all choices of $\alpha_1, \alpha_2 \in \mathfrak{S}_3$. We must also have $a_{\sigma_1^{-1}(2)} + b_{\sigma_2^{-1}(2)} = a_2 + b_2$, but we know that the largest a coordinate and the largest b coordinate have already occurred in $a_{\sigma_1^{-1}(1)} + b_{\sigma_2^{-1}(1)}$, so we must have $a_2 = a_{\sigma_1^{-1}(2)}$ and $b_2 = b_{\sigma_2^{-1}(2)}$. Similarly we have $a_{\sigma_1^{-1}(3)} + b_{\sigma_2^{-1}(3)} = a_3 + b_3$, but we know that the two largest a coordinates and two largest b coordinates have already occurred in forming the first two entries of $\sigma_1.a + \sigma_2.b$, and hence $a_3 = a_{\sigma_1^{-1}(3)}$ and $b_3 = b_{\sigma_2^{-1}(3)}$. □

Given $x, y, z \in \Lambda$, we say that $(x, y; z)$ is an *additive triple* if

$$d_{A^+}(x, z) + d_{A^+}(z, y) = d_{A^+}(x, y). \tag{IX.1.1}$$

Given x and y, we define $\operatorname{cone}(x, y)$, to be the intersection of all sectors based at x containing y. We define the *parallelogram* of x and y by $\operatorname{para}(x, y) = \operatorname{cone}(x, y) \cap \operatorname{cone}(y, x)$.

Proposition IX.2. *Suppose $x, y, z \in \Lambda$. The following are equivalent:*

(1) $(x, y; z)$ is an additive triple.

(2) For every sector $S(x; \sigma)$ based at x containing y, we have $z \in S(x; \sigma)$ and $y \in S(z; \sigma)$.

(3) $z \in \operatorname{para}(x, y)$.

Proof. (1) $\implies$ (2): Suppose $(x, y; z)$ is an additive triple. Without loss of generality suppose $x = 0$, and $y \in \mathfrak{a}^+$ (we can always apply some translation followed by some element in $\mathfrak{S}_3$ to achieve this; doing so preserves the Weyl chamber-valued metric). Then $S(x; \sigma) = \mathfrak{a}^+$. We now have reduced to showing that $z \in \mathfrak{a}^+$ and $y - z \in \mathfrak{a}^+$. We clearly have that $z + (y - z) = y = d_{A^+}(x, z) + d_{A^+}(z, y)$. Furthermore, clearly $z = \sigma_1.d_{A^+}(x, z)$ and $y - z = \sigma_2.d_{A^+}(y, z)$ for some $\sigma_1, \sigma_2 \in \mathfrak{S}_3$. However by Proposition IX.1 we conclude that $z = \sigma_1^{-1}.d_{A^+}(x, z) = d_{A^+}(x, z)$ and $y - z = \sigma_2^{-1}.d_{A^+}(y, z) = d_{A^+}(y, z)$. Therefore, we have that z and $y - z$ are in $\mathfrak{a}^+$.

(2) $\implies$ (1): Suppose z is contained in every sector based at x containing y, and y is in the parallel sector based at z. Again, we may assume that $x = 0$ and the sector based at x containing y is $\mathfrak{a}^+$. Then we have $d_{A^+}(x, z) = z$, $d_{A^+}(x, y) = y$, and $d_{A^+}(z, y) = y - z$, so we clearly have $d_{A^+}(x, z) + d_{A^+}(z, y) = d_{A^+}(x, y)$.

(2) $\implies$ (3): In general we have that $b \in S(a; \sigma)$ if and only if $a \in S(b; \sigma.\sigma^{\max})$. Suppose z is such that every for every $y \in S(x; \sigma)$, we have $z \in S(x; \sigma)$ and $y \in S(z; \sigma)$. Then clearly $z \in \operatorname{cone}(x, y)$. We also wish to show that $z \in \operatorname{cone}(y, x)$. Thus, suppose $x \in S(y; \tau)$

for some $\tau \in \mathfrak{S}_3$. Then $y \in S(x; \tau.\sigma^{\max})$, so $y \in S(z; \tau.\sigma^{\max})$, so $z \in S(y; \tau)$. Therefore $z \in \operatorname{cone}(y, x)$. Thus $z \in \operatorname{para}(x, y)$.

(3) $\implies$ (2): Suppose $z \in \operatorname{para}(x, y)$. Let $S(x; \sigma)$ be some sector containing y. Then $z \in S(x; \sigma)$. Furthermore, we know that $z \in S(y; \sigma.\sigma^{\max})$ as this sector contains x. However, $z \in S(y; \sigma.\sigma^{\max})$ in turn implies that $y \in S(z; \sigma)$.

□

Remark IX.3. Suppose $y = (r, s) \in \Lambda^+$ (in cone coordinates). Then clearly $d_{A^+}(0, y) = (r, s)$. Then clearly

$$\operatorname{para}(0, y) = \{(r', s') \in \Lambda : (0, 0) \preceq (r', s') \preceq (r, s)\}.$$

This follows from the fact any such point in the set on the right hand side satisfies Property (2) of Proposition IX.2, and, by Property (1) of IX.2, any point in $\operatorname{para}(x, y)$ must be as in the set on the right hand side. Because up to translation and the action of $\mathfrak{S}_3$, any pair of vertices in $\mathcal{X}$ can be converted to $(0, y)$ as above, we see that $\operatorname{para}(x, y)$ is indeed a (possibly degenerate) parallelogram in the Euclidean sense and with x and y at opposite corners.

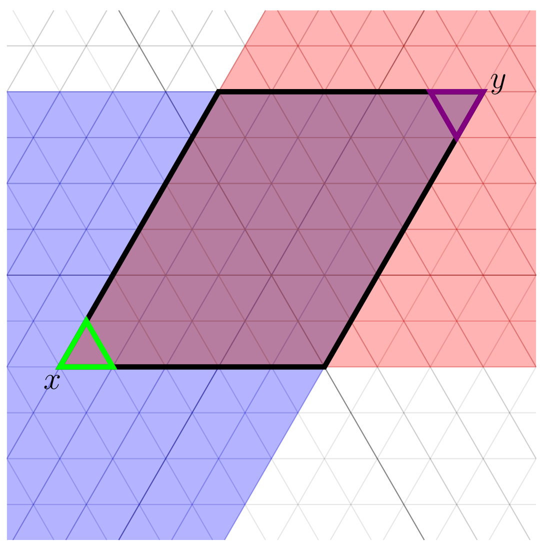


Figure 7: The parallelogram of x and y is the intersection of $\operatorname{cone}(x, y)$ (the red sector) and $\operatorname{cone}(y, x)$ (the blue sector). In this case $d_{A^+}(x, y) \in \mathfrak{a}^{++}$, and the germs of the cones give us $\mathfrak{c}_{x,y}$ (the green chamber) and $\mathfrak{c}_{y,x}$ (the magenta chamber).

Proposition IX.4. *Suppose* $z \in \operatorname{para}(x, y)$. *Then* $\operatorname{para}(x, z) \subseteq \operatorname{para}(x, y)$.

Proof. Without loss of generality, suppose $x = 0$, and $y \in \mathfrak{a}^+$. Then $z \in \mathfrak{a}^+$ and $y \in S(z; 1)$. Suppose $w \in \mathrm{para}(x, z)$. Then $z \in S(w; 1)$. Therefore $S(z; 1) \subseteq S(w; 1)$ and therefore $y \in S(w; 1)$. Therefore $w \in \mathrm{para}(x, y)$. □

Proposition IX.5. *Suppose $d_{A^+}(x, y) = (r, s)$ and $d_{A^+}(y, z) = (1, 0)$. Then*

$$d_{A^+}(x, z) \in \{(r + 1, s), (r, s - 1), (r - 1, s + 1)\}. \qquad \text{(IX.1.2)}$$

Suppose $d_{A^+}(y, z) = (0, 1)$. Then

$$d_{A^+}(x, z) \in \{(r, s + 1), (r - 1, s), (r + 1, s - 1)\}. \qquad \text{(IX.1.3)}$$

Proof. The vector $(1, 0)$ in cone coordinates corresponds to $(2/3, -1/3, -1/3)$ in $\mathfrak{a}$-coordinates. Its orbit under $\mathfrak{S}_3$ also contains $(-1/3, 2/3, -1/3)$ and $(-1/3, -1/3, 2/3)$ which corresponds to $(-1, 1)$ and $(0, -1)$ in cone coordinates. We clearly have $d_{A^+}(x, z) = d_{A^+}(x, y) + \sigma.d_{A^+}(y, z)$ for some $\sigma \in \mathfrak{S}_3$, from which the result follows (a similar calculation can be done for the $\mathfrak{S}_3$-orbit of $(0, 1)$). See also Figure 8. □

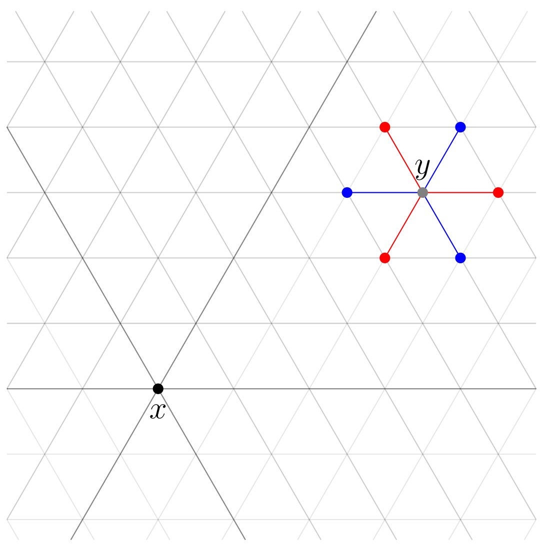


Figure 8: The figure illustrates Proposition IX.5: suppose the black point x has cone coordinates $(0, 0)$. The gray point y has three neighbors z for which $d_{A^+}(y, z) = (1, 0)$ (the red points), and three for which $d_{A^+}(y, z) = (0, 1)$ (the blue points). It is straightforward from this diagram to compute what $d_{A^+}(0, z)$ is for these points.

We define a *combinatorial path* in $\mathcal{X}$ to be any path between vertices along edges, and we define the length of such paths to be the number of edges traversed in the path. We define a *combinatorial geodesic* between two vertices x and y to be any shortest combinatorial path from x to y. We define the *combinatorial distance* between vertices, $d_c(x,y)$, to be the length of any combinatorial geodesic connecting them. Notice that $d_c(x,y)=1$ if and only if x and y are adjacent, if and only if $d_{A^+}(x,y)=(1,0)$ or $(0,1)$.

Proposition IX.6. *Suppose $x,y\in\Lambda$. Then $d_{A^+}(x,y)=(r,s)$ for some $r+s=n$ if and only if $d_c(x,y)=n$. Furthermore, the union of all vertices appearing along combinatorial geodesics from x to y is exactly* $\mathrm{para}(x,y)$.

Proof. We proceed by induction on n. The case of $n=1$ is obvious. Now suppose $d_{A^+}(x,y)=(r,s)$ with $r+s=n$. By the inductive hypothesis, we couldn't possibly have $d_c(x,y)\leq n-1$ as that would imply that $r+s\leq n-1$. Thus $d_c(x,y)\geq n$. On the other hand, y must have a neighbor z satisfying either $d_{A^+}(x,z)=(r-1,s)$ or $(r,s-1)$ (this can be seen by the fact that if $y=(r',s')\in\mathfrak{a}^+$ in cone coordinates, then $d_{A^+}(0,y)=(r',s')$ and at least one of the points in $\{(r'-1,s'),(r',s'-1)\}$ also lies in $\mathfrak{a}^+$). By the inductive hypothesis, this implies that there is a combinatorial geodesic from x to z of length $n-1$, and hence a combinatorial geodesic from x to z of length n. Hence $d_c(x,y)\leq n$, and thus must be exactly n.

Now suppose $d_c(x,y)=n$. Suppose $d_{A^+}(x,y)=(r,s)$. By the inductive hypothesis, we must have $r+s\geq n$. Let z be the penultimate vertex along some combinatorial geodesic from x to y. As this is a neighbor of y, by Proposition IX.5 we must have $d_{A^+}(x,z)\in\{(r-1,s),(r+1,s),(r-1,s+1),(r+1,s-1),(r,s+1),(r+1,s)\}$. Furthermore $d_c(x,z)=n-1$, so the sum of the entries of $d_{A^+}(x,z)$ must be exactly $n-1$ by the inductive hypothesis. Hence the only possibility is $d_{A^+}(x,z)\in\{(r-1,s),(r,s-1)\}$ in which case we must have $r+s=n$.

Now suppose z is the penultimate vertex of some combinatorial geodesic from x to y. Then by the inductive hypothesis, all but the last step of this combinatorial geodesic must lie in $\mathrm{para}(x,z)$. Furthermore, by the analysis in the preceding paragraph, it is clear that we have $d_{A^+}(x,z)+d_{A^+}(z,y)=d_{A^+}(x,y)$. Hence $z\in\mathrm{para}(x,y)$ by Proposition IX.2. Therefore all of this combinatorial geodesic lies in $\mathrm{para}(x,y)$. On the other hand, if $w\in\mathrm{para}(x,y)$ with $d_{A^+}(x,w)=(r',s')$, then by the fact that $(x,y;w)$ is an additive triple, we can find a combinatorial geodesic from x to w of length $r'+s'$, and a combinatorial geodesic from w to x of length $r+s-r'-s'$; the concatenation of these combinatorial geodesics has length $r+s$, implying that it is also a combinatorial geodesic. Therefore the elements in $\mathrm{para}(x,y)$ are exactly the vertices that appear along combinatorial geodesics joining x to y. ☐

Proposition IX.7. *Suppose $x,y\in\Lambda$ are such that $d_{A^+}(x,y)=(r,s)\in\mathfrak{a}^{++}$ (i.e. $r,s\geq 1$). Let γ be the (Euclidean) geodesic joining x and y. Let $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$ be the unique chambers* *that γ passes through which contain x and y, respectively. Then the combinatorial convex hull of the chambers $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$ is exactly* $\mathrm{para}(x,y)$.

Proof. See Appendix B.1 for the definition of the combinatorial convex hull and for the definition of roots. The combinatorial convex hull is the intersection of all roots containing $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$. Furthermore, all roots are of one of the following forms (using the coordinates on $\mathfrak{a}$ of the form $(\alpha_1, \alpha_2, \alpha_3)$; a is some constant):

(1) $\alpha_1 - \alpha_2 \geq a$ or $\alpha_1 - \alpha_2 \leq a$.

(2) $\alpha_2 - \alpha_3 \geq a$ or $\alpha_2 - \alpha_3 \leq a$.

(3) $\alpha_1 - \alpha_3 \geq a$ or $\alpha_1 - \alpha_3 \leq a$.

Without loss of generality, suppose $x = 0$ and $y \in \mathfrak{a}^+$. Then y satisfies $y_1 \geq y_2 \geq y_3$. The vertices of $\mathfrak{c}_{x,y}$ are 0 and $(2/3, -1/3, -1/3)$ and $(1/3, 1/3, -2/3)$. The vertices of $\mathfrak{c}_{y,x}$ are (y_1, y_2, y_3) and $(y_1 - 2/3, y_2 + 1/3, y_3 + 1/3)$ and $(y_1 - 1/3, y_2 - 1/3, y_3 + 2/3)$. A point z is in $\mathrm{para}(x,y)$ if and only if $z \in \mathfrak{a}^+$ and $y - z \in \mathfrak{a}^+$, i.e. $z_1 - z_2 \geq 0$ and $z_2 - z_3 \geq 0$ and $y_1 - y_2 \geq z_1 - z_2$ and $y_2 - y_3 \geq z_2 - z_3$.

We first show that if all vertices of $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$ are in a root, then z is also in that root. For roots of Type (1), if $\alpha_1 - \alpha_2 \geq a$ for $\alpha = 0$, then $a \leq 0$, in which case we clearly have $\alpha_1 - \alpha_2 \geq 0$ for z. If $\alpha_1 - \alpha_2 \leq a$ for $\alpha = y$, then $a \geq y_1 - y_2$, in which case z is also in the root. The other cases follow from a similar analysis. This show that $\mathrm{para}(x,y)$ is contained in the convex hull. On the other hand, it is clear that $\mathrm{para}(x,y)$ may be expressed as the intersection of finitely many roots, so the convex hull is contained in $\mathrm{para}(x,y)$. Hence they are equal. □

IX.2: Parallelograms in the building

Now suppose that x and y are two vertices in the building $\mathcal{B}$. Then there exists some apartment Σ containing both of them, and that apartment is itself a Coxeter complex, so it makes sense to talk about $\mathrm{para}(x,y)$ inside of Σ. It also makes sense to define combinatorial geodesics in $\mathcal{B}$ as any shortest path along edges connecting vertices (not necessarily only paths along edges which lie in some apartment). The following shows that $\mathrm{para}(x,y)$ is well-defined independently of the choice of Σ and that combinatorial geodesics in fact do lie in an apartment.

Proposition IX.8. *Suppose x and y are two vertices in $\mathcal{B}$. Let Σ be any apartment containing x and y. Then* $\mathrm{para}(x, y) \subset \Sigma$ *is contained in all apartments containing x and y, and its vertices are exactly the union of all vertices appearing along combinatorial geodesics connecting x and y in $\mathcal{B}$.*

Proof. Let γ be the unique Euclidean geodesic connecting x and y. Suppose first that γ purely consists of edges of $\mathcal{B}$. Then clearly the only combinatorial geodesic from x to y is exactly γ, and by [Bro89] p. 152, γ is independent of the choice of apartment.

Suppose instead that γ does not consist only of edges. Then $d_{A^+}(x, y) \in \mathfrak{a}^{++}$. Let $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$ be as in Proposition IX.7. Consider the union of all chambers appearing along minimal galleries from $\mathfrak{c}_{x,y}$ to $\mathfrak{c}_{y,x}$. Call this set B. Then B is the smallest convex subcomplex of $\mathcal{B}$ containing $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$, and it must lie in some apartment (by the building axioms some apartment contains $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$, and all apartments are convex). Since the intersection of any number of apartments is convex, the intersection of all apartments containing $\mathfrak{c}_{x,y}$ and $\mathfrak{c}_{y,x}$, or equivalently the intersection of all apartments containing x and y, must contain B. On the other hand B is exactly $\mathrm{para}(x, y)$ by Proposition IX.7.

What remains to be shown is that all combinatorial geodesics from x to y lie within $\mathrm{para}(x, y)$. Suppose P is some such combinatorial geodesic. Let Σ be some apartment containing x and y. Consider the retraction $\rho_{\mathfrak{c}_{x,y},\Sigma}$ from $\mathcal{B}$ onto Σ based at $\mathfrak{c}_{x,y}$ (see Appendix B.6.8). The image of P must again be a combinatorial path from x to y inside Σ of the same number of steps. However, all shortest paths from x to y in Σ are of length $r + s$ if $d_{A^+}(x, y) = (r, s)$ by Proposition IX.6. Hence the length of P must be equal to $r + s$.

Let z be the vertex along P right before the first time P leaves $\mathrm{para}(x, y)$. Let w be the next vertex after z along P. We must have that

$$d_c(y, w) = d_c(y, z) - 1. \tag{IX.2.1}$$

Suppose that $d_{A^+}(y, z) = (r', s')$. Then, using the retraction onto any sector based at y containing z (see Appendix B.6.9), any neighbor of z must have $d_{A^+}(y, w)$ belonging to one of the elements in (IX.1.2) or (IX.1.3) (replacing r with r' and s with s'). However, given (IX.2.1), we see that in fact $d_{A^+}(y, w) \in \{(r' - 1, s'), (r', s' - 1)\}$. On the other hand, by [CMo94] Lemma 2.1, there is at most one such w which is a neighbor of z for each of these possible values of $d_{A^+}(y, w)$. In all such cases, that unique such w lies is $\mathrm{para}(y, z)$ (as we can always find an element $w' \in \mathrm{para}(y, z)$ satisfying $d_{A^+}(y, w') = d_{A^+}(y, w)$). Therefore in fact $w \in \mathrm{para}(y, z) \subset \mathrm{para}(y, x) = \mathrm{para}(x, y)$ by Proposition IX.4. $\square$

IX.3: Classification of nearly opposite sectors

Suppose p is a vertex in $\mathcal{B}$. Recall that the link of p is itself a spherical building (see Appendix B.6.5). Hence given chambers $\mathfrak{c}_1$ and $\mathfrak{c}_2$ containing p, we may associate an element $d_W(\mathfrak{c}_1, \mathfrak{c}_2) \in \mathfrak{S}_3$ using the Coxeter group-valued metric. Up to simplicial automorphisms of the Coxeter complex of $\mathfrak{S}_3$, there are four relative positions that two chambers in a given apartment in the local spherical building can be in:

(1) $d_W(\mathfrak{c}_1, \mathfrak{c}_2) = 1$, i.e. $\mathfrak{c}_1 = \mathfrak{c}_2$.

(2) $d_W(\mathfrak{c}_1, \mathfrak{c}_2) = (12)$ or (23), i.e. $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are *adjacent*.

(3) $d_W(\mathfrak{c}_1, \mathfrak{c}_2) = (132)$ or (123), i.e. $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are *nearly opposite*.

(4) $d_W(\mathfrak{c}_1, \mathfrak{c}_2) = (13)$, i.e. $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are *opposite*.

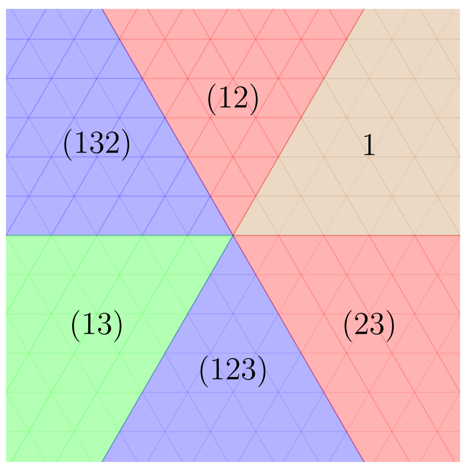


Figure 9: The red sectors are adjacent to the brown sector, the blue sectors are nearly opposite the brown sector, and the green sector is opposite the brown sector.

Now suppose x, y, p are vertices of $\mathcal{B}$. We say that $(x, y; p)$ is a *primitive triple* if $\operatorname{para}(x, p) \cap \operatorname{para}(y, p) = \{p\}$. If we have $d_{A^+}(x, p) \in \mathfrak{a}^{++}$ and $d_{A^+}(y, p) \in \mathfrak{a}^{++}$, then any sector $\mathcal{S}$ based at p containing x must have $\mathfrak{c}_1 := \mathfrak{c}_{p,x}$ as its germ, and similarly for $\mathfrak{c}_2 := \mathfrak{c}_{p,y}$. We clearly have $\mathfrak{c}_1 \cap \mathfrak{c}_2 \subset \operatorname{para}(x, p) \cap \operatorname{para}(y, p) = \{p\}$. Therefore $\mathfrak{c}_1$ and $\mathfrak{c}_2$ must be either opposite or nearly opposite (Cases (3) and (4) above).

If instead $(x, y; p)$ is a primitive triple but at least one of $d_{A^+}(p, x)$ and $d_{A^+}(p, y)$ is not in $\mathfrak{a}^{++}$, then we may find sectors $\mathcal{S}_1$ based at p containing x, and $\mathcal{S}_2$ based at p containing y such that their germs $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are either opposite or nearly opposite.

We now wish to classify primitive triples. There are a couple of facts about affine buildings that we shall utilize.

Proposition IX.9 ([Bro89] p. 169)**.** *Suppose $\mathfrak{H}$ is a half apartment in $\mathcal{B}$ with boundary wall $\partial\mathfrak{H}$. Suppose $\mathfrak{c}$ is a chamber in $\mathcal{B}$ with a panel lying in $\partial\mathfrak{H}$. Then there exists an apartment in $\mathcal{B}$ containing $\mathfrak{H}$ and $\mathfrak{c}$.*

Proposition IX.10 ([BS14], therein referred to as property (EC))**.** *Suppose Σ_1 and Σ_2 are two apartments that intersect in some half apartment $\mathfrak{H}$. Let $\mathfrak{H}_i$ be the other half apartment of Σ_i. Then $\mathfrak{H}_1$ and $\mathfrak{H}_2$ together form an apartment.*

Proposition IX.11 ([BS14], therein referred to as property (CO))**.** *Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are two sectors in $\mathcal{B}$ based at the same point p. Suppose the $\mathcal{S}_i$'s determine opposite chambers in the link of p (which is a spherical building). Then $\mathcal{S}_1$ and $\mathcal{S}_2$ are contained in a unique apartment.*

If $(x, y; p)$ is a primitive triple and $\mathfrak{c}_1 = \mathfrak{c}_{p,x}$ and $\mathfrak{c}_2 = \mathfrak{c}_{p,y}$ are opposite, then by Proposition IX.11 one can find an apartment containing all three points. It is then clear that in fact $p \in \mathrm{para}(x, y)$.

We thus are left with understanding the situation when $\mathfrak{c}_1$ and $\mathfrak{c}_2$ are nearly opposite. We extend the terminology and say that two sectors based at the same vertex are *nearly opposite* if their germs are.

IX.3.1: Strips and levels of sectors

We define a *wall* in the Coxeter complex $\mathcal{X}$ to be any union of edges which form an infinite line in the Euclidean sense. We define a *half-wall* to be any union of edges which form a ray in the Euclidean sense. We define a *combinatorial line segment* to be any union of edges which forms a line segment in the Euclidean sense. We define a *strip of width k* (with $k \geq 1$) in the Coxeter complex $\mathcal{X}$ to be the region bounded between two parallel walls whose combinatorial distance apart is k. We define a *half-strip* to be the intersection of a strip with any root in $\mathcal{X}$ which is not parallel to the walls defining the strip. Given a half-strip there is a unique chamber which has one of its edges along one of the defining walls, and the other edge along the boundary of the defining root; we call this the *germ of the half-strip*. See Figure 10.

Given a sector $\mathcal{S}$, and a choice of one of its bounding half-walls ℓ, we get an associated partition of $\mathcal{S}$ into *levels*, each of which is a half-strip of width one, which we shall index by

$k \in \mathbb{N}$ (starting from $k = 1$). Let $\mathrm{Level}(\mathcal{S}, k)$ denote the kth level, and let $\mathcal{S}^{(k)}$ denote the sector obtained by removing the first k levels (i.e. $\mathcal{S}^{(k)}$ is the union of levels $k+1, k+2, \dots$). See Figure 11.

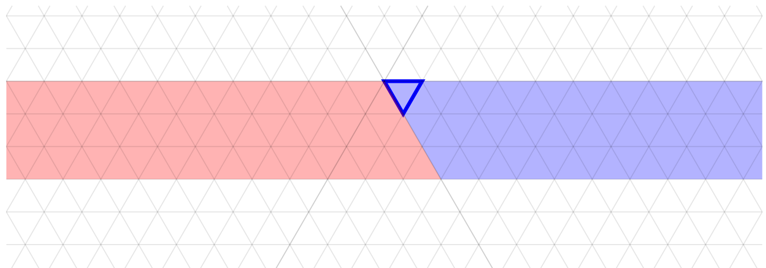


Figure 10: The red and blue regions together form a strip of width three. The red and blue regions are each half-strips. The solid blue chamber is the germ of the blue half-strip

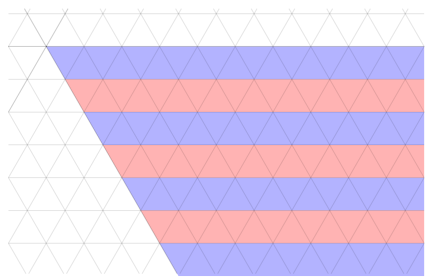


Figure 11: A sector may be partitions into levels, each of which is a half-strip of width one.

Suppose Σ is some apartment in $\mathcal{B}$. Suppose $\mathfrak{c} \subset \Sigma$ is a chamber and $\ell \subset \Sigma$ is a combinatorial line segment whose edges form part of a wall in Σ. Suppose the endpoints of ℓ are the vertices v_1 and v_2. Suppose the edge of ℓ containing v_1 is also one of the edges of $\mathfrak{c}$. Let w be the opposite vertex in $\mathfrak{c}$ to this edge. Then the convex hull of $\mathfrak{c}$ and ℓ is clearly $\mathrm{para}(w, v_2) \cup \mathfrak{c}$. See Figure 12.

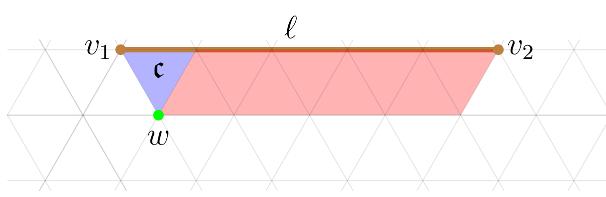


Figure 12: The convex hull of the blue chamber $\mathfrak{c}$ and the brown line ℓ is the union of $\mathfrak{c}$ and the red parallelogram $\mathrm{para}(w, v_2)$, where w is the opposite vertex of the blue chamber, and v_2 is the opposite vertex of the brown line.

Repeated application of this observation gives the following:

Lemma IX.12. *Suppose Σ is an apartment in $\mathcal{B}$, and $\mathfrak{c} \subset \Sigma$ is a chamber, and $\ell \subset \Sigma$ is a half-wall whose initial edge is one of the edges of $\mathfrak{c}$. Then the combinatorial convex hull of $\mathfrak{c}$ and ℓ is the half-strip in Σ of width one which has one of its bounding half-walls equal to ℓ and has germ equal to $\mathfrak{c}$. If instead ℓ is a wall, one of whose edges is one of the edges of $\mathfrak{c}$, then the convex hull of $\mathfrak{c}$ and ℓ is the strip of width one containing $\mathfrak{c}$ and which has ℓ as one of its bounding walls.*

Corollary IX.13. *Suppose $\mathfrak{H}$ is a half-apartment in $\mathcal{B}$. Suppose $\mathfrak{c}$ is a chamber not lying in $\mathfrak{H}$ but which shares a panel with $\partial\mathfrak{H}$. Then all apartments containing $\mathfrak{H}$ and $\mathfrak{c}$ (such apartments exist by Proposition IX.9) must contain the strip determined by $\partial\mathfrak{H}$ and $\mathfrak{c}$ as in Lemma IX.12.*

IX.3.2: Classification of nearly opposite sectors

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be nearly opposite sectors with common vertex p. Let $\mathcal{P}$ be the local spherical building obtained from the link of p. Let a_1, b_1 be the boundary half-walls of $\mathcal{S}_1$ and a_2, b_2 the boundary half-walls of $\mathcal{S}_2$ with a_1 and a_2 determining opposite vertices in the link of p. Let $\mathfrak{s}_1$ and $\mathfrak{s}_2$ be the germs of $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively.

We claim that the union of a_1 and a_2 forms an infinite (Euclidean) geodesic (and hence there exists some apartment containing $a_1 \cup a_2$ in which this set corresponds to a wall). Let $\bar{a}_1$ and $\bar{a}_2$ be the neighbors of p in a_1 and a_2 respectively. Notice that a_1 is a geodesic, a_2 is a geodesic, and the path $\bar{a}_1 \to p \to \bar{a}_2$ is a geodesic. Hence $a_1 \cup a_2$ is locally geodesic and therefore globally geodesic by the CAT(0) geometry.

In the sequel, any time we talk about levels of the sectors $\mathcal{S}_j$, it is with respect to the half-wall a_j.

Lemma IX.14. *Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are as above. Then there exists a half-apartment $\mathfrak{H}$ such that $\partial\mathfrak{H} = a_1 \cup a_2$, and $\mathfrak{H}$ does not otherwise intersect $\mathcal{S}_1$ or $\mathcal{S}_2$. Furthermore $\mathfrak{H}$ can be extended to a half-apartment $\mathfrak{H}_1$ by appending* $\mathrm{Level}(\mathcal{S}_1, 1)$, $\mathrm{Level}(\mathcal{S}_2, 1)$, *and a uniquely determined chamber $\mathfrak{t}$.*

Proof. Because the link of p is a spherical building, we know that there is some apartment Π in the local spherical building $\mathcal{P}$ containing $\mathfrak{s}_1$ and $\mathfrak{s}_2$. Let $\mathfrak{t}$ be the chamber connecting $\mathfrak{s}_1$ and $\mathfrak{s}_2$ along the minimal gallery in Π from $\mathfrak{s}_1$ to $\mathfrak{s}_2$. Then $\mathfrak{t}$ must have p as a vertex as well as the neighbors of p in b_1 and b_2. Since any three vertices determine at most one chamber, $\mathfrak{t}$ is well-defined independently of the choice of Π. Notice also that $\{\mathfrak{s}_1, \mathfrak{t}, \mathfrak{s}_2\}$ form

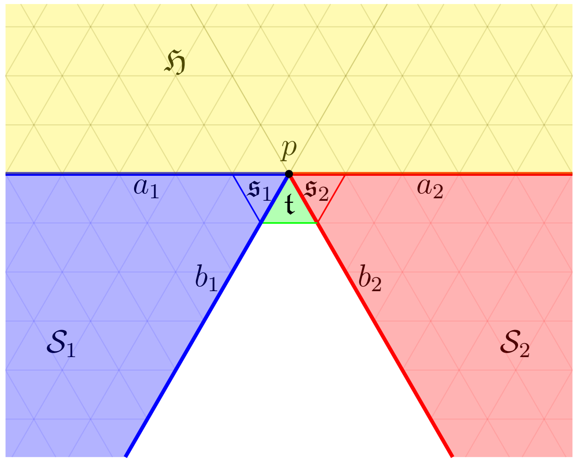


Figure 13: Given two nearly opposite sectors $\mathcal{S}_1$ (the blue sector) and $\mathcal{S}_2$ (the red sector), we may find a unique chamber $\mathfrak{t}$ (the green chamber) "connecting" the germs of $\mathcal{S}_1$ and $\mathcal{S}_2$, and a half-apartment $\mathfrak{H}$ (the yellow region) for which $\partial\mathfrak{H}$ is composed of the union of one of the bounding half-walls of $\mathcal{S}_1$ (denoted above as a_1) and one of the bounding half-walls of $\mathcal{S}_2$ (denoted above as a_2). Furthermore, the yellow region together with the green chamber and the first levels of the red and blue sectors also forms a half-apartment.

a half-apartment in $\mathcal{P}$, and therefore by [Ron89] Lemma 6.3, any chamber adjacent to $\mathfrak{s}_2$ which has $\mathfrak{s}_2 \cap a_2$ as one of its panels must be opposite to $\mathfrak{s}_1$.

Let Σ be any apartment in $\mathcal{B}$ containing $\mathcal{S}_2$. Let $\mathcal{R}$ be the sector based at p in Σ which shares the half-wall a_2 with $\mathcal{S}_2$. Let $\mathfrak{r}$ be the germ of $\mathcal{R}$. Because $\mathfrak{r}$ shares a panel with $\mathfrak{s}_2$, we conclude that it is opposite to $\mathfrak{s}_1$ in $\mathcal{P}$. Therefore by Proposition IX.11, there exists a unique apartment Σ' containing $\mathcal{S}_1$ and $\mathcal{R}$. Furthermore the half-apartment $\mathfrak{H}$ in Σ' containing both $a_1 \cup a_2$ and $\mathcal{R}$ clearly does not intersect $\mathcal{S}_1$ other than along a_1. We claim that it also cannot intersect $\mathcal{S}_2$ other than along a_2. This shall follow from the following in which we show that we may append a level to $\mathfrak{H}$ to obtain a bigger half-apartment as in the statement of the lemma.

Consider the convex hull of $\mathfrak{s}_1$ and $\partial\mathfrak{H}$. By Proposition IX.9, this adds on a strip to $\mathfrak{H}$. This strip clearly contains $\mathrm{Level}(\mathcal{S}_1, 1)$ by Corollary IX.13. Furthermore the convex hull of $\mathfrak{s}_1$ and $\mathfrak{r}$ must contain $\mathfrak{t}$ and $\mathfrak{s}_2$ (because they are opposite, they determine a unique apartment in $\mathcal{P}$ by [Ron89] Chapter 6.1). Since the convex hull of $\mathfrak{s}_2$ and a_2 is $\mathrm{Level}(\mathcal{S}_2, 1)$, we conclude the proof of the lemma. See Figure 13. $\square$

We define an *equilateral triangle* in $\mathcal{B}$ to be the intersection of a sector $\mathcal{S}$ with any half-apartment $\mathfrak{J}$ which has an extension to an apartment containing $\mathcal{S}$ and such that $\partial\mathfrak{J}$ is transverse (i.e. not parallel) to the half-walls of $\mathcal{S}$. Such a region lies in an apartment, and in that apartment it is a Euclidean equilateral triangle. The size of the equilateral triangle is the length of any one of its sides.

Lemma IX.15. *Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are as before. Let $\mathcal{S}_1^{(n)}$ and $\mathcal{S}_2^{(n)}$ be the sectors obtained by removing the first n levels. Let $p_1^{(n)}$ and $p_2^{(n)}$ be the base vertices of these sectors. Let $a_1^{(n)}$ and $a_2^{(n)}$ be the bouding half-walls of these sectors parallel to a_1 and a_2. Then either $\mathcal{S}_1$ and $\mathcal{S}_2$ are contained in an apartment, or there exists a k such that all of the following hold:*

(1) $\mathcal{S}_1^{(k)}$ and $\mathcal{S}_2^{(k)}$ are contained in an apartment Σ,

(2) $p_1^{(k)}$ and $p_2^{(k)}$ are the endpoints of a line segment of length k in Σ,

(3) $\ell \cup a_1^{(k)} \cup a_2^{(k)}$ is a wall in Σ,

(4) given $x \in \mathcal{S}_1^{(k)}$ and $y \in \mathcal{S}_2^{(k)}$, we have $\ell \subset \mathrm{para}(x, y) \subset \Sigma$,

(5) and $p_1^{(k)}$, $p_2^{(k)}$, and p form the vertices of an equilateral triangle of size k and one of whose sides is ℓ.

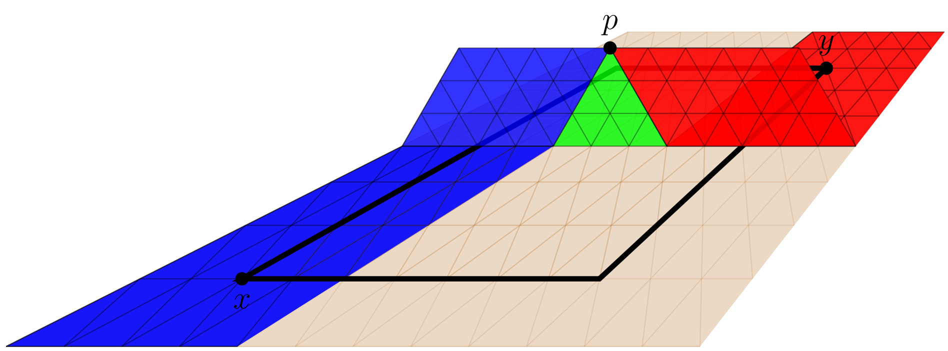


Figure 14: Given vertices x, y, and p, we can consider $\mathrm{para}(x, y)$, which is the black parallelogram in the figure. If $\mathrm{para}(x, p)$ and $\mathrm{para}(y, p)$ only intersect at p, and the associated chambers $\mathfrak{c}_{p,x}$ and $\mathfrak{c}_{p,y}$ are nearly opposite, then we may find nearly opposite sectors at p, one of which contains x (the blue sector) and the other of which contains y (the red sector). Certain half-strips of these sectors may be combined with an equilateral triangle (the green triangle) to form a strip. The remaining part of these sectors may then we placed in some apartment (containing x and y) which is the brown apartment in the figure. In this apartment, the remaining parts of these sectors are oriented opposite.

Proof. The strategy is to attempt to add levels to the half-apartment $\mathfrak{H}$ from Lemma IX.14. In fact Lemma IX.14 already tells us that we can add the first level. Let $\mathfrak{H}_k$ denote the half

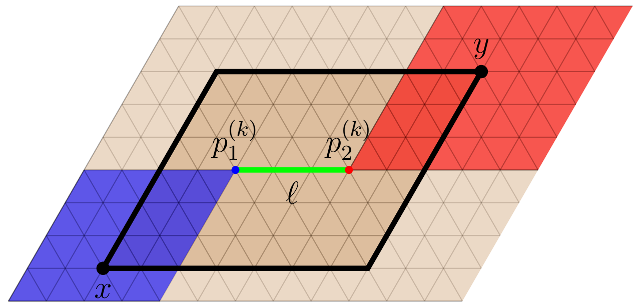


Figure 15: This shows essentially the same picture as Figure 14 from the "perspective" of the brown apartment. The blue sector is exactly $\mathcal{S}_1^{(k)}$ and the red sector is exactly $\mathcal{S}_2^{(k)}$ as in Lemma IX.15 (here $k = 3$).

apartment obtained by appending the first k levels of $\mathcal{S}_1$ and $\mathcal{S}_2$ to $\mathfrak{H}$, assuming we are able to do so. We thus know that $\mathfrak{H}_1$ exists.

If we are able to keep adding levels indefinitely, then, by taking the union of these levels together with $\mathfrak{H}$, we obtain a apartment containing $\mathcal{S}_1$ and $\mathcal{S}_2$ (the union is clearly an apartment in the complete apartment system; see Appendix B.4.2).

Suppose instead that n is such that we have successfully added n levels (and hence constructed $\mathfrak{H}_n$) but are not able to add the $(n+1)$st level. By Proposition IX.11, we may find apartments Σ_1 and Σ_2 such that Σ_j contains $\mathfrak{H}_n$ and $\mathcal{S}_j^{(n)}$. We clearly have $\Sigma_1 \cap \Sigma_2 \supset \mathfrak{H}_n$. Suppose some other chamber $\mathfrak{c}$ were in the intersection. Let $\mathfrak{c}'$ be the last chamber not in $\mathfrak{H}_n$ along any minimal gallery from $\mathfrak{c}$ to any chamber in $\mathfrak{H}_n$. This gallery must lie in the intersection $\Sigma_1 \cap \Sigma_2$ as the intersection is combinatorially convex. Therefore $\mathfrak{c}'$ is a chamber in the intersection which is not contained in $\mathfrak{H}_n$ but shares a panel with $\partial\mathfrak{H}_n$. By Corollary IX.13, any apartment containing $\mathfrak{H}_n$ and $\mathfrak{c}'$ must also contain the strip parallel to $\partial\mathfrak{H}_n$ containing $\mathfrak{c}'$ and hence this strip is in $\Sigma_1 \cap \Sigma_2$. However, this strip clearly contains $\mathrm{Level}(\mathcal{S}_j, n+1)$ for $j = 1, 2$, so we would be able to form $\mathfrak{H}_{n+1}$, which is a contradiction. Therefore, $\Sigma_1 \cap \Sigma_2 = \mathfrak{H}_n$.

Let $\mathcal{J}_1$ and $\mathcal{J}_2$ be the other half apartments of Σ_1 and Σ_2. By Proposition IX.10, we may form an apartment $\Sigma = \mathcal{J}_1 \cup \mathcal{J}_2$. It is clear from the construction that Σ satisfies Properties (1), (2), (3), and (5) (with $n = k$). Property (4) follows from the observation that in Σ, the sectors $\mathcal{S}_1^{(n)}$ and $\mathcal{S}_2^{(n)}$ have opposite orientations. $\square$

IX.4: Classification of primitive triples

IX.4.1: Existence of confluence points

Lemma IX.16. *Suppose x, y, and p are vertices of $\mathcal{B}$. Let $\mathcal{D} = \mathrm{para}(x,p) \cap \mathrm{para}(y,p)$. Then there exist points $z \in \mathcal{D}$ which simultaneously minimize $d_c(x,\cdot)$ and $d_c(y,\cdot)$ over all points in $\mathcal{D}$. Furthermore, for any such z, $(x, y; z)$ forms a primitive triple.*

Proof. Let $z \in \mathcal{D}$. By Proposition IX.8, this implies that there exist combinatorial geodesics from x to p and from y to p which both pass through z, and hence,

$$\begin{aligned} d_c(x,p) &= d_c(x,z) + d_c(z,p), \\ d_c(y,p) &= d_c(y,z) + d_c(z,p). \end{aligned}$$

Now suppose $z \in \mathcal{D}$ is as close to x with respect to $d_c(\cdot,\cdot)$ as any other point in $\mathcal{D}$. Suppose $w \in \mathcal{D}$ is some other point. We wish to show that $d_c(y,z) \leq d_c(y,w)$. By (IX.1.1) we have:

$$\begin{aligned} d_c(x,z) + d_c(z,p) &= d_c(x,w) + d_c(w,p), \\ d_c(y,z) + d_c(z,p) &= d_c(y,w) + d_c(w,p), \\ d_c(x,z) - d_c(y,z) &= d_c(x,w) - d_c(y,w), \\ d_c(x,z) - d_c(x,w) &= d_c(y,z) - d_c(y,w). \end{aligned}$$

By assumption $d_c(x,z) - d_c(x,w) \geq 0$. Therefore, the same holds for $d_c(y,z) - d_c(y,w)$, and therefore z is as close to x and y as any other point in $\mathcal{D}$.

We now show that $(x, y; z)$ is a primitive triple. Suppose $w \in \mathrm{para}(x,z) \cap \mathrm{para}(y,z)$. Then, on the one hand $w \in \mathcal{D}$ by Proposition IX.4, and on the other hand w lies along some combinatorial geodesic from x to z and hence w could not be further away from z than x is. But z is as close to x as any other point in $\mathcal{D}$. Therefore $w = z$. □

Given x, y, and p, we call any point $z \in \mathrm{para}(x,p) \cap \mathrm{para}(y,p)$ satisfying the conditions in Lemma IX.16 a *confluence point* of $(x, y; p)$.

Remark IX.17. In general $(x, y; p)$ may have several confluence points. For example suppose $d_{A^+}(x,y) = (1,1)$. Consider the unique edge $e \in \mathrm{para}(x,y)$ which passes through the interior of $\mathrm{para}(x,y)$. Let p be any point such that $e \cup p$ forms a chamber $\mathfrak{c} \subset \mathcal{B}$. Then

$\text{para}(x,p) \cap \text{para}(y,p) = \mathfrak{c}$. Therefore, either endpoint of e may be considered a confluence point of $(x,y;p)$.

IX.4.2: Directions of the bounding line segments of a parallelogram

Suppose $d_{A^+}(x,y) = (r,s)$. The parallelogram $\text{para}(x,y)$ viewed as a Euclidean parallelogram has as its corners x, y as well as two other vertices w, z. One of these, say w, satisfies $d_{A^+}(x,w) = (r,0)$, and the other, say z, satisfies $d_{A^+}(x,w) = (0,s)$. In such a case we say that the line from x to w is *in the* $(1,0)$*-direction with respect to* x, and the edge from x to z is *in the* $(0,1)$*-direction with respect to* x. Combinatorial line segments inside of $\text{para}(x,y)$ which are parallel to, e.g., the combinatorial line segment from x to w are also said to be in the $(1,0)$-direction with respect to x.

Given a combinatorial line segment ℓ we may consider its midpoint z. Then z is not necessarily in Λ, but it is inside of $\Lambda/2$, and we may extend $d_{A^+}(\cdot,\cdot)$ to also make sense for points in $\Lambda/2$.

IX.4.3: Branch lines and statement of the classification

The following is a summary of the content of Chapter IX.3 applied towards the classification of primitive triples (together with some straightforward calculations which have been suppressed).

Lemma IX.18. *Suppose $(x,y;p)$ is a primitive triple. Let $\mathcal{S}_x$ (and $\mathcal{S}_y$, resp.) be any sector based at p containing x (containing y, resp.).*

(1) If $\mathcal{S}_x$ and $\mathcal{S}_y$ are opposite, then $p \in \text{para}(x,y)$. We define $\ell = p$ to be the branch line of the primitive triple (we may consider a point to be a line of length 0).

(2) Suppose $\mathcal{S}_x$ and $\mathcal{S}_y$ are nearly opposite.

(a) Suppose there exists an apartment Σ containing $\mathcal{S}_x$ and $\mathcal{S}_y$. Let $\mathcal{T}$ be the sector joining these two sectors in Σ. Let ℓ be all points in $\mathcal{T} \cap \text{para}(x,y)$ minimizing $d_c(p,\cdot)$. Then ℓ consists of a line segment; we define ℓ to be the branch line of the primitive triple.

(b) Suppose there does not exist an apartment containing $\mathcal{S}_x$ and $\mathcal{S}_y$. Let ℓ be as in Lemma IX.15. Then ℓ is defined to be the branch line of the primitive triple.

Now suppose ℓ has midpoint z and ℓ has length k.

(1) If ℓ is in the $(1,0)$-direction, then

$$d_{A^+}(x,p) = d_{A^+}(x,z) - (k/2,0) + (0,k)$$
$$d_{A^+}(y,p) = d_{A^+}(y,z) - (0,k/2) + (k,0).$$

(2) If ℓ is in the $(0,1)$-direction, then

$$d_{A^+}(x,p) = d_{A^+}(x,z) - (0,k/2) + (k,0)$$
$$d_{A^+}(y,p) = d_{A^+}(y,z) - (k/2,0) + (0,k).$$

IX.5: Proof of Proposition II.6 and Corollary II.7

Recall the definition of the polytope H from Appendix C.6.1. We wish to prove Proposition II.6, namely that given $x, y \in G/K$, we have $xE_m \cap xE_m \neq \emptyset$ if and only if $d_{A^+}(x,y) \in H_m^\Lambda$.

Proof of Proposition II.6 (When Polytopal Ball Intersect). Let x and y be vertices in the building. We first claim the following: if $xE_m \cap yE_m \neq \emptyset$, then there exists a p in the intersection such that $(x,y;p)$ forms a primitive triple. To see this: first suppose $w \in xE_m \cap yE_m$. Suppose $(x,y;w)$ is not a primitive triple. Then, by Lemma IX.16, there exists a $p \in \mathrm{para}(x,w) \cap \mathrm{para}(y,w)$ such that $d_{A^+}(x,p) \preceq d_{A^+}(x,w)$, and $d_{A^+}(y,p) \preceq d_{A^+}(y,w)$, and $(x,y;p)$ is a primitive triple; we clearly have $p \in xE_m \cap yE_m$.

We now use the classification of primitive triples to show that there is a $w \in xE_m \cap yE_m$ such that $(x,y;w)$ is an additive triple. Let ℓ be the branch line of the primitive triple $(x,y;p)$, and suppose ℓ has length k (see Lemma IX.18). Suppose z is the midpoint of ℓ. Suppose ℓ is in the $(1,0)$-direction. Suppose $d_{A^+}(x,z) = (a,b)$. Then by Table 1:

$$d_{A^+}(x,p) = d_{A^+}(x,z) - (k/2,0) + (0,k),$$
$$|d_{A^+}(x,p)|_P = (a - k/2) + 2(b+k) = a + 2b + k/2 = |d_{A^+}(x,z)|_P + k/2$$

Suppose $d_{A^+}(y,z) = (c,d)$. Then:

$$d_{A^+}(y,p) = d_{A^+}(y,z) - (0,k/2) + (k,0),$$
$$|d_{A^+}(y,p)|_P = (c+k) + 2(d + k/2) = c + 2d = |d_{A^+}(y,z)|_P$$

Hence we conclude that if $p \in xE_m \cap yE_m$, namely $|d_{A^+}(x,p)|_P \leq m$ and $|d_{A^+}(y,p)|_P \leq m$, then z is also in $xE_m \cap yE_m$. A similar calculation may be performed in case ℓ is in the $(0,1)$-direction (the roles of x and y will be reversed). We therefore conclude that if xE_m and yE_m intersect, then the intersection contains an additive triple, namely $(x, y; z)$.

Without loss of generality, we may take $x = 1K \in \Lambda$ and assume that $y \in \mathfrak{a}^+$. By above discussion implies that $xE_m \cap yE_m \neq \emptyset$ if and only if $((xE_m) \cap \mathfrak{a}) \cap ((yE_m) \cap \mathfrak{a}) \neq \emptyset$, i.e. it suffices to determine when translated copies of P_m intersect P_m. By Figure 16 we conclude that when $m = 1$, this occurs if and only if $y \in H$. The case of general m follows immediately by simply scaling up Figure 16 by a factor of m. $\square$

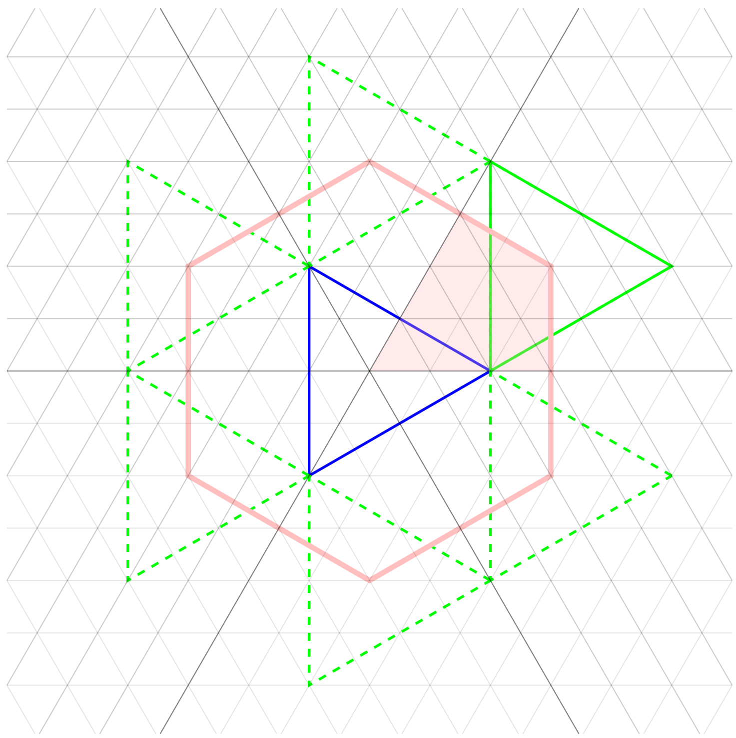


Figure 16: The blue triangle represents the $\mathfrak{S}_3$-orbit of the polytope P in $\mathfrak{a}$. The green triangles represent translated copies of this polytope. We wish to understand where the centers of the green triangles may be so that the intersection with the blue triangle is non-empty. This is exactly the pink hexagon. The restriction of the pink hexagon to $\mathfrak{a}^+$ tells us what the possible values of $d_{A^+}(x, y)$ are (with x equal to the center of the blue triangle, and y equal to the center of one of the green triangles) such that the triangles intersect; this is exactly the polytope H which is represented by the pink shaded region.

Remark IX.19. At its heart, the proof of Proposition II.6 relies on the convexity of the polytope obtained by taking the $\mathfrak{S}_3$ orbit of P in $\mathfrak{a}$. See Figure 25. Though we do not wish to discuss it at length here, using the theory of Hecke paths (see [KM08]), one may show that if one has a polytope $Q \subset \mathfrak{a}^+$ whose $\mathfrak{S}_3$ orbit is convex, then for vertex x in $\mathcal{B}$, the collection of values that $d_{A^+}(x,y)$ obtains for y such that the Q-shaped ball centered at x and the Q-shaped ball centered at y intersect is itself a convex polytope.

Proof of Corollary II.7 (Kernel Function is Supported Near the Diagonal). Notice that H is completely contained in the ball of radius $2\sqrt{3}$ centered at $1K$ (recall that we are using the normalization of the metric such that if $d_{A^+}(x,y) = (1,0)$ or $(0,1)$, then $d(x,y) = 1$). Therefore $\mathcal{L}_M(g,h) \neq 0$ implies that $gE_m \cap hE_m \neq 0$ for some $m \leq M$, which implies that $d_{A^+}(gK,hK) \in H_M$, which implies that $d(gK,hK) \leq 2\sqrt{3}M$. $\square$

CHAPTER X
Geometric Bound on the Size of the Intersection of Polytopal Balls

X.1: The number of equilateral triangles corresponding to a given combinatorial line segment

Proposition X.1. *The number of chambers in $\mathcal{B}$ containing a given edge is exactly $q+1$.*

Proof. This follows from counting the number of ways to extend a line in $\mathbb{F}_q^3$ to a plane. □

Given a combinatorial line segment (which is composed of finitely many edges), we call any edge containing one of the endpoints of the line segment a *boundary edge*.

Given an equilateral triangle $\mathcal{T}$ in $\mathcal{B}$ and a choice of one of its bounding combinatorial line segments, we may partition $\mathcal{T}$ into *levels* similarly to the way in which we partitioned sectors into levels. Each level is a trapezoid of width one as in Figure 12.

Lemma X.2. *Suppose ℓ is a combinatorial line segment in $\mathcal{B}$ of length k. The number of equilateral triangles which have ℓ as one of their sides is equal to $(q+1)q^{k-1}$.*

Proof. Let e_1 be one of the boundary edges of ℓ. By Proposition X.1, there are $q+1$ chambers containing e_1. Let $\mathfrak{c}_1$ be one such chamber. Then by Lemma IX.12, the convex hull of ℓ and $\mathfrak{c}_1$ is a trapezoid as in Figure 12; this shape exactly corresponds to the first level of an equilateral triangle containing e_1. We now wish to add a second level. Let e_2 be one of the boundary edges of the combinatorial line bounding this trapezoid which is parallel and opposite to ℓ. Again, e_2 is contained in $q+1$ chambers, one of which already occurs in the trapezoid we already constructed. For each of the q other choices, we may subsequently add another level to our trapezoid. We may continue this process until we end up with an equilateral triangle; at each step after the first there are q choices for how to add the next level. □

X.2: Coordinatizing the relative position of triples

Suppose $(x, y; z)$ is a triple of vertices in $\mathcal{B}$. Suppose $d_{A^+}(x, y) = (r, s)$. Let p be any confluence point of $(x, y; z)$. Let ℓ be the branch line of $(x, y; p)$. Suppose ℓ has length k and is in the α-direction with respect to x (i.e. $\alpha \in \{(1, 0), (0, 1)\}$). Let w be the point on ℓ closest to x (with respect to $d_c(\cdot, \cdot)$, or equivalently with respect to $d(\cdot, \cdot)$). Suppose $d_{A^+}(x, w) = (a_1, a_2)$ and $d_{A^+}(p, z) = (b_1, b_2)$. Then we assign to $(x, y; p; z)$ the "coordinates" $(r, s; a_1, a_2, k, \alpha; b_1, b_2)$. Note, that by Remark IX.17, a given point may get assigned multiple "coordinates" because there may be multiple choices of confluence points. Our next goal is to bound, for a given x, y with $d_{A^+}(x, y) = (r, s)$ and a give choice of $(r, s; a_1, a_2, k, \alpha; b_1, b_2)$, the number of z such that there exists a confluence point p for $(x, y; z)$ such that $(x, y; p; z)$ gets assigned the given coordinates.

Lemma X.3. *Suppose x, y are fixed with $d_{A^+}(x, y) = (r, s)$. Then*

$$\left\{ \begin{array}{c} \#z \text{ s.t. } \exists \text{ confluence point } p \text{ of } (x, y; z) \text{ s.t.} \\ \text{coordinates of } (x, y; p; z) \text{ are } (r, s; a_1, a_2, k, \alpha; b_1, b_2) \end{array} \right\} \leq \frac{2}{\nu_3(q^{-1})} (q^2)^{\frac{k}{2} + b_1 + b_2}. \quad \text{(X.2.1)}$$

Proof. The choice of coordinates uniquely determines the branch line ℓ. Let $\mathfrak{T}$ denote the collection of all equilateral triangles which have one of their bounding line segments equal to ℓ; the cardinality of this set is given by Lemma X.2. Given $\mathcal{T} \in \mathfrak{T}$, let $a(\mathcal{T})$ denote the vertex of $\mathcal{T}$ opposite to ℓ. Let X denote the set:

$$X := \{(\mathcal{T}, w) : \mathcal{T} \in \mathfrak{T} \text{ and } w \in G/K \text{ and } d_{A^+}(a(\mathcal{T}), w) = (b_1, b_2)\}.$$

Given $(x, y; p; z)$ on the left hand side of (X.2.1), we can associate to it an element of X, namely the equilateral triangle $\mathcal{T}'$ associated to the primitive triple $(x, y; p)$ as in Lemma IX.18 (clearly then $a(\mathcal{T}') = p$)) and the point z itself. This map is clearly an injection. On the other hand, for a fixed $\mathcal{T} \in \mathfrak{T}$, the number of points in X whose first entry is $\mathcal{T}$ is exactly $N_{(b_1, b_2)}$ which is given by (A.3.2) (note that that formula is given in partition coordinates, but here we are using cone coordinates). Therefore the cardinality of X is exactly given by the right hand side of (X.2.1). $\square$

Suppose $(x, y; p; z)$ has coordinates $(r, s; a_1, a_2, k, \alpha; b_1, b_2)$. Then if $\alpha = (1, 0)$, we have

$$\begin{aligned} d_{A^+}(x, p) &= (a_1, a_2 + k) \\ d_{A^+}(y, p) &= (s - a_2 + k, r - a_1 - k). \end{aligned}$$

If $\alpha = (0,1)$, we have

$$\begin{aligned} d_{A^+}(x,p) &= (a_1 + k, a_2) \\ d_{A^+}(y,p) &= (s - a_2 - k, r - a_1 + k). \end{aligned}$$

This follows from Lemma IX.18.

X.3: The polytope parametrizing allowable coordinates of triples of points

X.3.1: The defining inequalities of the polytope

Proposition X.4. *Suppose $(x, y; p; z)$ has coordinates $(r, s; a_1, a_2, k, \alpha; b_1, b_2)$ with $z \in xE_m \cap yE_m$. Then the following inequalities are satisfied:*

$$\begin{aligned} r + 2s &\leq 6m & & d_{A^+}(x,y) \in H_m \\ 2r + s &\leq 6m & & d_{A^+}(x,y) \in H_m \\ a_1 &\leq r & & \ell \subset \mathrm{para}(x,y) \\ a_2 &\leq s & & \ell \subset \mathrm{para}(x,y) \\ r, s, a_1, a_2, k, b_1, b_2, m &\geq 0. \end{aligned} \tag{X.3.1}$$

(1) Suppose $\alpha = (1,0)$. Then the following inequalities are also satisfied:

$$\begin{aligned} (a_1 + b_1) + 2(a_2 + k + b_2) &\leq 2m & & d_{A^+}(x,z) \in P_m \\ (s - a_2 + k + b_1) + 2(r - a_1 - k + b_2) &\leq 2m & & d_{A^+}(y,z) \in P_m \\ a_1 + k &\leq r & & \ell \subset \mathrm{para}(x,y). \end{aligned} \tag{X.3.2}$$

(2) Suppose $\alpha = (0,1)$. Then the following inequalities are also satisfied:

$$\begin{aligned} (a_1 + k + b_1) + 2(a_2 + b_2) &\leq 2m & & d_{A^+}(x,z) \in P_m \\ (s - a_2 - k + b_1) + 2(r - a_1 + k + b_2) &\leq 2m & & d_{A^+}(y,z) \in P_m \\ a_2 + k &\leq s & & \ell \subset \mathrm{para}(x,y). \end{aligned} \tag{X.3.3}$$

X.3.2: The dominating term of the sum over lattice points in the polytope

Lemma X.5. *All* $z \in xE_m \cap yE_m$ *with* $d_{A^+}(x, y) = (r, s)$ *have their coordinates* $(x, y; p; z)$ *(with any choice of confluence point* p*) satisfying:*

$$\frac{k}{2} + b_1 + b_2 \leq 2m - \frac{r}{2} - \frac{s}{2}. \tag{X.3.4}$$

Furthermore, when m, r, s *are fixed, there is at most one possible set of coordinates for such* $(x, y; p; z)$ *such that* (X.3.4) *becomes an equality.*

Proof. Suppose $(x, y; z; p)$ is as in the statement of the lemma. Depending on α, the coordinates of z must satisfy all of (X.3.1) plus either all of (X.3.2) or all of (X.3.3). In all cases we refer to these inequalities collectively as the coordinates inequalities. Suppose the coordinates of z also satisfy the inequality:

$$\frac{k}{2} + b_1 + b_2 \geq 2m - \frac{r}{2} - \frac{s}{2}. \tag{X.3.5}$$

The set of points in $\mathbb{R}^8$ satisfying the coordinates inequalities as well as (X.3.5) forms a convex polyhedron. Using the Polyhedra functionality in Sage, one realizes that, regardless of what α is, this polytope is the conical span of the following vectors in coordinates $(m, r, s, a_1, a_2, k, b_1, b_2)$:

$$u_1 = (1, 2, 2, 2, 0, 0, 0, 0)$$
$$u_2 = (1, 0, 0, 0, 0, 0, 2, 0).$$

Any vector in this conical span results in (X.3.5) becoming an equality. Hence no point in this polytope ever violates (X.3.4). Furthermore, when m, r, s are fixed, there is at most one point in the conical span whose first three coordinates are (m, r, s) because the first three coordinates of u_1, namely $(1, 2, 2)$, and the first three coordinates of u_2, namely $(1, 0, 0)$, are linearly independent in $\mathbb{R}^3$. In fact for such a point we must have

$$r = s$$
$$a_1 = r$$
$$a_2 = 0$$
$$k = 0$$
$$b_1 = 2m - r$$

$$b_2 = 0.$$

$\square$

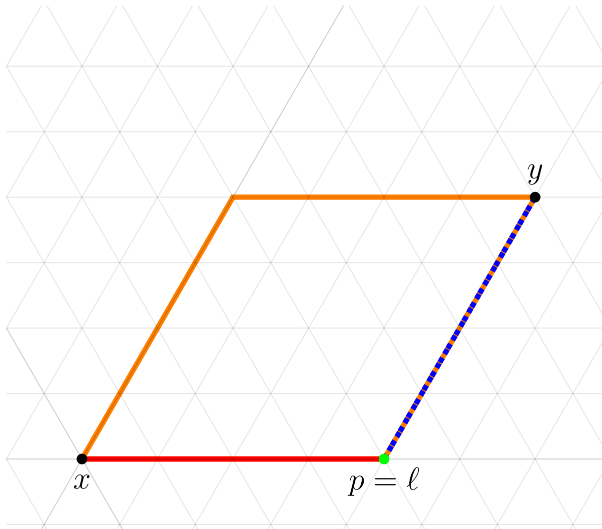


Figure 17: The only way that $(x, y; p; z)$ can have coordinates which results in (X.3.4) being an equality is when $r = s$ and $a_1 = r$ and $k = 0$. This means that the sides of $\mathrm{para}(x, y)$ are both the same, and that the branch line ℓ is a single point and is equal to p, and that this point is at the corner of $\mathrm{para}(x, y)$ which is obtained by moving from x in the $(1, 0)$-direction. We then see that $d_{A^+}(x, p) = d_{A^+}(y, p) = (r, 0)$ (in this case $r = s = 4$).

X.4: Proof of Proposition II.15

Proof of Proposition II.15 (Upper Bound on $\mathrm{card}(E_m^\lambda)$*).* Recall that $\tilde{E}_m^\lambda$ is the pullback to G of $E_m^\lambda = E_m \cap \varpi^\lambda E_m \subset G/K$ with $\lambda \in A^+$. Clearly $\mathrm{vol}(\tilde{E}_m^\lambda) = \mathrm{card}(E_m^\lambda)$. Suppose $\lambda = (r, s)$ in cone coordinates. We can associate to each $z \in E_m^\lambda$ the coordinates $(m; r, s; a_1, a_2, k, \alpha; b_1, b_2)$ (after choosing some confluence point p). Let $(\beta, \cdot)$ be the functional defined by

$$(\beta, (a_1, a_2, k, b_1, b_2)) = \frac{k}{2} + b_1 + b_2. \tag{X.4.1}$$

By Lemma X.3, the number of z which map to a given set of coordinates is at most

$$(q^2)^{\frac{k}{2} + b_1 + b_2} = (q^2)^{(\beta, (a_1, a_2, k, b_1, b_2))}. \tag{X.4.2}$$

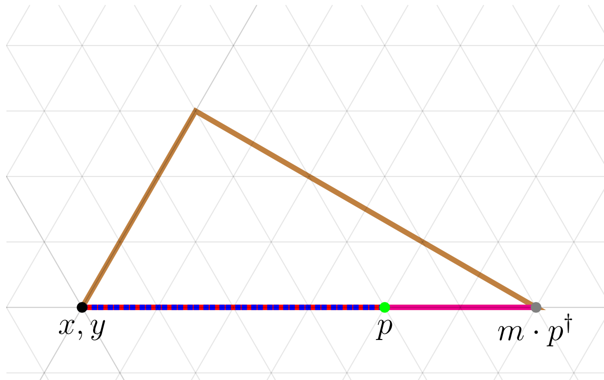


Figure 18: This is a continuation of Figure 17. Here the line from x to p (the red line) and from y to p (the dotted blue line) are moved into $\mathfrak{a}^+$ by applying the appropriate element of $\mathfrak{S}_3$. We then add the same vector, namely (b_1, b_2) to both of these line segments. In this case the line from p to $m \cdot p^\dagger$ (represented by the pink line) is the unique line we can add on that will maximize the dot product with δ and also keep us inside of P_m (represented by the brown polygon).

Let $P_\alpha(m, r, s)$ be the polytope defined by the inequalities in (X.3.1) and either (X.3.2) (if $\alpha = (1, 0)$) or (X.3.3) (if $\alpha = (0, 1)$). We can then bound $\mathrm{card}(E^\lambda_m)$ by summing (X.4.2) over all integer points (in coordinates (a_1, a_2, k, b_1, b_2)) in the polytopes $\mathcal{P}_{(1,0)}(m, r, s)$ and $\mathcal{P}_{(0,1)}(m, r, s)$.

For the remainder of the proof, we assume that we have fixed an α (the subsequent discussion does not depend on α). Notice that all of the defining inequalities of $P_\alpha(m, r, s)$ are of the form

$$f_i(a_1, a_2, k, b_1, b_2) \leq g_i(m, r, s),$$

where f_i and g_i are linear functionals. We have that (m, r, s) parametrizes some vector space which we call $V_{(m,r,s)}$.

We also have each of our polytopes $\mathcal{P}_\alpha(m, r, s)$ living inside of some five-dimensional space which we call W and which is naturally coordinatized via (a_1, a_2, k, b_1, b_2). We can partition up $V_{(m,r,s)}$ according to the type of the underlying polytope in W together with the information about which vertices maximize $(\beta, \cdot)$. We call each such component a *β-region*. Clearly there are only finitely many such regions.

Because all polytopes associated to a given β-region have the same type, it makes sense to discuss the "same" vertex for different polytopes in a given β-region. Some β-regions are such that their associated polytopes have a vertex v^* which, at least for some polytopes in the

β-region, satisfies $(\beta, v^*) = 2m - r/2 - s/2$; we shall refer to such β-regions as *extremal*. By Lemma X.5, in such cases v^* is the unique vertex satisfying this equation and is the unique vertex maximizing $(\beta, \cdot)$. In general, at the vertex w maximizing $(\beta, \cdot)$ in any polytope, we have that β is in the polar cone of the cone generated at w. Hence β is in the polar cone of $\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(v^*)$ (which we view as a cone based at the origin, not based at v^*). Furthermore, because v^* is the unique vertex maximizing $(\beta, \cdot)$, we must have β in the interior of the polar cone.

Recall that the structure of $\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(v^*)$ only depends on the type of the underlying polytope (see Appendix C.2.1). Hence this cone does not depend on (m, r, s) (as long as we stay in the same β-region). Because β is in the interior of the polar cone, we have that the sum of (X.4.2) over the lattice points in $\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(v^*)$ has a finite value; in fact this is exactly what is referred to as

$$\sigma(\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(v^*); 2\beta) \tag{X.4.3}$$

in Appendix C.3.2. If we translate this cone to be based at v^*, then the sum of (X.4.2) over the lattice points in this cone gives an upper bound on the sum of (X.4.2) over the lattice points in $\mathcal{P}_\alpha(m, r, s)$. On the other hand, this quantity is at most $(q^2)^{2m-r/s-s/2}$ times (X.4.3). Hence we obtain that for polytopes in extremal β-regions:

$$\mathrm{card}(E_m^\lambda) \leq \sigma(\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(v^*); 2\beta)(q^2)^{2m-r/2-s/2}.$$

We now consider β-regions which are adjacent to extremal ones, namely ones whose closure intersects the closure of an extremal β-region. If we have a sequence of points (m_i, r_i, s_i) in such a β-region which approaches the boundary of an extremal β-region, any vertex w_i in the associated polytopes which maximizes $(\beta, \cdot)$ must converge to v^* and hence the collection of coordinate inequalities which become equalities at w_i must be a subset of the inequalities that become equalities at v^*.

From the analysis in the proof of Lemma X.5, we must have

$$v^* = (r, 0, 0, 2m - r, 0)$$

with $r = s$. Therefore the collection of inequalities which become equalities at v^* is (supposing, for the moment, that $\alpha = (1, 0)$):

$$a_1 \leq r$$

$$
\begin{aligned}
a_2 &\geq 0 \\
k &\geq 0 \\
b_2 &\geq 0 \\
a_1 + 2a_2 + 2k + b_1 + 2b_2 &\leq 2m \\
-2a_1 - a_2 - k + b_1 + 2b_2 &\leq 2m - s - 2r \\
a_1 + k &\leq r.
\end{aligned}
$$

The cone of any such w_i in a region adjacent to an extremal β-region must be contained in the cone obtained from at least 5 of these inequalities (as the relevant polytope lives in a five-dimensional space). Furthermore this cone must contain β in its polar cone. Using the Sage Polyhedron package, we can enumerate all such possible cones; we in fact find that there are 4 possibilities and in all cases we have that β is in the interior of the polar cone. We may then use the same style of analysis as for the extremal β-regions to conclude that there exists a C_1 such that the following holds: (we can take C_1 to be the maximum over all possible $\sigma(\mathrm{Cone}_{\mathcal{P}_\alpha(m,r,s)}(w_i); 2\beta)$; by the preceding analysis this set is finite and hence the maximum is finite)

$$
\mathrm{card}(E_m^\lambda) \leq C_1 (q^2)^{(\beta, w_i)} \leq C_1 (q^2)^{2m - r/2 - s/2},
$$

for all (m, r, s) in a β-region adjacent to an extremal one. A similar analysis can be performed for $\alpha = (0, 1)$.

We are left now with analyzing the β-regions which are not extremal nor adjacent to an extremal one; we call such β-regions *tame*. Let B be a tame β-region. We seek to apply the degenerate case of Brion's formula for polytopes associated to points in B. However, we first note that Brion's formula requires that all vertices of the underlying polytope lie in some lattice. We are in particular interested in $\mathbb{Z}^5 \subset W$, but the vertices of the polytopes might not always be integer points. However, because all of the defining coordinate inequalities have integer coefficients, and because we are only really interested in the case when $(m, r, s) \in \mathbb{Z}^3$, there exists $L \in \mathbb{N}$ such that the vertices of the underlying polytopes always lie in $(\mathbb{Z}/L)^5$ (assuming m, r, s are integers). If we sum up (X.4.2) over the lattice points in this bigger lattice lying inside of the underlying polytope, then we still obtain an upper bound for $\mathrm{card}(E_m^\lambda)$.

We are now in a position to use degenerate Brion's formula (see Appendix C.5.2). This tells us that for all polytopes associated to points in B, we have

$$\sum_{x\in\mathcal{P}_\alpha(m,r,s)\cap(\mathbb{Z}/L)^5} = \sum_{\text{vertices } v \text{ of } \mathcal{P}_\alpha(m,r,s)} R_v(m,r,s)(q^2)^{h_v(m,r,s)},$$

where $R_v(m,r,s)$ is a polynomial in m,r,s and $h_v(m,r,s)$ is some linear functional in m,r,s. More specifically $h_v(m,r,s)$ is obtained by dotting β with a given vertex v in the family of polytopes associated to points in B; the coordinates of that vertex are in turn linear functionals in m,r,s so long as we are in a given β-region.

Consider the following locus in $V_{(m,r,s)}$:

$$A = \{(m,r,s) : m+r+s = 1 \text{ and } m,r,s \geq 0\}.$$

Consider the set $A\cap B$. Elements in this set are uniformly far away (in, say, the Euclidean metric) from the intersection of the extremal β-regions with A because all adjacent β-regions to the extremal ones are also in the complement of B. On $A\cap B$ consider the function $h_v(m,r,s)-(2m-r/2-s/2)$. By Lemma X.5, this function is negative on $A\cap B$, but since A is compact we in fact must have that there exists a C_2 such that

$$h_v(m,r,s)-(2m-r/2-s/2) \leq C_2 < 0$$

on $A\cap B$.

Suppose $m+r+s\geq 1$ (since we assume $(m,r,s)\in\mathbb{Z}^3$, this only excludes the case of $m=r=s=0$; in that case $\operatorname{card}(E_m^\lambda)=1$). We therefore get that there exists a $C_3, C_4 > 0$ and a $p \geq 0$ such that:

$$\begin{aligned} R_v(m,r,s)(q^2)^{h_v(m,r,s)} &\leq C_3(m+r+s)^p\frac{(q^2)^{h_v(m,r,s)}}{(q^2)^{2m-r/2-s/2}}(q^2)^{2m-r/2-s/2} \\ &\leq C_3(m+r+s)^p(q^2)^{C_2(m+r+s)}(q^2)^{2m-r/2-s/2} \\ &\leq C_4(q^2)^{2m-r/2-s/2}. \end{aligned}$$

By combining this bound over all vertices in v, we get that there exists a $C_5>0$ such that for all (m,r,s) in B:

$$\operatorname{card}(E_m^\lambda) \leq C_5(q^2)^{2m-r/2-s/2}.$$

Finally, by combining the bounds we obtained for the extremal, adjacent to extremal, and tame β-regions (which altogether consist of only finitely many regions), we get that there exists a C_6 such that for all $(m, r, s) \in \mathbb{N}^3$ we have:

$$\mathrm{card}(E_m^\lambda) \leq C_6(q^2)^{2m-r/2-s/2} = C_6(q^2)^{(\delta, m\cdot p^\dagger - \frac{\lambda}{2})}.$$

$\square$

X.5: The geometric bound using the Brumley-Matz polytope

We originally attempted to prove the analogue of Proposition II.15 using the polytope defined by Brumley-Matz in [BM21] which, somewhat coincidentally, is essentially exactly equal to the polytope that we call H. For simplicity assume that $\alpha = (1, 0)$. If we replace the first two equations of (X.3.2) in with the condition that $d_{A^+}(x, z) \in H_m$ and $d_{A^+}(y, z) \in H_m$, namely

$$\begin{aligned}
(a_1 + b_1) + 2(a_2 + k + b_2) &\leq 6m \\
2(a_1 + b_1) + (a_2 + k + b_2) &\leq 6m \\
(s - a_2 + k + b_1) + 2(r - a_1 - k + b_2) &\leq 6m \\
2(s - a_2 + k + b_1) + (r - a_1 - k + b_2) &\leq 6m,
\end{aligned}$$

and we replace (X.3.4) with the condition that

$$\frac{k}{2} + b_1 + b_2 \geq (\delta, m \cdot h^\dagger - \frac{d_{A^+}(x, y)}{2}) = 4m - \frac{r}{2} - \frac{s}{2}, \tag{X.5.1}$$

we get that such coordinates are exactly the conical row span of the following matrix in coordinates $(m, r, s, a_1, a_2, k, b_1, b_2)$:

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\
1 & 2 & 2 & 0 & 2 & 0 & 2 & 0 \\
1 & 2 & 2 & 2 & 0 & 0 & 0 & 2 \\
1 & 6 & 0 & 2 & 0 & 2 & 0 & 0 \\
1 & 4 & 4 & 2 & 2 & 0 & 0 & 0
\end{bmatrix}.$$

All such vectors in this cone in fact achieve equality in (X.5.1). Suppose for example we have (m, r, s) fixed such that $s = r = 2j$ for some $j \in \mathbb{N}$, and $2m \geq r$. Then for every

$0 \leq t \leq 2j = r$, the following element is in the cone:

$$\begin{bmatrix} m \\ 2j \\ 2j \\ t \\ 2j-t \\ 0 \\ 2m-t \\ 2m-2j+t \end{bmatrix} = (m-j)\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 2 \end{bmatrix} + \frac{t}{2}\begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} + (j-\frac{t}{2})\begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \\ 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

See Figures 19 and 20.

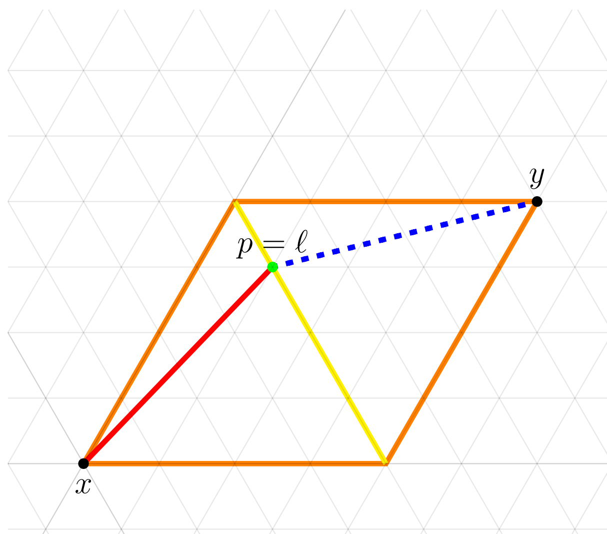


Figure 19: Suppose $d_{A^+}(x, y) = (r, s)$ with $r = s$. Suppose our confluence point p is some point along the "diagonal" of $\mathrm{para}(x, y)$ (the yellow line). Then $d_{A^+}(x, p)$ and $d_{A^+}(y, p)$ are equal.

This implies that for such (m, r, s), the cardinality of $E_m^{(r,s)}$ can only be upper bounded by something of size at least $r(q^2)^{(\delta, mh^\dagger - \frac{r}{2} - \frac{s}{2})}$. As mentioned in Remark IX.17, there is in fact some redundancy in the counting one obtains by this method. However, using the theory of Hecke paths [KM08], which we do not discuss here, one may in fact show that

$$\mathrm{card}(E_m^{(r,s)}) \gtrsim r(q^2)^{(\delta, mh^\dagger - \frac{r}{2} - \frac{s}{2})}.$$

This extra polynomial factor in fact results in an inability to complete the last step of the

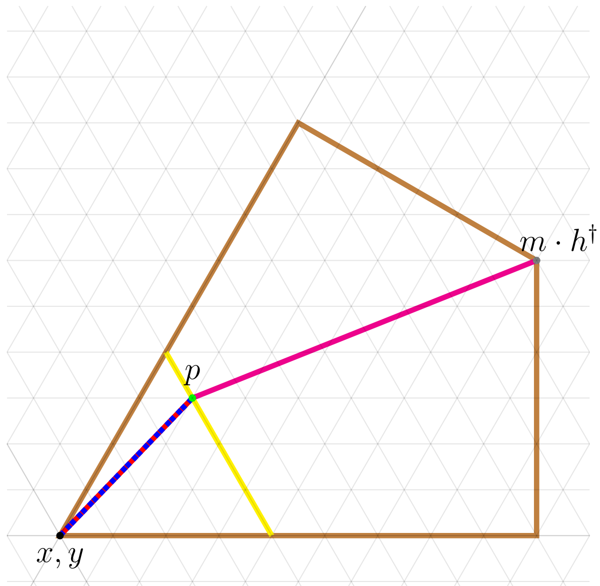


Figure 20: For each point p as in Figure 19 (here represented by the yellow line), we get that, after applying the appropriate element of $\mathfrak{S}_3$ to bring it into $\mathfrak{a}^+$, the line segments from x to p and from y to p are the same (represented by the solid red and blue dotted lines). For all such points p we have that $d_{A^+}(p, m \cdot h^\dagger)$ (represented by the pink line) has the same dot product with δ, namely $(\delta, m \cdot h^\dagger - r/2 - s/2)$, and such points result in (X.5.1) becoming an equation. Notice that the yellow line grows linearly with $r = s$, so we end up with an upper bound on E_m^λ of size $r(q^2)^{(\delta, m \cdot h^\dagger - r/2 - s/2)}$.

proof which is carried out in Chapter XI. It is for this reason that we ended up changing our polytope to P. In fact the important feature of p that circumvents this issue is that P has its vertex $p^\dagger$ "maximally singular", i.e. on the boundary of $\mathfrak{a}^+$.

CHAPTER XI

Bounding the Sum over H_M^Λ

XI.1: Proof of Proposition II.17

Proof of Proposition II.17 (Bounding the Sum over H_M^Λ). We now seek to prove:

$$\sum_{\lambda \in H_M^\Lambda} (q^2)^{\theta(\delta, \lambda - 2|\lambda|_H \cdot p^\dagger)} \lesssim M. \tag{XI.1.1}$$

We wish to use Brion's formula. However, the exponent in each summand is not quite a linear functional.

We now examine $|\lambda|_H$ more closely. Recall that H is defined by the following inequalities:

$$2r + s \leq 6,$$
$$r + 2s \leq 6.$$

We therefore have that

$$|\lambda|_H = \max\left\{\frac{2r+s}{6}, \frac{r+2s}{6}\right\},$$

where $\lambda = (r, s)$ in cone coordinates. The locus where $2r + s = r + 2s$ is exactly the locus $r = s$. We may partition H into two pieces: $H^{r \leq s}$ and $H^{r \geq s}$ by adding in the relevant constraint to the definition of H. See Figure 21. On $H^{r \leq s}$, we have that $|\lambda|_H = \frac{r}{6} + \frac{s}{3}$, and on $H^{r \geq s}$, we have $|\lambda|_H = \frac{r}{3} + \frac{s}{6}$. Hence we define:

$$\alpha_{r \leq s} := \frac{r}{6} + \frac{s}{3},$$
$$\alpha_{r \geq s} := \frac{r}{3} + \frac{s}{6}.$$

Hence we may split up the left hand side of (XI.1.1) into a sum over $(H^{r\leq s})^{\Lambda}_{M}$ and $(H^{r\geq s})^{\Lambda}_{M}$, and on each piece the summands become exactly of the form $(q^2)^{f(r,s)}$ where f is a linear functional in r, s.

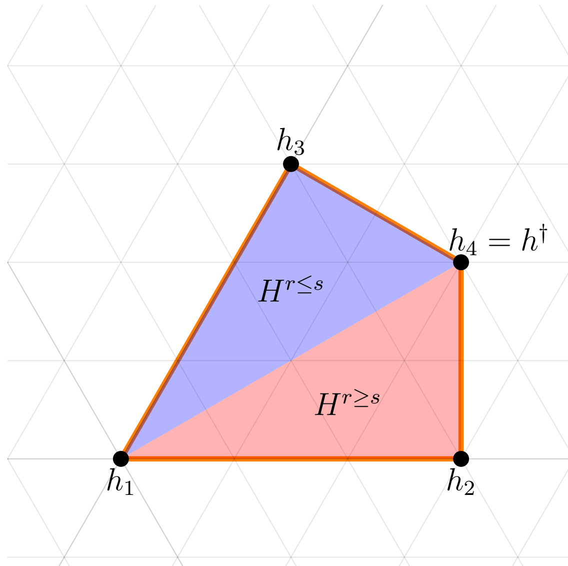


Figure 21: We may split up H (the brown polytope) into the locus where $r \geq s$ (the red region), and where $r \leq s$ (the blue region).

Let's focus now on $H^{r\leq s}$. On this polytope we have:

$$\begin{aligned}\theta(\delta, \lambda - 2|\lambda|_H \cdot p^\dagger) &= \theta\Big(\delta, \lambda - 2(\alpha_{r\leq s}, \lambda)\cdot p^\dagger\Big)\\ &= \theta\Big((\delta,\lambda) - 2(\alpha_{r\leq s}, \lambda)\cdot(\delta, p^\dagger)\Big)\\ &= \theta\Big(\lambda, \delta - 2(\delta, p^\dagger)\alpha_{r\leq s}\Big).\end{aligned}$$

In cone coordinates, we have

$$\begin{aligned}\delta &= (1,1),\\ p^\dagger &= (2,0),\\ \alpha_{r\leq s} &= \left(\frac{1}{3}, \frac{1}{6}\right),\\ \delta - 2(\delta, p^\dagger)\alpha_{r\leq s} &= \left(\frac{1}{3}, -\frac{1}{3}\right) =: \beta_{r\leq s}.\end{aligned}$$

Similarly we have

$$\alpha_{r\geq s} = \Big(\frac{1}{6}, \frac{1}{3}\Big),$$
$$\delta - 2(\delta, p^\dagger)\alpha_{r\geq s} = \Big(-\frac{1}{3}, \frac{1}{3}\Big) =: \beta_{r\geq s}.$$

The vertices of H are:

$$\begin{aligned} h_1 &= (0,0), \\ h_2 &= (3,0), \\ h_3 &= (0,3), \\ h_4 &= (2,2) =: h^\dagger. \end{aligned}$$

We have that the vertices of $H^{r\leq s}$ are h_1, h_3, h_4, and the vertices of $H^{r\geq s}$ are h_1, h_2, h_4.

On $H^{r\leq s}$ we have:

$$\begin{aligned} (h_1, \beta_{r\leq s}) &= 0 \\ (h_3, \beta_{r\leq s}) &= -\frac{1}{9} \\ (h_4, \beta_{r\leq s}) &= 0. \end{aligned}$$

Using degenerate Brion's formula (see Appendix C.5.2), we get that there exists a constant C_1 and a degree one polynomial $f(M)$ such that

$$\begin{aligned} \sum_{\lambda\in(H^{r\leq s})^\Lambda_M} (q^2)^{\theta(\delta,\lambda-2|\lambda|_H\cdot p^\dagger)} &= f(M) + C_1(q^2)^{-\frac{M\cdot\theta}{9}} \\ &\leq C_2 \cdot M \end{aligned}$$

for some C_2 and for all $M \geq 1$. An analogous analysis may be carried out on $(H^{r\geq s})^\Lambda_M$. $\square$

APPENDICES

APPENDIX A

Representation Theory Preliminaries

A.1: Notation and conventions

A.1.1: Notation relating to $\mathrm{PGL}(d, F)$ and $\mathrm{GL}(d, F)$

Let F be a non-archimedean local field, $\mathcal{O}$ its ring of integers, ϖ a uniformizer of $\mathcal{O}$, and q the order of the residue field. Let $G = \mathrm{PGL}(d, F)$ and $K = \mathrm{PGL}(d, \mathcal{O})$. Let $G' = \mathrm{GL}(d, F)$ and $K' = \mathrm{GL}(d, \mathcal{O})$. Let $(\mathcal{G}, \mathcal{K})$ denote either the pair (G, K) or the pair (G', K'). We denote the Haar measure by $\mathrm{vol}(\cdot)$ and assume that it is normalized so that $\mathrm{vol}(\mathcal{K}) = 1$. We let $\Gamma < \mathcal{G}$ denote a lattice. Let $\mathfrak{S}_d$ denote the symmetric group on d elements.

Let $T < G$ (and $T' < G'$, resp.) denote the subgroup of diagonal matrices. Let $A < T$ (and $A' < T'$, resp.) denote the subgroup of matrices all of whose diagonal entries are powers of ϖ. Let $A^+ \subset A$ (and $A'^+ \subset A'$, resp.) denote those elements for which the powers of ϖ along the diagonal are weakly decreasing. We use the shorthand $\mathcal{T}, \mathcal{A}, \mathcal{A}^+$ to denote one of either $\{T, T'\}, \{A, A'\}, \{A^+, A'^+\}$, respectively. To each element $a \in \mathcal{A}$ we can associate a tuple $\lambda = (\lambda_1, \ldots, \lambda_d)$ by recording the powers of ϖ along the diagonal (for A this tuple is only well-defined up to shifting all entries by the same integer). In case $a \in \mathcal{A}^+$, we have that $\lambda_1 \geq \cdots \geq \lambda_d$. We let ϖ^λ denote the matrix $\mathrm{diag}(\varpi^{\lambda_1}, \ldots, \varpi^{\lambda_d})$. We let (k^d) denote the tuple $(k, \ldots, k)$ of length d.

A.1.2: Notation relating to topological groups and vector spaces

We use M to denote a topological group. If $M_1, M_2 < M$, then we say that a function on M is (M_1, M_2)-invariant if it is invariant by M_1 multiplication on the left and by M_2 multipication on the right.

We use the convention that convolution is defined by:

$$f_1 * f_2(n) = \int_M f_1(m) f_2(m^{-1}n) dm.$$

Note that in case M is unimodular, all conventions for defining convolution are equivalent.

We shall use V to denote a (possibly infinite-dimensional) $\mathbb{C}$-vector space, and $\mathcal{H}$ to denote a separable Hilbert space. We let $(\cdot, \cdot)$ denote the pairing between V^* and V, and let $\langle \cdot, \cdot \rangle$ denote the inner product on $\mathcal{H}$. We use the convention that the inner product is linear in the second entry and sesquilinear in the first.

A.2: The structure of the spherical Hecke algebra

A.2.1: The spherical Hecke algebra

We define $H(\mathcal{G}, \mathcal{K})$ to be the algebra of compactly supported $(\mathcal{K}, \mathcal{K})$-invariant $\mathbb{C}$-valued functions on $\mathcal{G}$ with product given by convolution. This algebra is called the *spherical Hecke algebra* (of $\mathcal{G}$ with respect to $\mathcal{K}$).

A.2.2: The Cartan decomposition

Proposition A.1 (Cartan decomposition; [Mac95], p. 294)**.** *$\mathcal{G}$ is the disjoint union of the double cosets $\mathcal{K}\varpi^\lambda\mathcal{K}$ with $\varpi^\lambda \in \mathcal{A}^+$.*

This implies that each element of $H(\mathcal{G}, \mathcal{K})$ is completely determined by its restriction to $\mathcal{A}^+$. An obvious vector space basis for $H(\mathcal{G}, \mathcal{K})$ is given by the indicator functions for $\mathcal{K}\varpi^\lambda\mathcal{K}$ with $\varpi^\lambda \in A^+$. We shall denote these functions as c_λ and c'_λ respectively for G and G'.

A.2.3: Hall-Littlewood polynomials

Given $\lambda = (\lambda_1, \ldots, \lambda_d)$ with $\lambda_1 \geq \cdots \geq \lambda_d$, we let $P_\lambda(x_1, \ldots, x_d; t)$ be the *Hall-Littlewood polynomial* associated to λ defined as follows (note that in the definition, σ is permuting the x_i's rather than the λ_i's; see also [Mac95], Chapter III, Section 2):

$$P_\lambda(x_1, \ldots, x_d; t) = \frac{1}{\nu_\lambda(t)} \sum_{\sigma \in \mathfrak{S}_d} \sigma.\Big(x_1^{\lambda_1} \ldots x_d^{\lambda_d} \prod_{i<j} \frac{x_i - t x_j}{x_i - x_j}\Big), \tag{A.2.1}$$

$$\nu_\lambda(t) = \prod_{i \geq 0} \nu_{m_i}(t),$$

$$
\begin{aligned}
m_i &= \ \# \text{ of } \lambda_j \text{ equal to } i, \\
\nu_m(t) &= \prod_{i=1}^{m} \frac{1-t^i}{1-t}.
\end{aligned} \tag{A.2.2}
$$

A.2.4: The Satake isomorphism

Proposition A.2 (Satake isomorphism for $\mathrm{GL}(d,F)$; [Mac95], p. 296-297)**.** *Let $G' = \mathrm{GL}(d,F)$ and $K' = \mathrm{GL}(d,\mathcal{O})$. There is a $\mathbb{C}$-algebra isomorphism*

$$\theta : H(G',K') \to \mathbb{C}[x_1^{\pm 1},\dots,x_d^{\pm 1}]^{\mathfrak{S}_d}$$

given by

$$\theta(c_\lambda) = q^{-n(\lambda)} P_\lambda(x_1,\dots,x_d;q^{-1}), \tag{A.2.3}$$

where $n(\lambda) = \sum_i (i-1)\lambda_i$.

We have a surjective $\mathbb{C}$-algebra homomorphism from the spherical Hecke algebra for $\mathrm{GL}(d,F)$ to the spherical Hecke algebra for $\mathrm{PGL}(d,F)$ given by integrating a function over its $F^\times$-orbit (we shall prove that this is an algebra homomorphism as part of the proof of Proposition A.3 below). We call this map $\mathrm{sum}_{F^\times}$. The kernel of $\mathrm{sum}_{F^\times}$ is those functions whose $F^\times$-integral is zero. We have an explicit isomorphism $F^\times \simeq \mathcal{O}^\times \times \mathbb{Z}$ after choosing a uniformizer ϖ. Under the embedding $F^\times \hookrightarrow \mathrm{GL}(d,F)$, notice that $\mathcal{O}^\times \hookrightarrow \mathrm{GL}(d,\mathcal{O})$. We normalize the Haar measure on $F^\times$ so that $\mathcal{O}^\times$ has measure 1.

Proposition A.3 (Satake isomorphism for $\mathrm{PGL}(d,F)$)**.** *Let $G = \mathrm{PGL}(d,F)$ and $K = \mathrm{PGL}(d,\mathcal{O})$. The map θ in (A.2.3) restricted to the kernel of $\mathrm{sum}_{F^\times}$ has image equal to the ideal generated by $q^{-d(d-1)/2}x_1 \dots x_d - 1$. Hence we have a $\mathbb{C}$-algebra isomorphism*

$$\bar{\theta} : H(G,K) \to \mathbb{C}[x_1^{\pm 1},\dots,x_d^{\pm 1}]^{\mathfrak{S}_d}/(q^{-d(d-1)/2}x_1 \dots x_d - 1).$$

Proof. First we prove that $\mathrm{sum}_{F^\times}$ is an algebra homomorphism (that it is surjective is clear). Given $x \in \mathrm{GL}(d,F)$, we let $\bar{x} \in \mathrm{PGL}(d,F)$ denote its image under the natural map. On the other hand, given an element $y \in \mathrm{PGL}(d,F)$, we let $\tilde{y} \in \mathrm{GL}(d,F)$ be any lift of y. Given f in the spherical Hecke algebra for $\mathrm{GL}(d,F)$, we let $\bar{f}$ denote its image under $\mathrm{sum}_{F^\times}$. Hence we wish to show that $\overline{f * g} = \bar{f} * \bar{g}$.

Suppose h is in the spherical Hecke algebra for $\mathrm{GL}(d, F)$. Then

$$\begin{aligned}
\int_{F^\times} h(x\gamma)d\gamma &= \sum_m \int_{\mathcal{O}^\times} h(x\eta\varpi^{(m^d)})d\eta \\
&= \sum_m h(x\varpi^{(m^d)}), \\
\int_{\mathrm{GL}(d,F)} h(x)dx &= \int_{\mathrm{PGL}(d,F)} \sum_\ell h(\tilde{y}\varpi^{(\ell^d)})dy.
\end{aligned}$$

Therefore,

$$\begin{aligned}
f * g(x) &= \int_{\mathrm{PGL}(d,F)} \sum_\ell f(\tilde{y}\varpi^{(\ell^d)})g(\varpi^{((-\ell)^d)}\tilde{y}^{-1}x)dy, \\
\overline{f * g}(\bar{x}) &= \sum_m f * g(x\varpi^{(m^d)}) \\
&= \sum_m \int_{\mathrm{PGL}(d,F)} \sum_\ell f(\tilde{y}\varpi^{(\ell^d)})g(\tilde{y}^{-1}\varpi^{((-\ell)^d)}x\varpi^{(m^d)})dy \\
&= \int_{\mathrm{PGL}(d,F)} \sum_m \sum_\ell f(\tilde{y}\varpi^{(\ell^d)})g(\tilde{y}^{-1}x\varpi^{(m^d)})dy.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\bar{f} * \bar{g}(\bar{x}) &= \int_{\mathrm{PGL}(d,F)} \bar{f}(y)\bar{g}(y^{-1}\bar{x})dy \\
&= \int_{\mathrm{PGL}(d,F)} \Big(\sum_\ell f(\tilde{y}\varpi^{(\ell^d)})\Big)\Big(\sum_m g(\tilde{y}^{-1}x\varpi^{(m^d)})\Big)dy \\
&= \int_{\mathrm{PGL}(d,F)} \sum_m \sum_\ell f(\tilde{y}\varpi^{(\ell^d)})g(\tilde{y}^{-1}x\varpi^{(m^d)})dy.
\end{aligned}$$

Now we wish to prove that the image under θ of the kernel of $\mathrm{sum}_{F^\times}$ is the ideal generated by $q^{-n(n-1)/2}x_1 \ldots x_d - 1$. MacDonald computes that $\theta(c_{(1^d)}) = q^{-n(n-1)/2}x_1 \ldots x_d$ ([Mac95], p. 297). Clearly $\theta(c_{(0^d)}) = 1$. Hence we are really trying to show that the kernel of $\mathrm{sum}_{F^\times}$ is the ideal generated by $c_{(1^d)} - c_{(0^d)}$. It is immediate that this element is indeed in the kernel because $\varpi^{(1^d)} \cdot K\varpi^{(0^d)}K = K\varpi^{(1^d)}K$. More generally we have $\varpi^{(1^d)} \cdot K\varpi^\lambda K = K\varpi^{\lambda+(1^d)}K$.

Now suppose f is in the spherical Hecke algebra for $\mathrm{GL}(d, F)$. Then we can write

$$f = \sum_\lambda \alpha_\lambda c_\lambda$$

$$= \sum_{\lambda \text{ with } \lambda_d = 0} \Big(\sum_k \alpha_{\lambda+(k^d)} c_{\lambda+(k^d)}\Big),$$

$$\bar{f} = \sum_{\lambda \text{ with } \lambda_d = 0} \Big(\sum_k \alpha_{\lambda+(k^d)}\Big)\overline{c_\lambda} \tag{A.2.4}$$

with $\alpha_\lambda \in \mathbb{C}$. For a fixed λ with $\lambda_d = 0$, only finitely many of the $\alpha_{\lambda+(k^d)}$ are non-zero since f is compactly supported. Let N be the largest such k.

Let's now subtract off the following element from f (notice that $c_\lambda * c_{(k^d)} = c_{\lambda+(k^d)}$; [Mac95], p. 295):

$$\alpha_{\lambda+(N^d)} \left(c_{\lambda+(N^d)} - c_{\lambda+((N-1)^d)}\right) = \alpha_{\lambda+(N^d)}\ c_{\lambda+((N-1)^d)}\ \left(c_{(1^d)} - c_{(0^d)}\right).$$

Then the coefficient of $c_{\lambda+(N^d)}$ becomes zero and the coefficient of $c_{\lambda+((N-1)^d)}$ becomes $\alpha_{\lambda+((N-1)^d)} + \alpha_{\lambda+(N^d)}$.

Suppose L is the smallest k such that $d_{\lambda+(k^d)} \neq 0$. If we repeat the above process $N - L$ times, then we will get that all coefficients of $c_{\lambda+(k^d)}$ are zero except possibly for $c_{\lambda+(L^d)}$ whose coefficient will be $\sum_k \alpha_{\lambda+(k^d)}$. If f is in the kernel, then $\bar{f} = 0$, so by (A.2.4), this coefficient is also zero. If we repeat this process with each λ such that $\lambda_d = 0$, we end up with the zero function after subtracting off an element in the ideal generated by $c_{(1^d)} - c_{(0^d)}$. $\square$

A.3: Spherical functions

A.3.1: Spherical functions

A *($\mathcal{K}$-)spherical function* on $\mathcal{G}$ is a complex-valued continuous function ω on $\mathcal{G}$ satisfying:

(1) ω is $(\mathcal{K}, \mathcal{K})$-invariant.

(2) $\omega * f = \lambda_f \cdot \omega$ for each $f \in H(\mathcal{G}, \mathcal{K})$, where $\lambda_f \in \mathbb{C}$.

(3) $\omega(1) = 1$.

A.3.2: The spherical Fourier transform

We define the *spherical Fourier transform* of $f \in H(\mathcal{G}, \mathcal{K})$ to be the function $\hat{f}$ on the set of spherical functions defined by $\hat{f}(\omega) = (\omega * f)(1)$ (that is, $\omega * f = \hat{f}(\omega) \cdot \omega$). On the other hand, given a spherical function ω, we may take its spherical Fourier transform by defining a function $\hat{\omega}$ on $H(\mathcal{G}, \mathcal{K})$ by $\hat{\omega}(f) = (\omega * f)(1)$.

Proposition A.4 ([Mac71], Proposition (1.2.6), p. 6)**.** *The map $\hat{\omega} : H(\mathcal{G}, \mathcal{K}) \to \mathbb{C}$ is a $\mathbb{C}$-algebra homomorphism. Furthermore, there is a bijection from the set of spherical functions to the set* $\mathrm{Hom}_{\mathbb{C}-\mathrm{alg.}}(H(\mathcal{G}, \mathcal{K}), \mathbb{C})$ *given by the spherical Fourier transform.*

A.3.3: The variety $\Omega(\mathcal{G})$

Propositions A.2 and A.3 identify the spherical Hecke algebras with the coordinate rings of affine varieties. Recall that by the nullstellensatz, the points of these varieties are in bijection with $\mathbb{C}$-algebra homomorphisms of the coordinate rings to $\mathbb{C}$ (given by evaluation at a point). However, we have identified $\mathbb{C}$-algebra homomorphisms of the spherical Hecke algebra to $\mathbb{C}$ with spherical functions in Proposition A.4. Hence the points on these affine varieties can be naturally identified with spherical functions. Furthermore, since all the isomorphisms and bijections are explicit, we can derive an explicit formula for the spherical functions.

Let $\Omega(\mathcal{G})$ denote the affine variety associated to $\mathcal{G}$. Points on $\Omega(\mathrm{GL}(d, F))$ correspond to tuples $(z_1, \ldots, z_d)$ with $z_i \in \mathbb{C}^\times$ modulo permuting the coordinates. We can change variables by letting $z_i = q^{\frac{1}{2}(d-1)-s_i}$ with $s_i \in \mathbb{C} \pmod{2\pi i\mathbb{Z}/\ln(q)}$. Points on $\Omega(\mathrm{PGL}(d, F))$ under these coordinates are now tuples $s = (s_1, \ldots, s_d)$ such that $s_1 + \cdots + s_d = 0 \pmod{2\pi i\mathbb{Z}/\ln(q)}$.

A.3.4: Formula for the spherical functions

Consider c'_λ in the spherical Hecke algebra for $\mathrm{GL}(d, F)$. Let

$$\delta := \frac{1}{2}(d-1, d-3, \ldots, 3-d, 1-d);$$

this is exactly half the sum of positive roots for the Lie algebra $\mathfrak{sl}(d)$. Given $s = (s_1, \ldots, s_d)$ determining a point on $\Omega(\mathrm{GL}(d, F))$, let ω_s be the associated spherical function. Then $\hat{c}'_\lambda(\omega_s) = \hat{\omega}_s(c'_\lambda)$ is given by evaluating $\theta(c'_\lambda)$ at the point $(q^{\frac{1}{2}(d-1)-s_1}, \ldots, q^{\frac{1}{2}(d-1)-s_d})$ in $\Omega(\mathrm{GL}(d, F))$ which results in

$$\hat{c}'_\lambda(\omega_s) = q^{(\delta,\lambda)} P_\lambda(q^{-s_1}, \ldots, q^{-s_d}; q^{-1}). \tag{A.3.1}$$

On the other hand,

$$\hat{c}'_\lambda(\omega_s) = \omega_s(\varpi^{-\lambda}) \cdot \mathrm{vol}(K\varpi^\lambda K),$$

and it is known that ([Mac95], p. 298)

$$\operatorname{vol}(K\varpi^\lambda K) = q^{2(\delta,\lambda)}\frac{\nu_d(q^{-1})}{\nu_\lambda(q^{-1})}. \tag{A.3.2}$$

Therefore

$$\omega_s(\varpi^{-\lambda}) = \frac{q^{-(\delta,\lambda)}}{\nu_d(q^{-1})}\sum_{\sigma\in\mathfrak{S}_d}\sigma.\Big(q^{-(s,\lambda)}\prod_{i<j}\frac{q^{-s_i}-q^{-(s_j+1)}}{q^{-s_i}-q^{-s_j}}\Big). \tag{A.3.3}$$

The $\lambda \in A^+$ such that $\lambda_d = 0$ give elements c'_λ in the spherical Hecke algebra for $\mathrm{GL}(d,F)$ whose images $\overline{c'_\lambda}$ under $\mathrm{sum}_{F^\times}$ give a basis for the spherical Hecke algebra for $\mathrm{PGL}(d,F)$ (namely the elements c_λ). Similarly to before we can compute $\hat{\omega}_s(\overline{c'_\lambda})$, with s now determining a point in $\Omega(\mathrm{PGL}(d,F))$, by simply taking $\theta(c'_\lambda)$ and restricting it to points such that $s_1+\cdots+s_d = 0 \pmod{2\pi i\mathbb{Z}/\ln(q)}$. Hence (A.3.1) also gives a formula for the spherical functions for $\mathrm{PGL}(d,F)$.

A.4: Spherical representations

Recall that $G = \mathrm{PGL}(d,F)$ and $K = \mathrm{PGL}(d,\mathcal{O})$. We write Ω as shorthand for $\Omega(G)$. We shall now turn our attention to representations of G. Much of the content of this section can be found in Cartier [Car79].

A.4.1: Spherical representations

We say that a representation (ρ, V) is *smooth* if for each $v \in V$, $\mathrm{Stab}_G(v)$ contains a compact open subgroup. We say that it is *admissible* if V^H is finite dimensional for each compact open subgroup H. We shall be particularly interested in the smooth, admissible, irreducible representations of G which contain a non-zero K-fixed vector. Such representations are called *spherical representations.*

Proposition A.5 ([Car79], p. 152)**.** *If (ρ,V) is a smooth, admissible, irreducible representation of G, then V^K is at most one-dimensional. Hence for spherical representations, V^K is exactly one-dimensional.*

Any time we have a smooth, admissible, irreducible representation, we get an induced representation of $H(G,K)$. Furthermore, each $f \in H(G,K)$ acts as a projection operator onto V^K ([Car79], p. 117). Since V^K is one-dimensional in the case of spherical representations, we can in fact associate an element in $\mathbb{C}$ to each $f \in H(G,K)$. Hence given such a representation we get a $\mathbb{C}$-algebra homomorphism from $H(G,K)$ to $\mathbb{C}$.

Proposition A.6 ([Car79], p. 152)**.** *The above-described map defines a bijection:*

$$\big\{\text{spherical representations of } G\big\} \leftrightarrow \mathrm{Hom}_{\mathbb{C}\text{-alg.}}(H(G,K),\mathbb{C}).$$

A.4.2: Explicit description of the spherical representations

In fact we can explicitly construct all of these representations. A character of $F^\times$ is called *unramified* if it is invariant under $\mathcal{O}^\times$. Since the choice of a uniformizer gives an isomorphism $F^\times \simeq \mathbb{Z} \times \mathcal{O}^\times$, we may identify unramified characters with characters of $\mathbb{Z}$ (as an additive group).

Recall that T consists of all diagonal matrices in G; it is a maximal split torus. We say that a character of T is *unramified* if it is invariant under $K \cap T$. Since $T/(K \cap T) \simeq A \simeq \mathbb{Z}^{d-1}$, an unramified character is determined by just $d-1$ non-zero complex numbers, or, perhaps more naturally, d complex numbers whose product is 1. There is a natural $\mathfrak{S}_d$-action on such characters given by permuting this d-tuple.

Let $B < G$ be the collection of all upper triangular matrices and let χ be an unramified character of T. We can extend χ to B using the homomorphism $B \to T$ which simply reads off the diagonal entries. Let Δ be the modular character for B, that is, given $b \in B$ with diagonal entries $(a_1, \ldots, a_d)$,

$$\Delta(b) = |a_1|^{d-1}|a_2|^{d-3} \ldots |a_d|^{1-d},$$

where $|\cdot|$ is the norm induced by the valuation on F (which is normalized so that $|\varpi| = q^{-1}$). Notice that $\Delta(b)$ is actually the extension of an unramified character on T to B.

Let I_χ be defined as the space of locally constant functions $f : G \to \mathbb{C}$ such that

$$f(bg) = \Delta^{1/2}(b)\chi(b)f(g), \qquad b \in B,\ g \in G$$

with the action of G from the right ($g.f(x) = f(xg^{-1})$). This is simply the usual definition of induction $\mathrm{Ind}_B^G(\chi)$ twisted by the character $\Delta^{1/2}$.

Two spaces I_χ and $I_{\chi'}$ are isomorphic if and and only if $\chi' = \sigma.\chi$ for some $\sigma \in \mathfrak{S}_d$ ([Car79], Section 3.3). The Iwasawa decomposition tells us that $G = BK$, and it is clear that $B \cap K$ is those matrices which are upper triangular with entries in $\mathcal{O}$ and with diagonal entries in $\mathcal{O}^\times$. Thus, inside of I_χ we have an obvious choice of K-fixed vector, f_χ ([Car79], p. 143):

$$f_\chi(bk) := \Delta^{1/2}(b)\chi(b).$$

In general the I_χ are not irreducible. However, I_χ has a finite decomposition series as a G-module, and exactly one of its factors is spherical.

Proposition A.7 ([Car79], p. 152)**.** *Every spherical representation of G is isomorphic to a subquotient of some I_χ, with χ unique up to permutation.*

A.4.3: Spherical representations, spherical functions, and the variety Ω

Proposition A.3 identifies spherical functions with points on the variety Ω. On the other hand Proposition A.7 identifies spherical functions with unramified characters, i.e. $\mathrm{Hom}_{\mathrm{gps}}(A, \mathbb{C}^\times)$. A straightforward calculation shows that if $s = (s_1, \ldots, s_d)$, then combining these identifications gives:

$$\left(\chi : \varpi^\lambda \mapsto q^{(\lambda,s)}\right) \in \mathrm{Hom}_{\mathrm{gps}}(A, \mathbb{C}^\times) \quad \leftrightarrow \quad (q^{s_1}, \ldots, q^{s_d}) \in \Omega.$$

The tuple $(q^{s_1}, \ldots, q^{s_d})$ is called the *Satake parameters* of χ.

A.5: Unitary representations

A.5.1: Unitary representations

Suppose M is a topological group. Suppose $(\rho, \mathcal{H})$ is a representation of M on a Hilbert space $\mathcal{H}$ such that M acts by unitary transformations, and the map $M \to \mathcal{H}$ defined by $m \mapsto \rho(m).v$ is continuous for every $v \in \mathcal{H}$; then ρ is called a *unitary representation*. Given vectors $v, w \in \mathcal{H}$, we obtain a function on M via $m \mapsto \langle v, \rho(m).w\rangle$; such functions are known as *matrix coefficients*. When $v = w$, it is called a *diagonal matrix coefficient*. Two unitary representations $(\rho_1, \mathcal{H}_1)$ and $(\rho_2, \mathcal{H}_2)$ are equivalent if there is an intertwining Hilbert space isomorphism between $\mathcal{H}_1$ and $\mathcal{H}_2$. The collection of all irreducible unitary representations of M up to equivalence is called the *unitary dual*, denoted $\hat{M}$.

Given a unitary representation $(\rho, \mathcal{H})$ and a function $f \in L^1(M)$, we can construct a corresponding operator on $\mathcal{H}$ which we denote by $\hat{f}(\rho)$. It is defined by

$$\hat{f}(\rho).v := \int_M f(m)\rho(m^{-1}).v \; dm,$$

with $v \in \mathcal{H}$. In particular, elements of $H(G, K)$ are clearly in $L^1(G)$; hence given a unitary representation of G, we always get an associated $H(G, K)$-representation.

A.5.2: Functions of positive type

A function $\phi : M \to \mathbb{C}$ is said to be of *positive type* if it is continuous and for every $m_1, \ldots, m_n \in M$, the matrix $[\phi(m_j^{-1}m_i)]_{i,j}$ is positive semidefinite. Diagonal matrix coefficients are functions of positive type; in fact the converse is also true as formalized by the following theorem.

Theorem A.8 (GNS construction; [BdlHV08], Theorem C.4.10, p. 376)**.** *Suppose ϕ is a function of positive type on the topological group M. Then there exists a triple $(\rho_\phi, \mathcal{H}_\phi, v_\phi)$ consisting of a cyclic unitary representation ρ_ϕ acting on a Hilbert space $\mathcal{H}_\phi$ with cyclic vector v_ϕ such that $\phi(m) = \langle v_\phi, \rho_\phi(m).v_\phi\rangle$. Furthermore, this triple is unique.*

A.5.3: Weak containment

We can use functions of positive type to measure how similar two different representations are. Suppose $(\rho_1, \mathcal{H}_1)$ and $(\rho_2, \mathcal{H}_2)$ are two not necessarily irreducible unitary representations. We say that ρ_1 is *strongly contained* in ρ_2, written $\rho_1 < \rho_2$, if ρ_1 occurs as a subrepresentation of ρ_2. Notice that this would imply that every function of positive type associated to ρ_1 (i.e. diagonal matrix coefficients) are also functions of positive type associated to ρ_2. This motivates the following definition: we say that ρ_1 is *weakly contained* in ρ_2, written $\rho_1 \prec \rho_2$, if every function of positive type associated to ρ_1 can be uniformly approximated on compact subsets of M by finite sums of functions of positive type associated to ρ_2. Formally this means that for every $v \in \mathcal{H}_1$, every compact subset $Q \subset M$, and every $\varepsilon > 0$, there exist $w_1, \ldots, w_n \in \mathcal{H}_2$ such that, for all $x \in Q$:

$$\left| \langle v, \rho_1(x).v\rangle - \sum_{i=1}^{n} \langle w_i, \rho_2(x).w_i\rangle \right| < \varepsilon.$$

Proposition A.9 ([BdlHV08], Appendix F)**.** *We have $\rho_1 \prec \rho_2$ if and only if for every $f \in L^1(M)$, we have $||\hat{f}(\rho_1)|| \leq ||\hat{f}(\rho_2)||$ where $||\cdot||$ denotes the operator norm.*

A.5.4: The Fell topology

We may also use functions of positive type to define a topology on $\hat{M}$ called the *Fell topology*. Let $\phi_1, \ldots, \phi_n$ be functions of positive type associated to ρ, let Q be a compact subset of M, and let $\varepsilon > 0$. Let $U(\rho; \phi_1, \ldots, \phi_n; Q; \varepsilon)$ be the set of all elements $\iota \in \hat{G}$ such that for each ϕ_i, there exists a function ψ of positive type associated to ι such that $|\phi_i(x) - \psi(x)| < \varepsilon$ for all $x \in Q$. We takes these sets as a basis of open sets for the Fell topology.

Proposition A.10 ([BdlHV08], Appendix F)**.** *A net $(\rho_i)_i$ converges to ρ in the Fell topology if and only if $\rho \prec \oplus_j \rho_j$ for every subset $(\rho_j)_j$ of $(\rho_i)_i$.*

A.6: Class 1 representations

A.6.1: Unitary representations are smooth and admissible

The following proposition shows that unitary representations of $G = \mathrm{PGL}(d, F)$ (or in fact any semisimple algebraic group over a non-archimedean local field) are essentially also smooth and admissible.

Proposition A.11 ([Car79], Corollary 2.3, p. 133)**.** *Suppose $(\rho, \mathcal{H})$ is an irreducible unitary representation of G. Let $\mathcal{H}_\infty$ be the collection of vectors in $\mathcal{H}$ whose stabilizer contains a compact open subgroup. Then $\mathcal{H}_\infty$ is dense in $\mathcal{H}$, and $(\rho|_{\mathcal{H}_\infty}, \mathcal{H}_\infty)$ is a smooth, admissible, irreducible representation of G.*

A.6.2: Class 1 representations, principal series, and complementary series

We call a unitary representation of G *class 1* if it is irreducible and contains a K-fixed vector ($K = \mathrm{PGL}(d, \mathcal{O})$). The classification of class 1 representations is hence reduced to figuring out which spherical representations are unitarizable. This is in general a difficult question to answer, but there is a natural class of representations which are known to be unitarizable (recall the definition of I_χ from Appendix A.4):

Proposition A.12 ([Mac71], Proposition 3.3.1, p. 45)**.** *Suppose χ is an unramified unitary character, i.e., its Satake parameters lie on S^1. Then I_χ is unitarizable.*

Such unitary representations are known as the *principal series*. All the other unitary spherical representations are called the *complementary series*.

A.6.3: Class 1 representations and spherical functions of positive type

It turns out that the spherical functions of positive type are exactly the diagonal matrix coefficients for class 1 representations corresponding to the K-fixed vector.

Proposition A.13 ([Mac71], Theorem 1.4.4, p. 10)**.** *Suppose $(\rho, \mathcal{H})$ is a class 1 representation. Let v be a K-invariant unit vector. Then the function $\langle v, \rho(g).v\rangle$ is a spherical function and of positive type. On the other hand, if (ρ, V) is a spherical representation and the associated spherical function is of positive type, then (ρ, V) is unitarizable.*

A.7: Spectral parameters

Recall that we have the following bijections:

$$\mathrm{Hom}_{\mathbb{C}\text{-algs.}}(H(G,K),\mathbb{C}) \leftrightarrow \Omega(G) \leftrightarrow \text{spherical representations of } G.$$

Hence given a spherical representation, we get an associated point in $\Omega(G)$. We shall refer to this point as the *spectral parameter of the representation*. On the other hand, given a space with an $H(G,K)$-action (such as any representation of G) and an eigenvector v of $H(G,K)$, we get an associated algebra homomorphism of $H(G,K)$ by reading off the eigenvalue of each element. Hence this also gives us a point in $\Omega(G)$. We shall refer to this as the *spectral parameter of the eigenvector*.

A.8: Unitary representations from measure-preserving actions

Suppose M acts in a measure-preserving way on a measure space (X,μ). Then we get an associated unitary representation of M on $L^2(X,\mu)$. A common occurrence of this construction is the action of G on $\Gamma\backslash G$ where Γ is a lattice and the associated unitary representation is $L^2(\Gamma\backslash G)$. The case of Γ a cocompact is particularly nice.

Proposition A.14 ([DE09], Theorem 9.2.2)**.** *Suppose $\Gamma < G$ is a cocompact lattice. Then $L^2(\Gamma\backslash G)$ decomposes as a countable direct sum of orthogonal irreducible unitary representations, each of finite multiplicity.*

We can also consider the space $L^2(\Gamma\backslash G/K)$. Elements in this space may be thought of as (Γ,K)-invariant functions on G. When we convolve on the right with elements in $H(G,K)$, we end up with functions which are still (Γ,K)-invariant, so we have an $H(G,K)$-action.

In fact elements in $H(G,K)$ act as normal operators, so since $H(G,K)$ is a commutative algebra, $L^2(\Gamma\backslash G/K)$ has an orthonormal basis of joint eigenfunctions of $H(G,K)$.

Proposition A.15. *There is a bijection between the eigenspaces of $H(G,K)$ acting on $L^2(\Gamma\backslash G/K)$ and the isotypic components of class 1 representations that show up in the decomposition of $L^2(\Gamma\backslash G)$ into irreducibles. The dimension of each eigenspace is equal to the multiplicity of the corresponding class 1 representation.*

Proof. Let $(\rho,\mathcal{H})$ be a class 1 representation that shows up with multiplicity n_ρ. By decomposing the $(\rho,\mathcal{H})$-isotypic component into a direct sum of n_ρ copies of this representation, we may subsequently pick out a unique (up to scaling) unit vector in each copy which is K-invariant. Each of these functions define (Γ,K)-invariant functions on G. Furthermore, they must each be eigenfunctions of $H(G,K)$ all with the same spectral parameter. Hence they span an eigenspace of dimension n_ρ inside $L^2(\Gamma\backslash G/K)$. We claim that all of these $H(G,K)$-eigenspaces from all isotypic components of class 1 representations must span $L^2(\Gamma\backslash G/K)$. This is because any K-invariant vector in $L^2(\Gamma\backslash G)$, when projected orthogonally onto each isotypic components must only have non-zero component in the class 1 representations (because all other representations do not have any K-invariant vectors). Furthermore, the dimension of K-invariant vectors in each isotypic component is exactly n_ρ. Hence these eigenspaces span $L^2(\Gamma\backslash G/K)$. $\square$

If all isotypic components have multiplicity one, then all eigenspaces are one-dimensional and there is a canonical (up to scaling) orthonormal basis for $L^2(\Gamma\backslash G/K)$. If some isotypic component has multiplicity $n_\rho > 1$, then, on the one hand there are many ways to decompose the isotypic component into irreducible subspaces, and on the other hand there are many ways to refine the corresponding eigenspace into an orthonormal basis. In fact, it is straightforward to see that doing one of these further decompositions forces a corresponding decomposition on the other side.

A.9: Spherical Plancherel measure

A.9.1: Plancherel-Godement theorem

Let Ω^+ be the set of class 1 representations of G, or equivalently the set of spherical functions of positive type. Let $\mathcal{L}^p(G,K)$ denote the space of L^p functions on G which are (K,K)-invariant. Recall that we previously defined the spherical Fourier transform, which mapped elements of $H(G,K)$ to functions on Ω via

$$\hat{f}(\omega) = \int_G f(g)\omega(g^{-1})dg.$$

We now modify the definition so that we map elements in $\mathcal{L}^1(G,K)$ to functions on Ω^+ (by only integrating against spherical functions of positive type). We then give Ω^+ the weakest topology making all of the Fourier transforms of elements in $\mathcal{L}^1(G,K)$ continuous. This is equivalently the topology with basis given by $N(C,U) = \{\omega \in \Omega^+ : \omega(C) \subset U\}$ where $C \subset G$ is compact and $U \subset \mathbb{C}$ is open ([Mac71], p. 12). From this latter definition we see that this topology is equivalent to the restriction of the Fell topology on $\hat{G}$ to the class 1 representations.

Proposition A.16 (Plancherel-Godement theorem; [Mac71], Theorem 1.5.1, p. 13)**.** *There exists a unique measure μ on Ω^+ such that*

(1) $f \in H(G,K) \Rightarrow \hat{f} \in L^2(\Omega^+, \mu)$,

(2) $\int_G |f(g)|^2 dg = \int_{\Omega^+} |\hat{f}(\omega)|^2 d\mu(\omega)$ for all $f \in H(G,K)$.

Moreoever, the mapping $f \mapsto \hat{f}$ extends to an isomorphism of Hilbert spaces $\mathcal{L}^2(G,K) \to L^2(\Omega^+, \mu)$.

We call μ the *spherical Plancherel measure.*

A.9.2: Explicit formula for the spherical Plancherel measure

The measure μ is known explicitly for $G = \mathrm{PGL}(d,F)$ and $K = \mathrm{PGL}(d,\mathcal{O})$. This was essentially first computed by MacDonald ([Mac71], Theorem 5.1.2; see also Tadic [Tad83a], p. 230). Recall that T denotes the subgroup of diagonal matrices in G. Let $\hat{T}$ denote the collection of unramified unitary characters of T. As noted previously, given $\chi \in \hat{T}$, we get an associated class 1 representation of G which is unique up to the $\mathfrak{S}_d$-action on $\hat{T}$ coming from the $\mathfrak{S}_d$-action on T given by permuting diagonal elements. We can associate to χ a tuple $s = (s_1, \ldots, s_d)$ with $s_j \in (i\mathbb{R})/\left(\frac{2\pi i}{\ln(q)}\mathbb{Z}\right)$ such that $\chi(\varpi^\lambda) = q^{(s,\lambda)}$ (notice then that $\prod q^{s_j} = 1$ and hence $s_1 + \cdots + s_d = 0$).

We define the *c-function* on $\hat{T}$ by the formula

$$c(s) = \prod_{j<k} \frac{q^{s_j} - q^{-1}q^{s_k}}{q^{s_j} - q^{s_k}}.$$

Let $Q(t)$ be the *Poincaré polynomial* for $\mathfrak{S}_d$ (as a Coxeter group; see Appendix B.1), that is

$$Q(t) = \sum_{\sigma \in \mathfrak{S}_d} t^{\ell(\sigma)} = \prod_{i=1}^{d-1} \frac{1-t^{i+1}}{1-t},$$

with $\ell(\sigma)$ equal to the Coxeter length of σ. The spherical Plancherel measure is supported on $\hat{T}$ (or really $\hat{T}/\mathfrak{S}_d$) and is given by the formula

$$d\mu(\omega_s) = \frac{Q(q^{-1})}{n!} \frac{ds}{|c(s)|^2}. \tag{A.9.1}$$

The most salient feature of this formula for our purposes is simply that it shows that the spherical Plancherel measure is absolutely continuous with respect to the Haar measure on $\hat{T}$.

A.10: The Plancherel theorem

A.10.1: von Neumann algebras and type I groups

A *von Neumann algebra* is a $*$-algebra of bounded operators on a Hilbert space which is closed in the weak operator topology and contains the identity operator. Given a group M and a unitary representation $(\rho, \mathcal{H})$, we can consider the von Neumann algebra generated by $\rho(M) \subset B(\mathcal{H})$, the algebra of bounded operators. A von Neumann algebra is called a *factor* if its center is trivial, that is, only consists of scalar multiples of the identity. It is called a *type I factor* if it is isomorphic to the full algebra of bounded operators on some Hilbert space. A group M is called *type I* if every factor representation is type I. All reductive algebraic groups over p-adic fields are known to be type I [Ber74].

A.10.2: Direct integral representations

Suppose (X, ν) is a measure space with measure ν. The Hilbert space $L^2_\nu(X, \mathcal{H})$ is defined as all measurable functions from $f : X \to \mathcal{H}$ such that $\int_X ||f(x)||^2 d\nu < \infty$ modulo functions which are zero outside of a measure zero set.

A *measurable family of Hilbert spaces* over a Borel space X is a collection of (separable) Hilbert spaces $\{\mathcal{H}_x\}_{x \in X}$ such that the for every $n \in \mathbb{N} \cup \omega$, the set $X_n = \{x : \dim(\mathcal{H}_x) = n\}$ is measurable, and all these $\mathcal{H}_x$'s have been explicitly identified with the same n-dimensional Hilbert space $\mathcal{H}_n$. This is equivalent to choosing a collection of n functions $\zeta_i(x)$ for $x \in X_n$

with $\zeta_i(x) \in \mathcal{H}_x$ such that for each $x \in X_n$, the $\zeta_i(x)$ form an orthonormal basis of $\mathcal{H}_x$. This structure allows us to specify which sections $\{s_x\}_{x\in X}$ are measurable; namely we need that on each X_n, the map from X_n to $\mathcal{H}_n$ is measurable. Such sections are called *measurable vector fields*. If we further specify a measure ν on X, we can then define the *direct integral* $\int_X^\oplus \mathcal{H}_x d\nu(x)$ as the Hilbert space direct sum of the $L^2_\mu(X_n, \mathcal{H}_n)$.

Let M be a second countable, locally compact, unimodular, type I group. Let $\mathcal{H}_n$ be a fixed n-dimensional Hilbert space. Let Irr_n denote the set of all irreducible unitary representations of M on $\mathcal{H}_n$ (not up to equivalence!). Define a σ-algebra on Irr_n by taking the smallest σ-algebra such that $\langle v, \rho(m).w\rangle$ is measurable for all $m \in M$ and $v, w \in \mathcal{H}_n$. Let $\hat{G}_n$ denote the set of all equivalence classes of irreducible unitary n-dimensional representations of G. We clearly have a surjection from Irr_n to $\hat{G}_n$. Define a σ-algebra structure on $\hat{G}_n$ as the one induced from this surjection (a set is measurable if and only if its preimage is measurable). Take the union σ-algebra on all of $\hat{G}$ in this way (all irreducible representations of G are countable dimensional; [Car79] p. 118). This is called the *Mackey Borel structure*. In our case it is known to be a standard Borel space ([Fol95], Theorem 7.6).

Given a measurable family of Hilbert spaces $\{\mathcal{H}_x\}_{x\in X}$ and an operator on each Hilbert space, we say that this is a *measurable family of operators* if it maps measurable vector fields to measurable vector fields. A *measurable field of representations* of a group M is a measurable family of Hilbert spaces, each of which carries a unitary action of M, such that the field of operators $\{\rho_x(m)\}_{x\in X}$ is a measurable for all $m \in M$. If we also have a measure, we can construct the direct integral representation by letting M act on the direct integral in the natural way.

Associated to $\hat{M}$ (with the Mackey Borel structure), we have an obvious choice of measurable family of Hilbert spaces. Furthermore, in case the Mackey Borel structure is standard (which it is in case $M = G$) there is a measurable field of representations $\{\rho_p\}$ acting on this measure field of Hilbert spaces such that $\rho_p \in p$ for all p ([Fol95], Lemma 7.39). This is tautological if we work with Irr_n but is no longer obvious if we work with $\hat{M}$.

A.10.3: The Plancherel theorem

In addition to Proposition A.16, there is another theorem which is rightly called the Plancherel theorem and which also specifies the existence of a certain measure known as the Plancherel measure. Recall that the *left (right*, resp.) *regular representation* of M is the unitary representation $L^2(M)$ with M acting on the left (right, resp.) and is denoted λ_M (ρ_M, resp.). In fact the $M \times M$ action on $L^2(M)$ corresponding to both left and right translation encapsulates both the left and right regular representations; this is called the *two-sided regular representation* and is denoted τ. The Plancherel theorem is concerned with decomposing this representation into a direct integral of representations.

Given a unitary representation $(\rho, \mathcal{H})$ of M, we can form the *contragredient representation* $(\rho^*, \mathcal{H}^*)$ by letting $\rho^*(m) = \rho(m^{-1})^T$. Given representations $(\rho_1, \mathcal{H}_1)$ of M_1 and $(\rho_2, \mathcal{H}_2)$ of M_2, we can form a representation of $M_1 \times M_2$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ via $(m_1, m_2).v_1 \otimes v_2 = (\rho_1(m_1).v_1) \otimes (\rho_2(m_2).v_2)$. If $M_1 = M_2 = M$, $\mathcal{H}_1 = \mathcal{H}_2^* = \mathcal{H}$, and $\rho_1 = \rho_2^* = \rho$, then $M \times M$ acts on $\mathcal{H} \otimes \mathcal{H}^*$ which is canonically isomorphic to the vector space of Hilbert-Schmidt operators on $\mathcal{H}$ via $v \otimes \alpha \mapsto (w \mapsto \alpha(w)\, v)$.

Let $f \in L^1(M)$ and $(\rho, \mathcal{H}_\rho) \in \hat{M}$. Recall that we define $\hat{f}(\rho) = \int_M f(x)\rho(x^{-1})dx$. This operator is Hilbert-Schmidt, hence can be identified with an element of $\mathcal{H}_\rho \otimes \mathcal{H}_{\rho^*}$. With respect to the measurable family of Hilbert spaces over $\hat{M}$ discussed in A.10.2, $\hat{f}(\rho)$ is a measurable field of operators called the *Fourier transform* of f. Let $\mathcal{J}^1 = L^1(M) \cap L^2(M)$ and let $\mathcal{J}^2$ be the linear span of $f * g$ with $f, g \in \mathcal{J}^1$.

Theorem A.17 (Plancherel theorem; [Fol95], Theorem 7.44)**.** *Suppose M is a second countable, unimodular, type I group. There is a measure ν on $\hat{M}$, uniquely determined once the Haar measure on M is fixed, with the following properties. The Fourier transform $f \to \hat{f}$ maps*

$$\mathcal{J}^1 \to \int_{\hat{G}}^{\oplus} \mathcal{H}_\rho \otimes \mathcal{H}_{\rho^*}, d\nu(\rho)$$

and it extends to a unitary isomorphism

$$L^2(M) \to \int_{\hat{G}}^{\oplus} \mathcal{H}_\rho \otimes \mathcal{H}_{\rho^*} d\nu(\rho)$$

that intertwines the two-sided regular representation τ with $\int_{\hat{G}}^{\oplus} \rho \otimes \rho^* d\nu(\rho)$.

For $f, g \in \mathcal{J}^1$ one has the Parseval formula

$$\int_M f(x)g(\bar{x})dx = \int_{\hat{M}} \mathrm{Tr}[\hat{f}(\rho)\hat{g}(\rho^*)]d\nu(\rho),$$

and for $h \in \mathcal{J}^2$ one has the Fourier inversion formula

$$h(x) = \int_{\hat{M}} \mathrm{Tr}[\rho^*(x)\hat{h}(\rho)]d\nu(\rho).$$

A.10.4: The relationship between the spherical Plancherel measure and the Plancherel measure

Proposition A.18. *The measure μ in Proposition A.16 is the same as the restriction of ν in Theorem A.17 to the set $\Omega^+ \subset \hat{G}$.*

Proof. Suppose $f \in H(G, K)$. We shall compute its image under the Fourier transform as in Theorem A.17. Let $(\rho, \mathcal{H}) \in \hat{G}$. First off we observe that $\hat{f}(\rho)$ projects onto the K-invariant subspace of $\mathcal{H}$. If we let $v \in \mathcal{H}$ and $k \in K$ then

$$\begin{aligned} k.(\hat{f}(\rho).v) &= \int_G f(x)\rho(kx^{-1}).vdx \\ &= \int_G f(ku)\rho(u^{-1}).vdu \\ &= \int_G f(u)\rho(u^{-1}).vdu \\ &= \hat{f}(\rho).v. \end{aligned}$$

We know that all elements in $\hat{G}$ have their K-fixed subspace of dimension either 0 or 1 depending on whether it is class 1 (see Proposition A.5). If its dimension is 0, then $\hat{f}(\rho) = 0$. Otherwise $\hat{f}(\rho)$ is the composition of projection onto the one-dimensional K-invariant subspace and some scaling operator. We now compute this scaling factor.

Let v_ρ be a unit K-fixed vector in $\mathcal{H}$. Then

$$\begin{aligned} \langle v_\rho, \hat{f}(\rho).v_\rho\rangle &= \int_G f(x)\langle v_\rho, \rho(x^{-1}).v_\rho\rangle \\ &= \int_G f(x)\omega_\rho(x^{-1})dx. \end{aligned}$$

Thus we see that the underlying scaling operator is $\hat{f}(\omega_\rho)$.

From Theorem A.17 and Proposition A.16 we know now that

$$\begin{aligned} \int_G |f(x)|^2 dx &= \int_{\hat{G}} \mathrm{Tr}[\hat{f}(\rho)\hat{f}(\rho^*)]d\nu(\rho) \\ &= \int_{\Omega^+} |\hat{f}(\omega_\rho)|^2 d\nu(\rho). \end{aligned}$$

Since the measure μ in Proposition A.16 is uniquely characterized by satisfying the above identity, and ν restricted to Ω^+ also satisfies the identity, we get that they must be equal. □

It is also worth noting that the image of $L^2(G/K)$, thought of as a subspace of $L^2(G)$, under the Fourier transform is equal to

$$\int_{\Omega^+}^{\oplus} v_\rho \otimes \mathcal{H}_\rho^* d\nu(\rho).$$

The proof of the above proposition also shows that $L^2(K\backslash G/K)$ is isomorphic to

$$\int_{\Omega^+}^{\oplus} v_\rho \otimes v_\rho^* d\nu(\rho).$$

A.11: Tempered representations, integrability exponents, and property (T)

A.11.1: Tempered representations and the tempered spectrum

An irreducible unitary representation $(\rho, \mathcal{H})$ is called *tempered* if it is weakly contained in the regular representation. In light of the Plancherel theorem we see that this is equivalent to ρ lying in the support of the Plancherel measure on $\hat{M}$.

In case $G = \mathrm{PGL}(d, F)$, we have by Proposition A.6 and (A.9.1) that the tempered representations which are also spherical are exactly those appearing as (the unitarization of) the unique spherical subquotient of I_χ with χ a unitary character. This defines a distinguished sublocus of Ω^+ whose points correspond to these tempered class 1 representations. We call this sublocus the *tempered spectrum*, denoted Ω^+_{temp}; explicitly points in Ω^+_{temp} are of the form $(q^{s_1}, \ldots, q^{s_d})$ with $q^{s_j} \in S^1$ and $\Pi q^{s_j} = 1$.

A.11.2: Integrability exponents

In [CHH88], an alternative criterion for a representation to be tempered is given:

Proposition A.19. *Let ρ be a unitary representation of a locally compact group M on a Hilbert space $\mathcal{H}$. Let v be a cyclic vector for $(\rho, \mathcal{H})$. If the diagonal matrix coefficient $\langle v, \rho(m).v\rangle$ is in $L^{2+\varepsilon}(M)$ for every $\varepsilon > 0$, then $(\rho, \mathcal{H})$ is tempered.*

This suggests studying which L^p-space matrix coefficients of unitary representations lie in. Given a unitary representation $(\rho, \mathcal{H})$, we define its *integrability exponent* $q(\rho)$ to be

$$\inf\{p : \langle v, \rho(m).v\rangle \in L^p(M) \text{ for } v \in V \subset \mathcal{H} \text{ a dense subset}\}.$$

Proposition A.20 ([CHH88])**.** *Suppose M is a semisimple algebraic group over a local field. Let $(\rho, \mathcal{H})$ be a unitary representation with integrability exponent $q(\rho) \leq 2$. Then $(\rho, \mathcal{H})$ is tempered.*

A topological group M is said to have *property (T)* if the trivial representation is an isolated point in $\hat{M}$ with respect to the Fell topology. We say that a unitary representation $(\rho, \mathcal{H})$ has a *spectral gap* if there exists $\varepsilon > 0$ and a compact subset $C \subset M$ such that for every unit vector $v \in \mathcal{H}$ and for every $m \in C$, we have $||\rho(m).v - v|| > \varepsilon$. Property (T) is in fact equivalent to all non-trivial representations having a spectral gap. For semisimple algebraic groups over local fields, having a spectral gap is equivalent to $q(\rho) < \infty$ ([GN10]). The group $G = \mathrm{PGL}(d, F)$ is known to have property (T) if $d \geq 3$.

Proposition A.21 ([GN10])**.** *Suppose M is a semisimple algebraic group over a local field and has property (T). Then there exists a $q_0 < \infty$ such that for all non-trivial $\rho \in \hat{M}$, we have $q(\rho) \leq q_0$.*

APPENDIX B

Coxeter Complexes and Buildings

B.1: Coxeter groups and Coxeter complexes

The material of this section is standard and may be found in Chapter 2 of [Ron89].

B.1.1: Coxeter groups

A *Coxeter group* W is any group which has a presentation of the form:

$$W = \langle r_i | r_i^2 = (r_i r_j)^{m_{ij}} = 1 \text{ for all } i, j \in I \rangle$$

where I is a finite set. The m_{ij}'s are also allowed to be ∞. The r_i's are called the *Coxeter generators*.

B.1.2: The Coxeter complex

Given a Coxeter group W, there is a way of constructing an associated cell complex. Each $w \in W$ defines a top-dimensional cell of dimension $|I| - 1$. The top-dimensional cells are called *chambers* (sometimes also referred to in the literature as *alcoves*). Two chambers are adjacent, i.e. share a codimension one face, if they differ by a Coxeter generator. More specifically, we say that chambers w_1 and w_2 are *i-adjacent* if $w_2 = w_1 r_i$. Codimension one faces are called *panels*. It is clear from the construction that W acts simply transitively on chambers in such a way that the I-valued adjacency relations are preserved (i.e. W acts by multiplication on the left) and that each chamber has exactly one neighbor of each type for $i \in I$. This cell complex is called the *Coxeter complex* associated to W.

B.1.3: Coloring of the vertices

If we pick a coloring of the vertices of some fixed chamber, we can use this coloring and the W-action to color each vertex one of $|I|$ colors such that W preserves the coloring.

B.1.4: The W-valued metric

Let X be the Coxeter complex associated to W. There is a *W-valued metric* $d_W(\cdot,\cdot)$ on chambers: given two chambers represented by elements $w_1, w_2 \in W$, we define $d_W(w_1, w_2) = w_1^{-1}w_2$. Then if two chambers are i-adjacenct, $d_W(w_1, w_2) = r_i$. Notice that d_W is preserved by the left W-action.

B.1.5: Galleries

A *gallery* from w_1 to w_2 is any sequence of chambers

$$(\mathfrak{c}_1 = w_1, \mathfrak{c}_2, \ldots, \mathfrak{c}_{k-1}, \mathfrak{c}_k = w_2)$$

such that $\mathfrak{c}_{j+1} = \mathfrak{c}_j r_{i_j}$, i.e. consecutive chambers are adjacent. Then in fact $r_{i_1} r_{i_2} \ldots r_{i_{k-1}} = d_W(w_1, w_2)$ for all galleries. A *minimal gallery* is a gallery of minimal length. Minimal galleries are in some sense "combinatorial chamber geodesics".

B.1.6: Convex sets in Coxeter complexes

We say that a subset Y of chambers of X is *combinatorially convex* if every minimal gallery between any two elements in Y lies entirely in Y. We define the *combinatorial convex hull* of two chambers $\mathfrak{c}_1, \mathfrak{c}_2$ to be the intersection of all combinatorially convex subsets of X which contain $\mathfrak{c}_1$ and $\mathfrak{c}_2$.

B.1.7: The Coxeter length of a word

We may convert $d_W(\cdot,\cdot)$ to an actual metric $\tilde{d}(\cdot,\cdot)$ between chambers by simply recording the length of any minimal gallery. This also allows us to define the *length* of elements in $w \in W$ which we denote by $\ell(w) := \tilde{d}(1, w)$.

B.1.8: Reflections and roots

A *reflection* $r \in W$ is any conjugate of a Coxeter generator. Its *wall* M_r is all simplices fixed by r. We can define an equivalence relation on chambers by specifying two chambers to be equivalent if any (or equivalently all) galleries connecting them never cross M_r (that this is well-defined is standard). This partitions chambers into two distinct sets. Each such set is called a *root* (also referred to as a *half-apartment*).

B.2: Spherical Coxeter complexes

Material for this section may also be found in Chapter 2 of [Ron89].

B.2.1: Geometric realization of a spherical Coxeter group

Let W be a finite Coxeter group. Then W has a faithful representation as isometries of $\mathbb{R}^d$ (for some d) which fix the origin and such that each Coxeter generator corresponds to reflection across some hyperplane through the origin. Consider the collection of all hyperplanes obtained as the image under W of one of the generating hyperplanes. This partitions $\mathbb{R}^d$ into polyhedral cones called *Weyl chambers*. By restricting to the unit sphere, we may identify each Weyl chamber with a simplex in S^{d-1}, the $(d-1)$-sphere. These simplices give a triangulation of S^{d-1} which we may identify with the associated Coxeter complex. Such Coxeter groups and Coxeter chambers are called *spherical*.

Each reflection corresponds to the restriction to S^{d-1} of reflection across one of the hyperplanes, and the associated wall is the intersection of S^{d-1} with that hyperplane. The roots correspond to all chambers on a given side of one of these walls. There is a unique element of W of greatest length. Because of this each chamber has a unique *opposite chamber* which is obtained by multiplying on the right by this longest element.

B.2.2: The symmetric group as a spherical Coxeter group

An important class of spherical Coxeter groups are the symmetric groups $\mathfrak{S}_d$ with generators $(1\ 2), (2\ 3), \ldots, (d-1\ d)$. In fact all Weyl groups of roots systems are spherical Coxeter groups.

We shall in particular be interested in the case of $\mathfrak{S}_3$, in which case the associated Coxeter complex is a triangulation of the unit circle into six equal pieces.

B.3: Affine Coxeter complexes

Material for this section may be found in Chapter 9.1 of [Ron89].

B.3.1: Geometric realization of an affine Coxeter group

In other cases a Coxeter group W has a faithful representation as isometries of some d-dimensional Euclidean space E^d (with $|I| = d+1$) in such a way that each Coxeter generator corresponds to reflection across some affine hyperplane. If we again consider all hyperplanes obtained by the orbit under W of the generating hyperplanes, we end up with a tessellation, this time of E^d, into simplices. We may identify this tessellation with the associated Coxeter complex. We call the Coxeter group and the associated Coxeter complex *affine*.

Each reflection corresponds to reflection across one of the hyperplanes, and the associated wall is exactly that hyperplane. The roots correspond to all chambers on a given side of one of these walls.

The group W may hence be realized as a collection of affine isometries of a Euclidean space; this latter group, after choosing an origin, may be identified with $\mathbb{R}^d \rtimes \mathrm{O}(d)$.

B.3.2: The associated spherical Coxeter group W_0 and special vertices

The group $\mathbb{R}^d \rtimes \mathrm{O}(d)$ projects to $\mathrm{O}(d)$. The image of W inside of $\mathrm{O}(d)$ is a discrete, and hence finite, group denoted W_0. It is in fact a spherical Coxeter group generated by d Coxeter generators. Suppose $\alpha_1, \ldots, \alpha_d \in W$ are reflections whose images generate W_0. Then the intersection of the associated walls of the α_i's must intersect in a unique vertex. Any vertex that arises in this way is called a *special vertex*. Hence the special vertices have associated "residual" spherical Coxeter groups and in fact the link of a special vertex is isomorphic to the Coxeter complex associated to W_0.

B.3.3: Sectors

Suppose now p is a special vertex. Let $\mathfrak{c}$ be a chamber having p as one of its vertices. The panels of $\mathfrak{c}$ containing p determine d roots $r_1, \ldots, r_d$ each containing $\mathfrak{c}$. Their intersection $\mathcal{S}$ is called the *sector* based at vertex p with germ $\mathfrak{c}$. We may generate a copy of W_0 by reflection across the walls of $r_1, \ldots, r_d$. From this perspective we see that sectors are the same thing as the Weyl chambers for the spherical Coxeter group W_0 thought of as acting on the vector space E^d with origin at p. Thus sectors are in some sense "affine Weyl chambers."

B.3.4: Coxeter groups of type $\tilde{A}_d$

Particularly important examples of affine Coxeter groups are those of type $\tilde{A}_d$ which arise by reflections across the $d+1$ faces of a regular d-simplex sitting inside $\mathbb{R}^d$. The group W_0 is the Coxeter group of type A_d which is just an alternate name for $\mathfrak{S}_{d+1}$. For the associated Coxeter complex, all vertices are special.

B.4: Buildings

Material for this section may be found in [Bro89].

B.4.1: The axioms of a building

A *building* is a simplicial complex Δ which can be expressed as the union of subcomplexes Σ (called *apartments*) satisfying:

(1) Each apartment is a Coxeter complex.

(2) For any two simplices $\mathfrak{c}_1, \mathfrak{c}_2 \in \Delta$, there is an apartment Σ containing both of them.

(3) If Σ and Σ' are two apartments containing $\mathfrak{c}_1$ and $\mathfrak{c}_2$, then there is an isomorphism $\Sigma \to \Sigma'$ fixing $\mathfrak{c}_1$ and $\mathfrak{c}_2$ pointwise.

Each building has an associated Coxeter group W. We shall in particular be concerned with *thick buildings*, meaning that every codimension one face is contained in at least 3 chambers.

When the underlying Coxeter group is spherical (affine, resp.), the building is called *spherical* (*affine* or *Euclidean*, resp.).

B.4.2: Apartment systems

Any time we have a collection of apartments $\mathcal{C}$ whose union is Δ and whose elements Σ satisfy the above axioms, we call $\mathcal{C}$ a *system of apartments*. There is a unique maximal system of apartments which is referred to as the *complete system of apartments*.

B.4.3: The W-valued metric on chambers and coloring of vertices in Δ

If we fix a particular chamber in a particular apartment and color each of its vertices one of $|I|$ colors, then this coloring extends uniquely to all vertices of Δ using the W-action on each apartment. The isomorphisms $\Sigma \to \Sigma'$ in the above axioms can be taken to be color-preserving. The Coxeter group-valued metric $d_W(\cdot,\cdot)$ also extends to pairs of chambers in Δ.

B.5: The spherical building associated to $\mathrm{SL}(d, \mathbb{K})$

Material in this section may be found in [Ron89].

Given any field $\mathbb{K}$, we may construct a spherical building Π of type A_{d-1}. We define the vertices of Π to be the non-trivial subspaces of $\mathbb{K}^d$. A collection of vertices form a simplex if there is some ordering of the underlying vector spaces such that they formed a nested containment. From this we see that the top-dimensional simplices (i.e. chambers) correspond to full flags in $\mathbb{K}^d$. The apartments correspond to bases of $\mathbb{K}^d$, and the individual chambers in an apartment correspond to all ways of creating a full flag using elements in a given basis. That $\mathcal{B}$ satisfies the axioms of a (spherical) building (in particular that any two chambers are contained in an apartment) essentially follows from the Bruhat decomposition. We shall be particularly interested in the case when $\mathbb{K} = \mathbb{F}_q$, in which case Π is a finite simplicial complex.

B.6: The affine building associated to $\mathrm{SL}(d, F)$

Material in this section may be found in [Ron89].

B.6.1: The construction of $\mathcal{B}$

Now suppose F is a non-archimedean local field with ring of integers $\mathcal{O}$, uniformizer ϖ, and residue field of order q. Let $\mathcal{L}$ denote the set of $\mathcal{O}$-modules of rank d inside of F^d; elements of $\mathcal{L}$ are called *lattices*. We have a surjection $\mathrm{GL}(d, F) \twoheadrightarrow \mathcal{L}$ by taking the lattice spanned by the columns. Hence every $L \in \mathcal{L}$ can be represented by a matrix $M_L \in \mathrm{GL}(d, F)$.

We now form the quotient set $\mathcal{B}$ of $\mathcal{L}$ by declaring two elements L_1, L_2 of $\mathcal{L}$ equivalent if for some $c \in F^\times$ we have $cL_1 = L_2$. We can furthermore put a simplicial complex structure on $\mathcal{B}$ by declaring a collection of elements of $\{[L_1], \dots, [L_m]\}$ of $\mathcal{B}$ to form a simplex if for some ordering of these elements and some choice of representative $L_i \in [L_i]$, we have

$$L_{i_1} \subsetneq L_{i_2} \subsetneq \cdots \subsetneq L_{i_m} \subsetneq \varpi L_{i_1}.$$

Since $L_{i_1}/(\varpi L_{i_1}) \simeq \mathbb{F}_q^d$, we see that under this identification we can associate to each L_i a subspace of $\mathbb{F}_q^d$. The largest any simplex can be is thus a set of d elements (a $d-1$-dimensional simplex) corresponding to some full flag in $\mathbb{F}_q^d$.

The resulting object is an affine building with associated Coxeter group of type $\tilde{A}_{d-1}$. It is known as the *Bruhat-Tits building* associated to $\mathrm{SL}(d, F)$. In fact, given any reductive algebraic group M over a non-archimedean local field, one may construct its associated Bruhat-Tits building. In the case that M is simply connected, one may construct it in general using the theory of BN-pairs; however, the case of $\mathrm{SL}(d, F)$ has a more intuitive construction which we have described above. The Bruhat-Tits building only depends on the isogeny class of M, so it makes sense to also call the above $\mathcal{B}$ the Bruhat-Tits building associated to $\mathrm{PGL}(d, F)$. That $\mathcal{B}$ satisfies the building axioms essentially follows from the Iwahori decomposition.

B.6.2: The vertices of $\mathcal{B}$ as the set G/K

Let $G = \mathrm{PGL}(d, F)$ and $K = \mathrm{PGL}(d, \mathcal{O})$. Notice that G acts transitively on the vertices of $\mathcal{B}$, and the stabilizer of the lattice represented by the identity matrix has stabilizer equal to K. Hence we can identify the vertices of $\mathcal{B}$ with G/K.

B.6.3: The coloring of the vertices of $\mathcal{B}$

We may color the vertices as follows: each lattice L may be represented (non-uniquely) by some element in $M_L \in \mathrm{GL}(d, F)$. We then color L with the valuation of $\det(M_L) \bmod d$ (hence the color takes values in $\mathbb{Z}/d\mathbb{Z}$). Each chamber has exactly one vertex of each color. We can also color directed edges by taking the difference in colors between the tip and the base of the directed edge mod d.

The group $\mathrm{SL}(d, F)$ acts on $\mathcal{B}$ simplicially and in such a way that coloring of vertices is preserved. In fact it acts transitively on all vertices of a given color. On the other hand, the group $\mathrm{PGL}(d, F)$ also acts on $\mathcal{B}$ simplicially. It acts transitively on vertices but it does not preserve the coloring of the vertices. However it does preserve the coloring of directed edges.

B.6.4: Apartments in $\mathcal{B}$

Apartments in $\mathcal{B}$ correspond to maximal split tori in $G = \mathrm{PGL}(d, F)$. As long as F is complete with respect to the valuation, this collection of apartments is the complete apartment system. We define a *sector* in $\mathcal{B}$ to be any sector in any apartment of $\mathcal{B}$.

Recall that $T < G$ consists of all diagonal matrices, $A < T$ consists of those whose diagonal entries are powers of ϖ, and $A^+ \subset A$ consists of those for which the powers of ϖ along the diagonal are weakly decreasing. The subgroup T is a maximal split torus defining an apartment which we call the standard apartment $\mathcal{X}$. Concretely, this apartment consists of all lattices which may be expressed as the $\mathcal{O}$-span of $\{\varpi^{\lambda_1} e_1, \ldots, \varpi^{\lambda_d} e_d\}$. There is a bijection between the vertices of $\mathcal{X}$ and elements of A. Furthermore the image of A^+ defines a sector $\mathcal{W}$ in $\mathcal{X}$ based at the vertex o corresponding to the lattice $\mathcal{O}e_1 + \cdots + \mathcal{O}e_d$ and with germ equal to the chamber corresponding to the chain

$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \supseteq \begin{bmatrix} \varpi & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \supseteq \begin{bmatrix} \varpi & & & & \\ & \varpi & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \supseteq \cdots \supseteq \begin{bmatrix} \varpi & & & & \\ & \varpi & & & \\ & & \ddots & & \\ & & & \varpi & \\ & & & & 1 \end{bmatrix}.$$

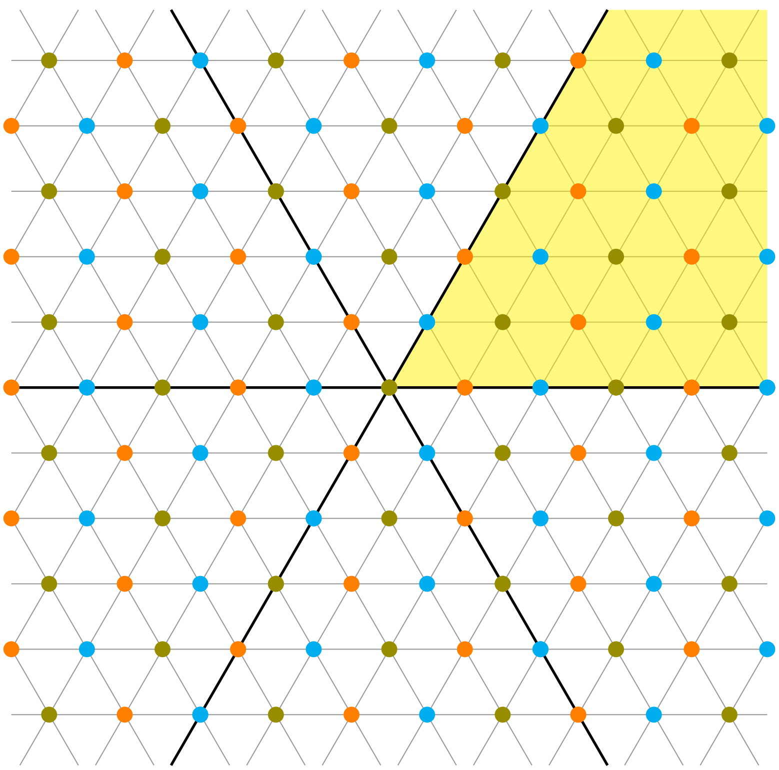


Figure 22: This shows the Coxeter complex of type $\tilde{A}_2$, or equivalently an apartment in the Bruhat-Tits building associated to $\mathrm{PGL}(3, F)$. We may color the vertices one of three colors; this coloring is preserved by the action of the underlying affine Coxeter group which acts simply transitively on chambers. Given a fixed vertex, we can partition the Coxeter complex into six Weyl chambers which are parametrized by elements in $\mathfrak{S}_3$.

B.6.5: The local spherical building of a vertex

The link of a vertex p in $\mathcal{B}$ is a spherical building which we call the *local spherical building at* p. In fact it is exactly the spherical building Π constructed in Appendix B.5 for the field $\mathbb{F}_q$.

B.6.6: The Weyl-chamber valued metric

We may use the Cartan decomposition to define a "Weyl chamber-valued metric" $d_{A^+}(\cdot,\cdot)$ between vertices in the Bruhat-Tits building $\mathcal{B}$. Given two vertices $v_1, v_2 \in \mathcal{B}$, choose representatives $x_1, x_2 \in G$ and define $d_{A^+}(v_1, v_2) = \lambda$ where $x_1^{-1}x_2 = k_1\varpi^\lambda k_2$ in the Cartan decomposition (with $\varpi^\lambda \in A^+$). It is invariant under multiplication by G (on the left).

B.6.7: The Euclidean metric on $\mathcal{B}$

Given two points in $\mathcal{B}$, we can find some apartment containing both of them and consider the straight line segment in that apartment connecting them. By measuring the length of that line segment we may define a metric $d(\cdot,\cdot)$ on the building. With respect to this metric the building is a CAT(0) space. In particular there is a unique geodesic connecting any two points. We normalize $d(\cdot,\cdot)$ so that the distance between two vertices in a given chamber is exactly 1.

B.6.8: The retraction onto an apartment centered at a chamber

Given any fixed chamber $\mathfrak{c}$ and a fixed apartment Σ containing $\mathfrak{c}$, there is a unique simplicial retraction $\rho_{\mathfrak{c},\Sigma}$ which maps $\mathcal{B}$ onto Σ in such a way that for any chamber $\mathfrak{d} \in \mathcal{B}$ we have $d_W(\mathfrak{c},\mathfrak{d}) = d_W(\mathfrak{c}, \rho_{\mathfrak{c},\Sigma}(\mathfrak{d}))$. However, in general $\rho_{\mathfrak{c},\Sigma}$ has the property that it does not increase distances between chambers (and it preserves the distance between $\mathfrak{c}$ and any other chamber).

B.6.9: The retraction onto a sector based at a vertex

Suppose p is a special vertex and $\mathcal{S}$ is some sector based at p. There is a unique simplicial retraction $R_{p,\mathcal{S}}$ which maps $\mathcal{B}$ onto $\mathcal{S}$ in such a way that for any vertex $x \in \mathcal{B}$ we have $d_{A^+}(p,x) = d_{A^+}(p, R_{p,\mathcal{S}}(x))$. In general, $R_{p,\mathcal{S}}$ has the property that it does not increase distances between vertices (and it preserves the distance between p and any other vertex).

B.7: The $H(G, K)$-action on $\mathcal{B}$

B.7.1: Geometric interpretation of convolving with c_λ

A function f on the vertices of the building is the same thing as a $(1, K)$-invariant function on G. Elements in $H(G, K)$ act on these functions. Recall that a vector space basis for $H(G, K)$ is the functions c_λ, which is the indicator function for $K\varpi^\lambda K$ where $\varpi^\lambda = \mathrm{diag}(\varpi^{\lambda_1}, \ldots, \varpi^{\lambda_d})$. Given $\lambda = (\lambda_1, \ldots, \lambda_d)$ with $\lambda_1 \geq \cdots \geq \lambda_d = 0$, let $\lambda^\vee$ be:

$$\lambda^\vee := (\lambda_1 - \lambda_d, \lambda_1 - \lambda_{d-1}, \ldots, \lambda_1 - \lambda_2, 0). \tag{B.7.1}$$

Notice then that $(K\varpi^\lambda K)^{-1} = K\varpi^{\lambda^\vee} K$. Furthermore we have

$$d_{A^+}(x, y) = \lambda \iff d_{A^+}(y, x) = \lambda^\vee.$$

We claim that

$$f * c_\lambda(x) = \sum_{yK \text{ such that } d_{A^+}(xK, yK) = \lambda^\vee} f(yK). \tag{B.7.2}$$

First off note that $K\varpi^\lambda K$ is the disjoint union of finitely many cosets $\{Kz_i\}$. Then $K\varpi^{\lambda^\vee} K$ consists of the cosets $\{z_i^{-1}K\}$. Notice that $d_{A^+}(1K, xK) = \lambda^\vee$ if and only if $xK \in K\varpi^{\lambda^\vee} K$. Hence the points $\{z_i^{-1}K\} \subset G/K$ exhaust the points in G/K such that $d_{A^+}(1K, xK) = \lambda^\vee$. Since G preserves $d_{A^+}(\cdot, \cdot)$, the points y such that $d_{A^+}(xK, yK) = \lambda^\vee$ for a fixed xK are exactly $\{xz_i^{-1}K\}$. So we have

$$\begin{aligned} f * c_\lambda(x) &= \int_G f(y) c_\lambda(y^{-1}x) dy \\ &= \sum_i \int_K f(xz_i^{-1}k) dk \\ &= \sum_i f(xz_i^{-1}). \end{aligned}$$

B.7.2: Coloring on directed edges

We may color every directed edge in $\mathcal{B}$ with colors in $\{1, 2, \ldots, d-1\}$ depending on the difference in color between the end point and the starting point (this takes values in $(\mathbb{Z}/d\mathbb{Z})^\times$). Define $\mathcal{A}_k$ to be the operator corresponding to summing up over all adjacent vertices which can be reached by going along a directed edge of color k. These operators generate $H(G,K)$; in terms of the standard basis for $H(G,K)$, this operator corresponds to the element $c_{(1,1,\ldots,1,0,\ldots,0)}$ where the associated partition has $d-k$ ones.

B.7.3: The $H(G,K)$-action on $L^2(\Gamma\backslash G/K)$

Now suppose Γ is a cocompact torsionfree lattice inside of G. Then Γ acts on G/K (on the left), and we may quotient the building by this action to get a finite simplicial complex. Functions on this simplicial complex may be identified with (Γ, K)-invariant functions on G. This space is preserved under the $H(G,K)$-action described above (convolution on the right). There's a natural Hilbert space structure on functions on the vertices of the quotient of the building by declaring functions supported at different vertices to be orthogonal. This identifies this space with $L^2(\Gamma\backslash G/K)$. The spherical Hecke algebra acts by normal operators since $(\cdot * c_\lambda)^* = (\cdot * c_{\lambda^\vee})$ and these two operators commute (since $H(G,K)$ is commutative). Thus $L^2(\Gamma\backslash G/K)$ has an orthogonal basis of joint eigenfunctions of $H(G,K)$. By Proposition A.15 we know that these correspond to the class 1 representations that show up in the representation $L^2(\Gamma\backslash G)$.

B.8: The Bruhat-Tits tree for $\mathrm{PGL}(2,F)$

In the case of $d = 2$, the Bruhat-Tits building is the infinite regular $(q+1)$-regular tree. The spherical Hecke algebra is generated (as an algebra) by one element, namely $c_{(1,0)}$. Geometrically this corresponds to the adjacency operator on the tree. If we quotient by a cocompact torsionfree lattice, we get a finite $(q+1)$-regular graph. Studying the $\mathbb{C}$-algebra homomorphisms of $H(G,K)$ in this case reduces to studying eigenvalues of the adjacency operator (or the Laplacian operator). The image of A gives a bi-infinite path through the tree, and the image of A^+ gives a one-sided infinite path.

B.9: Specific computations the building associated to $\mathrm{PGL}(3,F)$

B.9.1: Coordinates on the standard apartment $\mathcal{X}$

We now focus on the case of $d = 3$. Let $\mathcal{X}$ denote the standard apartment. We can identify A with the vertices of $\mathcal{X}$; on the other hand we may identify A with lattice points for a certain lattice $\Lambda \subset \mathfrak{a}$ (see Chapter II.1.2). Explicitly,

$$\begin{aligned}\Lambda &= \{(x_1, x_2, x_3) : x_1 - x_2 \in \mathbb{Z} \text{ and } x_1 - x_3 \in \mathbb{Z} \text{ and } x_1 + x_2 + x_3 = 0\} \\ &= \mathbb{Z} \cdot \left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right) + \mathbb{Z} \cdot \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right).\end{aligned}$$

We shall refer to these coordinates as the $\mathfrak{a}$-*coordinates*.

There are other natural ways to coordinatize A and $\mathcal{X}$: to each element in A was can associate a tuple $\lambda = (\lambda_1, \lambda_2, 0)$ (i.e. shift all entries so that the last entry is zero) or we can write λ as a linear combination, $\lambda = r(1,0,0) + s(1,1,0)$, and assign it the coordinates (r, s). The vectors $(1,0,0)$ and $(1,1,0)$ give a basis for the cone corresponding to A^+. We shall call the first set of coordinates *partition coordinates* and the second set *cone coordinates*. We let $\preceq$ denote the partial ordering on $\mathfrak{a}$ defined in cone coordinates by:

$$(r_1, s_1) \preceq (r_2, s_2) \iff r_1 \leq r_2 \text{ and } s_1 \leq s_2.$$

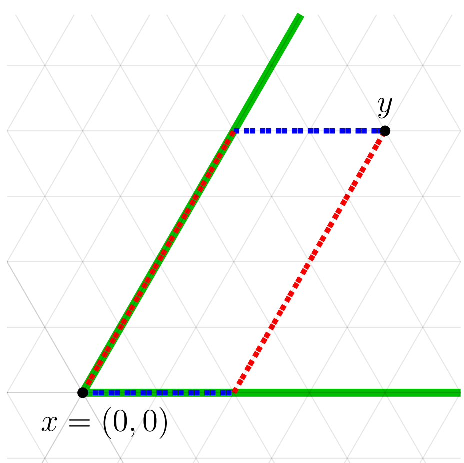


Figure 23: Suppose $x = (0,0)$ in cone coordinates. If we have $d_{A^+}(x, y) = (2,4)$ in cone coordinates, then we may reach y from x by going two steps in the direction of $(1,0)$ and 4 steps in the direction of $(0,1)$. In cone coordinates the vectors $(1,0)$ and $(0,1)$ generate the Weyl chamber $\mathfrak{a}^+$.

APPENDIX C

Polytopes and Brion's Formula

C.1: Convex polytopes

Suppose $\{\beta_i(x)\}_{i\in I}$ is a finite set of linear functionals on $\mathbb{R}^d$, and $\{b_i\}_{i\in I}$ is a finite set of real numbers; we call the elements b_i *cutoffs*. The set of solutions to the system of inequalities $\{\beta_i(x) \leq b_i\}_{i\in I}$ is called a *(convex) polytope*, denoted Q.

From the definition we see that Q is in fact the intersection of finitely many affine halfspaces. The interior of Q is those points satisfying $\{\beta_i(x) < b_i\}_{i\in I}$. The set of all points in Q along which some subset of the defining inequalities becomes equalities is called a *face* $\mathfrak{f}$. We can associate to $\mathfrak{f}$ the collection of functionals $\mathrm{Func}(\mathfrak{f})$ for which the associated inequality becomes an equality. A *vertex of* Q is a zero-dimensional face (hence consists of exactly one point) and an *edge* is a one-dimensional face.

Let $\mathfrak{f}$ be a face of Q. Let $\mathrm{Aff}(\mathfrak{f})$ denote the affine span of points in $\mathfrak{f}$. The *dimension* of $\mathfrak{f}$ is defined as the dimension of $\mathrm{Aff}(\mathfrak{f})$. The dimension of Q is the dimension of the affine space spanned by all points in Q. We shall always assume that we are working with d-dimensional polytopes (i.e. the dimension is equal to the dimension of the underlying Euclidean space; all polytopes have representations for which this is the case up to an appropriate definition for isomorphism of polytopes).

Proposition C.1. *Let B be the set of simultaneous solutions to $\beta_j(x) = b_j$ for all $\beta_j \in \mathrm{Func}(\mathfrak{f})$. Then $B = \mathrm{Aff}(\mathfrak{f})$.*

C.2: The type of a polytope and the cone of a vertex

C.2.1: The type of a polytope

By associating to each face $\mathfrak{f}$ the collection of functionals $\mathrm{Func}(\mathfrak{f})$, we obtain a subset of the power set of the functionals. We call this the *type of* Q. We shall later on work with families of polytopes for which the functionals remain constant but the cutoffs change. In such a setting it makes sense to talk about when two polytopes in the family have the same type.

Suppose $\mathcal{Q}$ is a family of polytopes all of the same type. If $Q \in \mathcal{Q}$, and $\mathfrak{f}$ is a face of Q, then there is an analogous face $\mathfrak{f}'$ for each $Q' \in \mathcal{Q}$. Let v be a vertex of Q and e an edge associated to v.

Proposition C.2. *The direction of the ray based at v in the direction of e is completely determined by the type of Q.*

Proof. By Proposition C.1, we can pick out from $\mathrm{Func}(e)$ a subset of $d-1$ linearly independent functionals, call them $\beta_1, \ldots, \beta_{d-1}$, such that $\mathrm{Aff}(e)$ is the set of solutions to $\beta_i(x) = b_i$ for $1 \leq i \leq d-1$. We then can realize $\mathrm{Aff}(e)$ as the solutions to $Mx = b$ for M an $(d-1) \times d$ matrix of rank $d-1$. The solutions to this may be described as any particular solution plus vectors in the kernel of M. However we can take the same matrix M for all polytopes of the same type.

All that is left to do now is orient this line to get a ray. Choose some $\beta_j \in \mathrm{Func}(v)$ which is linearly independent from the $\beta_1, \ldots, \beta_{d-1}$ above. We orient the ray so it points in the direction so that motion in this direction decreases the value of the functional β_j. □

C.2.2: The cone of a vertex

Suppose we translate Q so that v is moved to the origin. Let $\mathrm{Rays}_Q(v)$ be the collection of rays associated to v now thought of as based at the origin. If we take the conical span of $\mathrm{Rays}_Q(v)$ we get a polyhedral cone which we shall denote by $\mathrm{Cone}_Q(v)$. The main purpose of the above discussion is the following conclusion: $\mathrm{Cone}_Q(v)$ only depends on the type of Q and not on Q itself!

A cone is called *simplicial* if it has exactly d generators. A polytope is called simplicial if the cone at every vertex is simplicial.

C.3: Brion's formula

C.3.1: Lattice polytopes

Now suppose Λ is a lattice inside $\mathbb{R}^d$. Suppose Q is a polytope such that all vertices are in Λ; such polytope are called *lattice polytopes*. Let $Q^\Lambda := Q \cap \Lambda$. Consider $C := \mathrm{Cone}_Q(v)$ for some vertex v. Then to each ray $r \in \mathrm{Rays}_Q(v)$, we can associate the smallest non-zero element in Λ on the ray. We will call this the *coprime generator of the ray*. Let $\mathrm{Coprime}(C)$ denote the union of all coprime generators for all rays of C. If C is simplicial, we let $\mathrm{PP}(C)$ be the parallelpiped formed by taking $[0,1)\cdot r_1 + \ldots [0,1)\cdot r_d$ for coprime generators r_k.

C.3.2: Summing an exponential function over a polyhedral cone

Suppose C is a polyhedral cone based at the origin such that all of its extremal rays intersect Λ. Let C^* be the polar cone of C, that is $C^* = \{x \in (\mathbb{R}^n)^* : (x,y) \leq 0 \text{ for all } y \in C\}$. Let α be a (possibly $\mathbb{C}$-valued) functional. Define

$$\sigma(C;\alpha) := \sum_{\gamma \in C\cap\Lambda} q^{(\alpha,\gamma)}.$$

This quantity converges if $\mathrm{Re}(\alpha)$ is in the interior of C^*.

Suppose now that C is simplicial. Then using power series it is easy to see that

$$\sigma(C;\alpha) = \left(\sum_{\beta\in\mathrm{PP}(C)\cap\Lambda} q^{(\alpha,\beta)}\right)\left(\prod_{r\in\mathrm{Coprime}(C)} \frac{1}{1-q^{(\alpha,r)}}\right). \tag{C.3.1}$$

If C is not simplicial, then there exists a partition of C into simplicial cones such that each simplicial cone has its ray generators among the extremal rays of C. This implies that in all cases

$$\sigma(C;\alpha) = \frac{R(\alpha)}{\prod_{r\in\mathrm{Coprime}(C)} 1-q^{(\alpha,r)}},$$

where $R(\alpha)$ is a Laurent polynomial in q^{α_i} (where $\alpha = (\alpha_1,\ldots,\alpha_n)$). Hence the whole expression is a rational function in these variables.

C.3.3: Brion's formula

Theorem C.3 (Brion's formula [Bri88]; see also [Bar93] Theorem 4.5)**.** *Suppose Λ is a lattice and Q is a polytope with vertices $\mathcal{V}$, and $\mathcal{V} \subset \Lambda$. Let α be a functional for which no ray* *associated to any vertex is in its kernel (i.e. α is not orthogonal to any face of Q). Then,*

$$\sum_{\gamma \in Q^{\Lambda}} q^{(\alpha,\gamma)} = \sum_{v \in \mathcal{V}} \sigma(\mathrm{Cone}_Q(v);\alpha) \cdot q^{(\alpha,v)}. \tag{C.3.2}$$

Notice that each term in the sum is a product of two terms. The first term depends only on the structure of the cone at each vertex (and on Λ and α) and the second term is purely exponential.

C.3.4: Brion's formula as a generalization of geometric series

One may think of Brion's formula as a vast generalization of the geometric series formula. First we can rewrite a geometric series as:

$$1 + x + \cdots + x^n = q^{\log_q(x)\cdot 0} + q^{\log_q(x)\cdot 1} + \cdots + q^{\log_q(x)\cdot n}.$$

Hence if we let α be $\log_q(x) \in \mathbb{R}$ (treating x as a constant), let $\Lambda = \mathbb{Z}$, and let $Q = [0, n]$, then the above is exactly of the form of the left hand side of Brion's formula. Q has two vertices, 0 and n. At 0, the only coprime ray generator is 1, and at n the only coprime ray generator is -1. The only lattice point in $\mathrm{PP}(0)$ is 0 and the only lattice point in $\mathrm{PP}(n)$ is 0. Hence the right hand side of Brion's formula gives

$$\frac{1}{1 - q^{\log_q(x)}} + \frac{1}{1 - q^{-\log_q(x)}} \cdot q^{\log_q(x)\cdot n} = \frac{1}{1-x} + \frac{x^n}{1 - x^{-1}} = \frac{1 - x^{n+1}}{1-x}.$$

C.3.5: The asymptotics of Brion's formula

One important takeaway of Brion's formula is that the left hand side of (C.3.2) is dominated by its largest term. More specifically, if we fix Q and then dilate it by integer multiples, then the cone and coprime ray generators at each vertex stay the same; all such polytopes have the same type. The nondegeneracy assumption on α implies that there is a unique vertex $v^{\dagger}$ maximizing the dot product with α (if we assume that α is an $\mathbb{R}$-valued functional). Then the dominating term on the left hand side of (C.3.2) is $q^{(\alpha, n\cdot v^{\dagger})}$. On the right hand side of (C.3.2) we always get something of the form $\sum_v C(v) q^{(\alpha, n\cdot v)}$ where $C(v)$ does not depend on n. Hence as $n \to \infty$, the right hand side of (C.3.2) grows like $q^{(\alpha, n\cdot v^{\dagger})}$ as well.

C.4: Motivation for degenerate Brion's formula

We shall ultimately also be interested in the case when α is orthogonal to some face of Q. We call this the *degenerate case*. We would like a formula which allows us to understand the growth rate of the right hand side as some parameter in a family of polytopes goes to infinity. In such a case we will have potentially an entire face $\mathfrak{f}$ of Q for which on $\mathfrak{f}\cap\Lambda$ the dot product with α is maximized. Call this maximal dot product value M. Then heuristically we should expect the growth to be something of the form $\mathrm{Area}(\mathfrak{f}) \cdot q^M$.

An illustrative example is the polytope $Q = [0,1] \times [0,1]$, the lattice $\Lambda = \mathbb{Z}^2$, and the functional $\alpha = (0, \log_q(x))$ for some value of $x > 1$. Then Q can be defined by $\{x_1 \geq 0; x_2 \geq 0; x_1 \leq 1; x_2 \leq 1\}$. The dot product with α is maximized at both $(0,1)$ and $(1,1)$. Let $Q_{m,n}$ be the polytope defined by $\{x_1 \geq 0; x_2 \geq 0; x_1 \leq m; x_2 \leq n\}$. All $Q_{m,n}$ have the same type. It is easy to calculate that

$$\sum_{\gamma \in Q_{m,n}^{\Lambda}} q^{(\alpha,\gamma)} = (m+1) \cdot \frac{1 - x^{n+1}}{1 - x}.$$

This grows like $m \cdot x^n = m \cdot q^{\log_q(x) \cdot n}$ as $m, n \to \infty$, matching our above heuristic (the length of the side in $Q_{m,n}$ along which the dot product with α is maximized is m, and the value of that dot product is $\log_q(x) \cdot n$).

C.5: Degenerate Brion's formula

C.5.1: Degeneracy of a vertex

Suppose α is a functional. For each $v \in Q$ define the *degeneracy of v with respect to α* as the dimension of the span of the rays which are orthogonal to α. We say that a vertex is *good* if its degeneracy is zero; other it is *bad*. We say that a ray at a bad vertex is bad if it is orthogonal to α.

C.5.2: Degenerate Brion's formula

We shall now derive a degenerate form of Brion's formula. It is likely that some form of this already exists in the literature but we have been unable to find it. It is also likely that more elegant formulations along the same lines exist but for our purposes the below suffices.

Proposition C.4. *Suppose Λ is a lattice and $\mathcal{Q}(b_1, \ldots, b_\ell)$ is a family of lattice polytopes of the same type defined by $\beta_i(x) \leq b_i$. Let α be some functional. For all Q in this family*

$$\sum_{\gamma \in Q^\Lambda} q^{(\alpha,\gamma)} = \sum_{v \in \mathcal{V}} R_v(b_1, \ldots, b_\ell) q^{(\alpha,v)},$$

where each R_v is a polynomial in the b_j whose degree is bounded by greatest degeneracy of any vertex in V (which is at least the largest dimension of any face orthogonal to α) and which is degree 0 if v has degeneracy 0.

Proof. Let τ be any functional which is not orthogonal to any edge in Q. Theorem C.3 allows us to express for any ε small enough

$$\sum_{\gamma \in Q^\Lambda} q^{(\alpha+\varepsilon\cdot\tau,\gamma)} = \sum_{v} \sigma(\mathrm{Cone}_Q(v); \alpha + \varepsilon \cdot \tau) q^{(\alpha+\varepsilon\cdot\tau,v)}. \tag{C.5.1}$$

We shall further separate out the sum on the left hand side into the sum over good and bad vertices. The terms coming from the good vertices define holomorphic functions in ε near $\varepsilon = 0$. When $\varepsilon = 0$, the term corresponding to a good vertex v is exactly

$$\sigma(\mathrm{Cone}_Q(v); \alpha) q^{(\alpha,v)}.$$

The coefficient of $q^{(\alpha,v)}$ only depends on the type of Q, and hence is a non-zero constant when viewed as a function of the b_j's.

The left hand sum in (C.5.1) is also clearly a holomorphic function in ε as it is a finite sum of exponential functions. Consequently, the sum over the bad vertices also has a holomorphic extension to $\varepsilon = 0$.

Recall that $\sigma(\mathrm{Cone}_Q(v); \alpha + \varepsilon \cdot \tau)$ corresponds to the (analytic continuation) of the sum of $q^{(\alpha+\varepsilon\cdot\tau,\lambda)}$ over $\lambda \in \Lambda \cap \mathrm{Cone}_Q(v)$. We may partition this cone into simplicial cones, then partition the intersection of two of these cones into simplicial cones (of smaller dimension), and so on. In this way, by using inclusion-exclusion, we can express the sum of $q^{(\alpha+\varepsilon\cdot\tau,\lambda)}$ over the lattice points of $\mathrm{Cone}_Q(v)$ as a weighted sum of $q^{(\alpha+\varepsilon\cdot\tau,\lambda)}$ over lattice points in simplicial subcones, each of whose rays are contained among the rays of $\mathrm{Cone}_Q(v)$.

Suppose now v is a bad vertex and we have partitioned $\mathrm{Cone}_Q(v)$ into simplicial cones C

as described above. We then have

$$\sigma(\mathrm{Cone}_Q(v);\alpha+\varepsilon\cdot\tau)=\sum_{(v,C)\text{ with }C\text{ simplicial}} B_{(C,v)}\cdot\frac{\mathrm{PP}(C;\alpha+\varepsilon\cdot\tau)}{\prod_{\text{rays } r \text{ of C}} 1-q^{(\alpha+\varepsilon\cdot\tau,r)}}$$

for some constants $B_{(C,v)}$.

We now analyze the following sum:

$$\sum_{(v,C)\text{ s.t. } v\text{ bad, } C\text{ simplicial}} B_{(C,v)}\cdot\frac{\mathrm{PP}(C;\alpha+\varepsilon\cdot\tau)}{\prod_{\text{rays } r \text{ of C}} 1-q^{(\alpha+\varepsilon\cdot\tau,r)}}q^{(\alpha+\varepsilon\cdot\tau,v)}. \tag{C.5.2}$$

This is simply a rewriting of the right hand side of (C.5.1) with the terms correspond to good vertices removed.

Let v^{bad} be the vertex whose associated cone C^{bad} contains the greatest number of bad rays. Let BR denote these corresponding bad rays. Let k be the cardinality of BR. Notice then that k is at most the maximal degeneracy of any vertex. We shall multiply (C.5.2) by a "fancy one" given by the expression

$$\frac{\prod_{t\in\mathrm{BR}}\left(1-q^{(\alpha+\varepsilon\cdot\tau,t)}\right)}{\prod_{t\in\mathrm{BR}}\left(1-q^{(\alpha+\varepsilon\cdot\tau,t)}\right)},$$

which has a removable singularity at $\varepsilon=0$. Let

$$\nu(v,C;\alpha,\varepsilon,\tau):=B_{(C,v)}\cdot\mathrm{PP}(C;\alpha+\varepsilon\cdot\tau)q^{(\alpha+\varepsilon\cdot\tau,v)}.$$

We thus obtain

$$\frac{\sum_{(v,C)\text{ s.t. } v\text{ bad}}\left(\nu(v,C;\alpha,\varepsilon,\tau)\prod_{t\in\mathrm{BR}}\left(1-q^{(\alpha+\varepsilon\cdot\tau,t)}\right)\prod_{r\in\mathrm{Rays}(C)}\left(1-q^{(\alpha+\varepsilon\cdot\tau,r)}\right)^{-1}\right)}{\prod_{t\in\mathrm{BR}}\left(1-q^{(\alpha+\varepsilon\cdot\tau,t)}\right)}. \tag{C.5.3}$$

We wish to take a limit as $\varepsilon\to 0$. We shall pair every term in the product over Rays(C) in the numerator of (C.5.3) coming from a bad ray with some term in the product over BR; we can do this because the number of bad rays at each v is at most k. These give expressions of the form $\frac{1-q^{a\varepsilon}}{1-q^{b\varepsilon}}$ with a and b non-zero (each term has a different a and b), and these functions extend holomorphically to $\varepsilon=0$ (it is a removable singularity). All the remaining terms in the numerator are also holomorphic at $\varepsilon = 0$.

We now wish to apply L'Hopital's rule. We know that the limit at $\varepsilon \to 0$ must exist because, as discussed above, this function must extend holomorphically to $\varepsilon = 0$. The denominator vanishes to order k. Hence the numerator must also vanish to order k. We now apply L'Hopital k times.

When we differentiate the denominator k times and set $\varepsilon = 0$ we get

$$\prod_{t \in \mathrm{BR}} \Big(-(\tau, t) \Big) \ln(q)^k.$$

This expression only depends on the type.

Now let's rewrite the numerator in the form

$$\sum_{(v,C) \text{ s.t. } v \text{ bad}} F(C; \alpha, \varepsilon, \tau) q^{(\alpha + \varepsilon \cdot \tau, v)}.$$

The functions F depends only on α, ε, τ, and the cone at v, but not on the specific coordinates of v (that it, for a family of polytopes with the same fixed type, F stays the same). The kth derivative of this expression is

$$\sum_{\ell=0}^{k} \frac{\partial^{k-\ell} F(C; \alpha, \varepsilon, \tau)}{\partial \varepsilon^{k-\ell}} q^{(\alpha+\varepsilon\cdot\tau, v)} (\tau, v)^\ell \ln(q)^\ell$$

and when we set $\varepsilon = 0$, we get

$$\left(\sum_{\ell=0}^{k} \left. \frac{\partial^{k-\ell} F(C; \alpha, \varepsilon, \tau)}{\partial \varepsilon^{k-\ell}} \right|_{\varepsilon=0} (\tau, v)^\ell \ln(q)^\ell \right) q^{(\alpha, v)}.$$

The coordinates of v are in turn linear functionals in the b_i. Hence these terms are degree (at most) k polynomials in the b_i. □

C.6: Polytopal balls and polytopal norms on affine buildings

Let $\mathcal{B}$ be the Bruhat-Tits building associated to $G = \mathrm{PGL}(3, F)$. Recall that we may identify the vertices of $\mathcal{B}$ with G/K where $K = \mathrm{PGL}(3, \mathcal{O})$. The Weyl-chamber valued metric $d_{A^+}(\cdot, \cdot)$ allows us to define certain sets in the building $\mathcal{B}$ using polytopes defined in some apartment.

Recall that $A < G$ denotes the collection of diagonal matrices whose diagonal entries are powers of ϖ, and A^+ denotes those elements in A for which the diagonal powers of ϖ are weakly decreasing. We may identify A with elements in the lattice $\Lambda \subset \mathfrak{a}$. Recall that $\mathfrak{a}^+ \subset \mathfrak{a}$ denotes the positive Weyl chamber, and we may identify elements in A^+ with $\mathfrak{a}^+ \cap \Lambda$.

Suppose $Q \subset \mathfrak{a}^+$ is a convex polytope whose vertices lie in Λ. Given a vertex $x \in \mathcal{B}$, we may consider the polytopal ball centered at x whose "shape" is determined by Q; we call this the *Q-shaped ball centered at x*:

$$B_Q(x) := \{y \in G/K : d_{A^+}(x, y) \in Q^\Lambda\}.$$

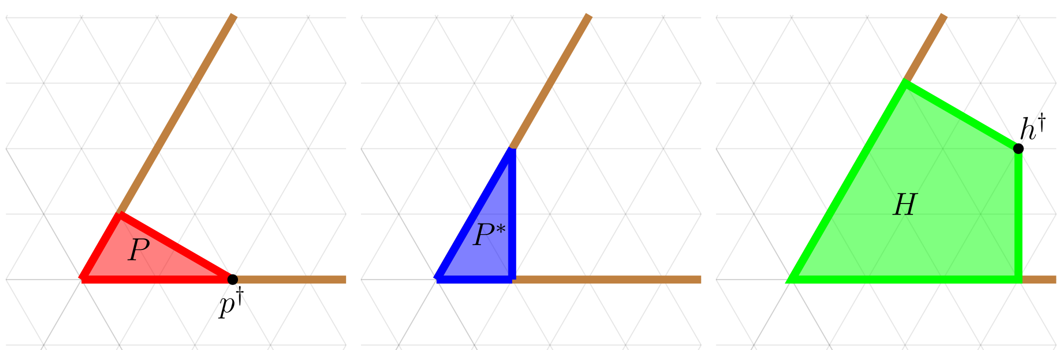
Similarly we can define a *polytopal norm* by

$$|x|_Q := \inf\{m \in \mathbb{R}_{\geq 0} : d_{A^+}(1K, x) \in Q_m^\Lambda\}$$

where Q_m is the mth dilate of Q. This tells us what the smallest dilate of Q is such that x is contained in the associated polytopal ball centered at $1K$.

Analogues of such polytopal balls for symmetric spaces appeared previously in the work of Brumley-Matz [BM21]. Also note that in rank one (such as the case of a regular tree or the hyperbolic plane), such polytopal balls are simply balls with respect to the usual metric.

C.6.1: The polytopes P, P^*, and H



(a) P (b) P^* (c) H

Figure 24: Geometric realization of the polytopes P, P^*, and H.

In this work with shall be concerned with polytopes which we call P, P^*, and H. Their defining linear inequalities in cone coordinates (r, s) are given in Table 1 (see Appendix B.9.1

for the definition of cone coordinates). Notice that Table 1 implies that if $d_{A^+}(1K, x) = (r, s)$, then

$$|x|_P = \frac{r + 2s}{2},$$
$$|x|_H = \max\left\{\frac{2r + s}{6}, \frac{r + 2s}{6}\right\}.$$

Polytope P	Polytope P^*	Polytope H
$r \geq 0$ $s \geq 0$ $r + 2s \leq 2$	$s \geq 0$ $r \geq 0$ $2r + s \leq 2$	$r \geq 0$ $s \geq 0$ $2r + s \leq 6$ $r + 2s \leq 6$

Table 1: Defining inequalities for P, P^* and H.

Notice that P^* is obtained from P by simply swapping the coordinates r and s. Recall the definition of $\lambda^\vee$ from (B.7.1). If $\lambda = (r, s)$ is cone coordinates, then $\lambda^\vee = (s, r)$. Therefore, by the defining property of $\lambda^\vee$, we get that

$$K\varpi^\lambda K \subset B_{P_m}(1K) \iff (K\varpi^\lambda K)^{-1} = K\varpi^{\lambda^\vee} K \subset B_{P_m^*}(1K).$$

The polytope P has a distinguished vertex which we denote by $p^\dagger$. What distinguishes

$\mathfrak{a}$-coordinates	$(\frac{4}{3}, -\frac{2}{3}, -\frac{2}{3})$
Partition coordinates	$(2, 0, 0)$
Cone coordinates	$(2, 0)$

(a) Coordinates of $p^\dagger$.

$\mathfrak{a}$-coordinates	$(2, 0, -2)$
Partition coordinates	$(4, 2, 0)$
Cone coordinates	$(2, 2)$

(b) Coordinates of $h^\dagger$.

Table 2: The points $p^\dagger$ and $h^\dagger$ in various coordinate systems.

$p^\dagger$ is that its dot product with $\delta = (1, 0, -1)$ is maximal among all vertices of P. Recall that δ is equal to $(1, 0, -1)$ in $\mathfrak{a}$-coordinates.

The polytope H also has a distinguished vertex which we denote by $h^\dagger$. What distinguishes $h^\dagger$ is that its dot product with δ is maximal among all vertices of H.

An important property of all 3 polytopes is that their orbit under the $\mathfrak{S}_3$-action on $\mathfrak{a}$ forms a convex set. See Figure 25.

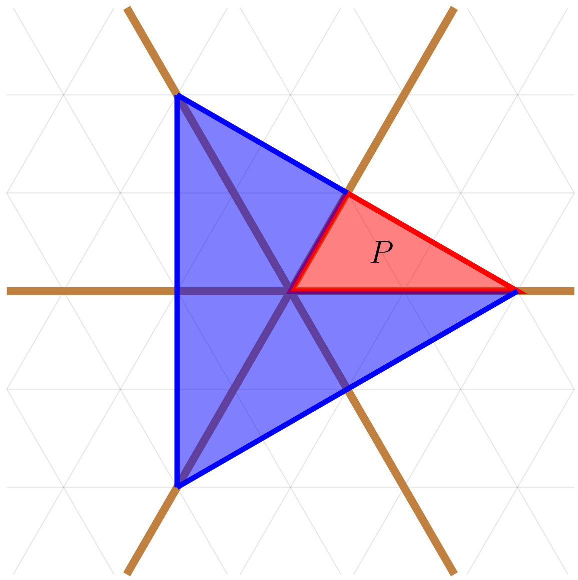


Figure 25: The blue polytope represents the orbit of P under $\mathfrak{S}_3$. Notice that this resulting polytope is convex.

BIBLIOGRAPHY

[ABB$^+$17] Miklos Abert, Nicolas Bergeron, Ian Biringer, Tsachik Gelander, Nikolay Nikolov, Jean Raimbault, and Iddo Samet. On the growth of L^2-invariants for sequences of lattices in Lie groups. *Ann. of Math. (2)*, 185(3):711–790, 2017.

[ABLM18] Miklos Abert, Nicolas Bergeron, and Etienne Le Masson. Eigenfunctions and random waves in the Benjamini-Schramm limit, 2018. arxiv:1810.05601.

[ALM15] Nalini Anantharaman and Etienne Le Masson. Quantum ergodicity on large regular graphs. *Duke Math. J.*, 164(4):723–765, 2015.

[Ana17] Nalini Anantharaman. Quantum ergodicity on regular graphs. *Comm. Math. Phys.*, 353(2):633–690, 2017.

[AS13] Nalini Anantharaman and Lior Silberman. A Haar component for quantum limits on locally symmetric spaces. *Israel J. Math.*, 195(1):393–447, 2013.

[AS19] Nalini Anantharaman and Mostafa Sabri. Recent results of quantum ergodicity on graphs and further investigation. *Ann. Fac. Sci. Toulouse Math. (6)*, 28(3):559–592, 2019.

[Bar93] Alexander I. Barvinok. Computing the volume, counting integral points, and exponential sums. *Discrete Comput. Geom.*, 10(2):123–141, 1993.

[BdlHV08] Bachir Bekka, Pierre de la Harpe, and Alain Valette. *Kazhdan's property (T)*, volume 11 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2008.

[Bee91] R. J. Beerends. Chebyshev polynomials in several variables and the radial part of the Laplace-Beltrami operator. *Trans. Amer. Math. Soc.*, 328(2):779–814, 1991.

[Ber74] I. N. Bernstein. All reductive p-adic groups are of type I. *Funkcional. Anal. i Priložen.*, 8(2):3–6, 1974.

[BKS07] G. Berkolaiko, J. P. Keating, and U. Smilansky. Quantum ergodicity for graphs related to interval maps. *Comm. Math. Phys.*, 273(1):137–159, 2007.

[BL13] Shimon Brooks and Elon Lindenstrauss. Non-localization of eigenfunctions on large regular graphs. *Israel J. Math.*, 193(1):1–14, 2013.

[BL14] Shimon Brooks and Elon Lindenstrauss. Joint quasimodes, positive entropy, and quantum unique ergodicity. *Invent. Math.*, 198(1):219–259, 2014.

[BLML16] Shimon Brooks, Etienne Le Masson, and Elon Lindenstrauss. Quantum ergodicity and averaging operators on the sphere. *Int. Math. Res. Not. IMRN*, (19):6034–6064, 2016.

[BM00] M. Bachir Bekka and Matthias Mayer. *Ergodic theory and topological dynamics of group actions on homogeneous spaces*, volume 269 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2000.

[BM21] Farrell Brumley and Jasmin Matz. Quantum ergodicity for compact quotients of $\mathrm{SL}_d(\mathbb{R})/\mathrm{SO}(d)$ in the Benjamini-Schramm limit. *Journal of the Institute of Mathematics of Jussieu*, page 1–41, 2021.

[Bri88] Michel Brion. Points entiers dans les polyèdres convexes. *Ann. Sci. École Norm. Sup. (4)*, 21(4):653–663, 1988.

[Bro89] Kenneth S. Brown. *Buildings*. Springer-Verlag, New York, 1989.

[BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001.

[BS14] Curtis D. Bennett and Petra N. Schwer. On axiomatic definitions of non-discrete affine buildings. *Adv. Geom.*, 14(3):381–412, 2014. With an appendix by Koen Struyve.

[BT72] F. Bruhat and J. Tits. Groupes réductifs sur un corps local. *Inst. Hautes Études Sci. Publ. Math.*, (41):5–251, 1972.

[Car79] P. Cartier. Representations of p-adic groups: a survey. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 111–155. Amer. Math. Soc., Providence, R.I., 1979.

[Cas95] William Casselman. Introduction to the theory of admissible representations of p-adic reductive groups. Unpublished Lecture Notes, May 1995.

[CdV85] Y. Colin de Verdière. Ergodicité et fonctions propres du laplacien. *Comm. Math. Phys.*, 102(3):497–502, 1985.

[CHH88] M. Cowling, U. Haagerup, and R. Howe. Almost L^2 matrix coefficients. *J. Reine Angew. Math.*, 387:97–110, 1988.

[CMo94] Donald I. Cartwright and Wojciech Mł otkowski. Harmonic analysis for groups acting on triangle buildings. *J. Austral. Math. Soc. Ser. A*, 56(3):345–383, 1994.

[Cow78] Michael Cowling. The Kunze-Stein phenomenon. *Ann. of Math. (2)*, 107(2):209–234, 1978.

[DE09] Anton Deitmar and Siegfried Echterhoff. *Principles of harmonic analysis*. Universitext. Springer, New York, 2009.

[Dei18] Anton Deitmar. Benjamini-Schramm and spectral convergence. *Enseign. Math.*, 64(3-4):371–394, 2018.

[Fol95] Gerald B. Folland. *A course in abstract harmonic analysis*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.

[Fra21] Mikolaj Fraczyk. Strong limit multiplicity for arithmetic hyperbolic surfaces and 3-manifolds. *Invent. Math.*, 224(3):917–985, 2021.

[GL18] Tsachik Gelander and Arie Levit. Invariant random subgroups over non-Archimedean local fields. *Math. Ann.*, 372(3-4):1503–1544, 2018.

[GN10] Alexander Gorodnik and Amos Nevo. *The ergodic theory of lattice subgroups*, volume 172 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2010.

[GN15] Alexander Gorodnik and Amos Nevo. Quantitative ergodic theorems and their number-theoretic applications. *Bull. Amer. Math. Soc. (N.S.)*, 52(1):65–113, 2015.

[HS10] Roman Holowinsky and Kannan Soundararajan. Mass equidistribution for Hecke eigenforms. *Ann. of Math. (2)*, 172(2):1517–1528, 2010.

[Kes59] Harry Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.

[KM08] Michael Kapovich and John J. Millson. A path model for geodesics in Euclidean buildings and its applications to representation theory. *Groups Geom. Dyn.*, 2(3):405–480, 2008.

[Kna86] Anthony W. Knapp. *Representation theory of semisimple groups*, volume 36 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1986. An overview based on examples.

[KS60] R. A. Kunze and E. M. Stein. Uniformly bounded representations and harmonic analysis of the 2×2 real unimodular group. *Amer. J. Math.*, 82:1–62, 1960.

[KS97] Tsampikos Kottos and Uzy Smilansky. Quantum chaos on graphs. *Phys. Rev. Lett.*, 79:4794–4797, 1997.

[Leu06] Enrico Leuzinger. Entropy of the geodesic flow for metric spaces and Bruhat-Tits buildings. *Adv. Geom.*, 6(3):475–491, 2006.

[Lin01] Elon Lindenstrauss. On quantum unique ergodicity for $\Gamma\backslash\mathbb{H}\times\mathbb{H}$. *Internat. Math. Res. Notices*, (17):913–933, 2001.

[Lin06] Elon Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)*, 163(1):165–219, 2006.

[LM14] Etienne Le Masson. Pseudo-differential calculus on homogeneous trees. *Ann. Henri Poincaré*, 15(9):1697–1732, 2014.

[LMS17] Etienne Le Masson and Tuomas Sahlsten. Quantum ergodicity and Benjamini-Schramm convergence of hyperbolic surfaces. *Duke Math. J.*, 166(18):3425–3460, 2017.

[LMS20] Etienne Le Masson and Tuomas Sahlsten. Quantum ergodicity for Eisenstein series on hyperbolic surfaces of large genus, 2020. arxiv:2006.14935.

[LSV05a] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Explicit constructions of Ramanujan complexes of type $\tilde{A}_d$. *European J. Combin.*, 26(6):965–993, 2005.

[LSV05b] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Ramanujan complexes of type $\tilde{A}_d$. *Israel J. Math.*, 149:267–299, 2005. Probability in mathematics.

[Lub14] Alexander Lubotzky. Ramanujan complexes and high dimensional expanders. *Jpn. J. Math.*, 9(2):137–169, 2014.

[Mac71] I. G. Macdonald. *Spherical functions on a group of p-adic type*. Publications of the Ramanujan Institute, No. 2. University of Madras, Centre for Advanced Study in Mathematics, Ramanujan Institute, Madras, 1971.

[Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[Mar91] G. A. Margulis. *Discrete subgroups of semisimple Lie groups*, volume 17 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991.

[McK81] Brendan D. McKay. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.*, 40:203–216, 1981.

[Nel11] Paul D. Nelson. Equidistribution of cusp forms in the level aspect. *Duke Math. J.*, 160(3):467–501, 2011.

[Nel18] Paul D. Nelson. Microlocal lifts and quantum unique ergodicity on $GL_2(\mathbb{Q}_p)$. *Algebra Number Theory*, 12(9):2033–2064, 2018.

[Nev98] Amos Nevo. Spectral transfer and pointwise ergodic theorems for semi-simple Kazhdan groups. *Math. Res. Lett.*, 5(3):305–325, 1998.

[NPS14] Paul D. Nelson, Ameya Pitale, and Abhishek Saha. Bounds for Rankin-Selberg integrals and quantum unique ergodicity for powerful levels. *J. Amer. Math. Soc.*, 27(1):147–191, 2014.

[PT21] Lam Pham and François Thilmany. Arithmetic groups and the Lehmer conjecture. *Israel J. Math.*, 245(2):839–868, 2021.

[Ron89] Mark Ronan. *Lectures on buildings*, volume 7 of *Perspectives in Mathematics*. Academic Press, Inc., Boston, MA, 1989.

[RS94] Zeév Rudnick and Peter Sarnak. The behaviour of eigenstates of arithmetic hyperbolic manifolds. *Comm. Math. Phys.*, 161(1):195–213, 1994.

[Sau97] François Sauvageot. Principe de densité pour les groupes réductifs. *Compositio Math.*, 108(2):151–184, 1997.

[Ser79] Jean-Pierre Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1979. Translated from the French by Marvin Jay Greenberg.

[Ser80] Jean-Pierre Serre. *Trees*. Springer-Verlag, Berlin-New York, 1980. Translated from the French by John Stillwell.

[Sni74] A. I. Snirelman. Ergodic properties of eigenfunctions. *Uspehi Mat. Nauk*, 29(6(180)):181–182, 1974.

[Sou10] Kannan Soundararajan. Quantum unique ergodicity for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. *Ann. of Math. (2)*, 172(2):1529–1538, 2010.

[SV07] Lior Silberman and Akshay Venkatesh. On quantum unique ergodicity for locally symmetric spaces. *Geom. Funct. Anal.*, 17(3):960–998, 2007.

[SV19] Lior Silberman and Akshay Venkatesh. Entropy bounds and quantum unique ergodicity for Hecke eigenfunctions on division algebras. In *Probabilistic methods in geometry, topology and spectral theory*, volume 739 of *Contemp. Math.*, pages 171–197. Amer. Math. Soc., [Providence], RI, 2019.

[Tad83a] M. Tadić. Harmonic analysis of spherical functions on reductive groups over $\mathfrak{p}$-adic fields. *Pacific J. Math.*, 109(1):215–235, 1983.

[Tad83b] M. Tadić. The topology of the dual space of a reductive group over a local field. *Glas. Mat. Ser. III*, 18(38)(2):259–279, 1983.

[Tad87] Marko Tadić. Topology of unitary dual of non-Archimedean $\mathrm{GL}(n)$. *Duke Math. J.*, 55(2):385–422, 1987.

[Vec02] Alessandro Veca. *The Kunze-Stein phenomenon*. PhD thesis, Universita degli Studi di Milano, 2002.

[Wol01] Scott A. Wolpert. The modulus of continuity for $\Gamma_0(m)\backslash\mathbb{H}$ semi-classical limits. *Comm. Math. Phys.*, 216(2):313–323, 2001.

[Zel87] Steven Zelditch. Uniform distribution of eigenfunctions on compact hyperbolic surfaces. *Duke Math. J.*, 55(4):919–941, 1987.

[Zel06] Steven Zelditch. Quantum ergodicity and mixing of eigenfunctions. In *Encyclopedia of Mathematical Physics*. Academic Press/Elsevier Science, 2006.