Canad. J. Math. Vol. **51** (2), 1999 pp. 266–293

Spectral Estimates for Towers of Noncompact Quotients

Anton Deitmar and Werner Hoffman

Abstract. We prove a uniform upper estimate on the number of cuspidal eigenvalues of the Γ-automorphic Laplacian below a given bound when Γ varies in a family of congruence subgroups of a given reductive linear algebraic group. Each Γ in the family is assumed to contain a principal congruence subgroup whose index in Γ does not exceed a fixed number. The bound we prove depends linearly on the covolume of Γ and is deduced from the analogous result about the cut-off Laplacian. The proof generalizes the heat-kernel method which has been applied by Donnelly in the case of a fixed lattice Γ.

Introduction

If Δ is a self-adjoint operator with discrete spectrum bounded from below and $N(\lambda)$ is the number of its eigenvalues (counted with multiplicity) not exceeding λ, we call N the spectral counting function of Δ. Let X be a symmetric space of the noncompact type and Γ_0 an arithmetic lattice in the isometry group G of X. For each torsion-free subgroup Γ of finite index in Γ_0, let $\Delta^{\Gamma}_{\mathrm{cus}}$ be the restriction of the Laplacian on $\Gamma \setminus X$ to the cuspidal subspace of $L^2(\Gamma \setminus X)$. In this paper we derive an upper bound for the spectral counting function of $\Delta^{\Gamma}_{\mathrm{cus}}$ which is uniform in Γ. In fact, we consider the slightly more general case $L^2(\Gamma \setminus X, \Gamma \setminus E)$, where E is a homogeneous hermitian vector bundle over X. The emphasis is on noncompact quotients $\Gamma \setminus X$. For a single lattice Γ, the result has been proved by Donnelly [6], and we will use his approach.

Our spectral estimate for the cuspidal Laplacian will be an immediate consequence of an analogous estimate for the cut-off Laplacian (also called pseudo-Laplacian) Δ^{Γ}_{T} introduced in [10, p. 489]. For a fixed lattice Γ, the estimate for Δ^{Γ}_{T} was proved in [10] by an easy adaption of the method of [6] and played a significant role in the proof of the trace-class conjecture. Our generalization of this result is a prerequisite for a forthcoming paper, where we study limit multiplicities of subsets of the unitary dual of G when $\Gamma \setminus X$ goes up in a tower.

Actually, our estimates apply to more general families than just towers. However, we cannot consider arbitrary families of lattices and not even arbitrary towers. So we have to be more precise. Recall that the principal congruence subgroup of level N in $\mathrm{GL}_n(\mathbb{Z})$ is defined as the kernel $\Gamma_n(N)$ of the residue map $\mathrm{GL}(n, \mathbb{Z}) \to \mathrm{GL}_n(\mathbb{Z}/N\mathbb{Z})$. Let $\mathcal{G}$ be a connected reductive linear algebraic $\mathbb{Q}$-group.

Definition A family $\mathcal{T}$ of subgroups of $\mathcal{G}(\mathbb{Q})$ will be called a *family of bounded depth* in $\mathcal{G}(\mathbb{Q})$ (with respect to a faithful $\mathbb{Q}$-rational representation $\eta\colon \mathcal{G} \to \mathrm{GL}_n$) if there exists a

Received by the editors February 26, 1998; revised September 4, 1998.
AMS subject classification: Primary: 11F72; secondary: 58G25, 22E40.

natural number D with the following property: For each $\Gamma \in \mathcal{T}$ there is a natural number N such that $\Gamma_n(N) \cap \eta\big(\mathcal{G}(\mathbb{Q})\big)$ is a subgroup of $\eta(\Gamma)$ of index at most D.

It is easy to see that this notion is independent of the choice of η. Now assume that $\mathcal{G}$ is semisimple (this assumption will be weakened below) and put $G = \mathcal{G}(\mathbb{R})$, $X = G/K$, where K is a fixed maximal compact subgroup of G. We fix a unitary representation τ of K on a finite-dimensional Hilbert space V_τ. Then $E_\tau := G \times_K V_\tau$ is a homogeneous hermitian vector bundle over X endowed with a canonical G-invariant connection ∇ and a Bochner-Laplace operator $\Delta = \nabla^*\nabla$. If Γ is torsion-free, then $E_\tau^\Gamma := \Gamma \setminus E_\tau$ is a hermitian vector bundle over the locally-symmetric space $\Gamma \setminus X$ inheriting a Laplacian Δ^Γ. Let Δ_T^Γ be the corresponding cut-off Laplacian with coefficients in the bundle E_τ^Γ. Its definition will be recalled in the first section. We shall prove (see Corollary 17 in Section 6) that for a family $\mathcal{T}$ of bounded depth in Γ_0 the spectral counting function $N_T^\Gamma(\lambda)$ of Δ_T^Γ satisfies the estimate

$$N_T^\Gamma(\lambda) \le C[\Gamma_0 : \Gamma](1 + \lambda)^{d/2},$$

where $d = \dim X$ and $C > 0$ is independent of $\Gamma \in \mathcal{T}$ and $\lambda \ge 0$. In a final section, we give an adelic version, which is slightly more general if $\mathcal{G}$ does not have the strong approximation property.

An inspection of our proof shows that the bounded depth assumption is unnecessary for $\Gamma_0 \setminus X$ compact. We do not know whether in the general case some restriction of this kind is necessary or whether this is only a drawback of our method.

Following [6], we shall derive our result by von Neumann bracketing from analogous spectral estimates for cuspidal Laplacians with von Neumann boundary conditions on certain submanifolds of $\Gamma \setminus X$. However, the method does not immediately carry over. Starting from the heat kernel on the symmetric space X, Donnelly obtains the heat kernel $\underline{F}(t, x, y)$ of the quotient $\Gamma \cap P \setminus X$ for a parabolic P by averaging. Next he produces a kernel $\bar{F}(t, x, y)$ by projecting F on the cuspidal subspace. Lastly he modifies $\bar{F}$ by adding a single-layer potential to end up with the kernel $\bar{E}(t, x, y)$ satisfying the boundary conditions. Actually, the kernel $\bar{F}$ does not have the short-range asymptotic of a heat kernel because of the averaging over horospheres implicit in its construction, and Propositions 5.6 and 5.7 of [6] become valid only after replacing the distance in X by the distance in $N \setminus X$. However, the method of [13] Donnelly refers to relies on the jump relations for single-layer potentials which have only been proved for the undisturbed heat kernel. Moreover, it would not be easy to carry out the proof of Proposition 6.1 in [6] because the boundary does not have the asserted simple description outside a compact set. Thus, the proof of the main result of [6] is incomplete. Our present paper fills this gap in the prerequisites of [10], because any single lattice in G constitutes a family of bounded depth.

Since we are going to let Γ vary, we would face the additional problem of keeping track of the growing boundary in doing the necessary estimates. Therfore we proceed differently in that we incorporate the boundary conditions already on the universal covering and then follow the other steps. The method of single-layer potentials seems to become unmanageable for such noncompact manifolds with boundary, and we construct the heat kernel by a new method, which may also be useful in other situations. Namely, we construct a parametrix which already satisfies the boundary conditions.

As in [6], a majorant for the heat kernel on the universal cover yields an estimate for the heat kernel on the quotient by averaging over $\Gamma \cap P$. However, instead of bounds obtained from the compactness of $\Gamma \cap N \setminus N$ and of other quotients we have to include additional arguments to see how the bounds vary if the quotient changes. It is for this reason that we can only admit families of bounded depth. A key step in [6] consists in expressing a difference of heat kernel values as a path integral. For fixed Γ, the length of this path is uniformly bounded and can therefore be estimated by the distance of its endpoints. However, the subdomains we consider are not geodesically closed in general, and for varying Γ the minimal length of a path connecting two points is not bounded by a fixed multiple of their distance in X. Since the heat kernel majorant is expressed in terms of that latter distance, we have to compare various metrics. Finally, the proof of Proposition 5.6 in [6] uses the boundedness of the averaging operator over horospheres, but the norm of this operator varies with Γ. In effect, we have to rewrite the whole argument.

Following [10], we use a decomposition of $\Gamma \setminus X$ into pieces indexed by all Γ-conjugacy classes of parabolic Q-subgroups, as this seems most natural and admits the easiest description of the boundaries. If the truncation parameter grows, however, only the pieces associated to the minimal parabolics will shrink. Thus, if one wanted to recover the results of [6] concerning the upper limit of the spectral counting function, one would have to use, as in that paper, another covering of $\Gamma \setminus X$ consisting of a compact part and its complements in the Siegel domains for the minimal parabolics. One should also be able to make the estimates uniform in the K-type and to handle the Laplacian on $\Gamma \setminus G$ using the ideas of [11] and [8].

Acknowledgement The second author would like to express his gratitude to the Institute for Advanced Study in Princeton for hospitality and financial support during the fall term 1997, when this paper was completed.

1 A Parametrix Construction

In this section we will construct a parametrix for the heat equation on certain noncompact manifolds with boundary. Our purpose is to obtain uniform estimates. The new feature is that the parametrix will already satisfy the boundary conditions. Certainly, our method would work under some general assumptions on bounded geometry, but we did not explore what their exact formulation should be. Instead, we require a certain group invariance, which will be granted in our applications (*cf.* Prop. 2), albeit in a non-straightforward way.

We consider a hermitian vector bundle E endowed with a metric connection over a possibly noncompact Riemannian manifold Ω with boundary. We assume that Ω is given as a subdomain with smooth boundary in a complete Riemannian manifold X and that E extends over X. Then, by imposing Dirichlet or von Neumann boundary conditions, we obtain a self-adjoint extension Δ_Ω of the Bochner-Laplace operator $\Delta = \nabla^*\nabla$ in the Hilbert space of square-integrable sections. This can be proved by adapting the method of [4]. Indeed, Proposition 1.1 in [4] remains true if one replaces the truncated cone considered there by its intersection with $\mathrm{R} \times \Omega$. In the proof, this produces new boundary terms in Green's formula, which vanish, however, due to the boundary conditions. In this way one proves propagation speed one with respect to the distance in the surrounding manifold X. By subdividing a given time interval finer and finer, one gets the same with respect to the distance inside Ω. Having proved the existence of the self-adjoint extension Δ_Ω, it is easy to deduce from the spectral theorem and Sobolev's embedding theorem that the Schwartz kernel of the heat operator $\exp(-t\Delta_\Omega)$, $t > 0$, on the interior of $\Omega \times \Omega$ is a smooth section $E_\Omega(t, x, y)$ of $E \boxtimes E^*$. This heat kernel is clearly determined by the property that, for fixed y, it is a solution of the Cauchy problem for the heat equation with initial value equal to the delta distribution at y.

In order to obtain bounds on the heat kernel, we will now construct a right parametrix. This is a smooth section $H_\Omega(t, x, y)$ of $E \boxtimes E^*$ depending smoothly on $t > 0$ and such that

$$B_\Omega(t, x, y) := \left(\Delta_x + \frac{\partial}{\partial t}\right) H_\Omega(t, x, y)$$

is smooth, too. We also require that $H_\Omega(t, x, y)$ as a function of x satisfy the boundary conditions for each fixed y and t, be compactly supported and tend to the delta distribution at y as $t \to 0$.

Lemma 1 *Suppose that there exists a subgroup S of the isometry group of X leaving Ω invariant. Suppose further that $S \setminus \Omega$ is compact and that the action of S lifts to an action on E by automorphisms. Then, given $c > 4$, $\delta > 0$ and $t_0 > 0$, there exists a right parametrix $H_\Omega(t, x, y)$ for the heat equation on Ω with von Neumann boundary conditions and coefficients in E satisfying the estimates*

$$|H_\Omega(t, x, y)| \leq Ct^{-d/2} e^{-\rho(x,y)^2/ct},$$

$$|\nabla_x H_\Omega(t, x, y)| \leq Ct^{-(d+1)/2} e^{-\rho(x,y)^2/ct},$$

$$|B_\Omega(t, x, y)| \leq C$$

for all $0 < t \leq t_0$ and some $C > 0$, where $d = \dim X$ and ρ denotes the geodesic distance in X. Moreover, the value $H_\Omega(t, x, y)$ depends only on the geometry of the 2δ-neighborhood of y in X and vanishes for $\rho(x, y) \geq \delta$.

Proof As a preparation for our later construction, let us consider the case when Ω is a compact domain in X with smooth boundary. The results of [7] easily extend to the bundle case and show that $E_\Omega(t, x, y)$ exists and satisfies the estimates we require of the parametrix. (The case of the differential form bundle with different boundary conditions is done in [13].) Let us denote the smallest possible constant in the estimates by $C(\Omega)$. If ω_z is an embedding of a d-dimensional ball B into X preserving the interior unit normal and depending smoothly on $z \in Z$ for some manifold Z, then $C\big(\omega_z(B)\big)$ is locally bounded on Z. In fact, pulling back everything under ω_z and trivializing the bundle, one reduces the proof to the case of a fixed domain and a differential operator with coefficients depending smoothly on z, which can be handled by the methods of [7]. Moreover, the pulled-back heat kernel depends smoothly on z.

Before proceeding to the actual construction for an unbounded domain Ω, let us outline the basic idea. We manufacture the parametrix, considered as a function of the first variable x, from the heat kernel of a bounded subdomain by means of a cut-off function. For this subdomain we take a neighborhood $\Omega(y)$ of the varying point y, which is simply a ball with center y if y is far from the boundary of Ω, but which nestles to that boundary when y comes close to it. The deformation of the ball into $\Omega(y)$ is done by a retraction ψ_+

of some neighborhood of Ω onto Ω. We use the heat kernel with von Neumann boundary conditions on the boundary of $\Omega(y)$, which coincides with the boundary of Ω in a small neighborhood of y. To obtain the parametrix, we multiply by a cut-off function χ supported in that neighborhood. The boundary conditions are not affected because we make sure that the normal derivative of χ vanishes on the boundary. For this purpose, we use a retraction ψ_- of Ω onto some subdomain.

We will now carry out the construction under certain assumptions on Ω which we will formulate in due course. Afterwards we will see that they are justified in the situation of the lemma. At first, we assume that there exists $\varepsilon_1 > 0$ such that the map $\varphi\colon \partial\Omega \times]-\varepsilon_1, \varepsilon_1[\to X$ given by $\varphi(x, r) = \exp_x(r\nu_x)$ is an embedding and that the distance of any $\varphi(x, r)$ from $\partial\Omega$ equals $|r|$. Here ν_x denotes the inner unit normal vector of Ω at x and $\exp_x$ the exponential map of X at the point x. Let Ω_+ be the union of Ω with the range of φ.

We choose two functions $\eta_+, \eta_- \in C^\infty(\mathrm{R})$ such that $\eta_\pm(r) = r$ for $r \geq \varepsilon_\pm$, where $0 < \varepsilon_\pm < \varepsilon_1$. We require that $\eta_+(r) = 0$ for $r \leq -\varepsilon_+$, and $\eta_+'(r) > 0$ for $r > -\varepsilon_+$ while $\eta_-(r) = \varepsilon_-/2$ for $r \leq 0$ and $\eta_-'(r) > 0$ for $r > 0$. Now we define smooth maps $\psi_+\colon \Omega_+ \to \Omega$ and $\psi_-\colon \Omega \to \Omega$ as follows. We put $\psi_\pm\bigl(\varphi(x, r)\bigr) = \varphi\bigl(x, \eta_\pm(r)\bigr)$ and $\psi_\pm(x) = x$ if $x \in \Omega$ is not in the range of φ.

For each $y \in \Omega$, let $\Omega(y)$ be the image under ψ_+ of the ball around y with radius ε_2, where $\varepsilon_+ < \varepsilon_2 < \varepsilon_1$. We assume that ε_2 can be chosen as close to ε_+ as to ensure that, for any $y \in \Omega$, no normal vector to $\varphi(\partial\Omega \times \{-\varepsilon_+\})$ is tangent to that ball. Then $\Omega(y)$ has a smooth boundary, and each point in Ω has a neighborhood Z in X such that $\Omega(y) = \omega_y(B)$ for $y \in Z$, where ω_y is some smooth family of embeddings. Thus the remarks made at the beginning of the proof apply to $\Omega(y)$. It is clear that $\Omega(y)$ contains a neighborhood of y in Ω.

Now we want to multiply $E_{\Omega(y)}(t, x, y)$ by a cut-off function. Thus, let $\eta \in C_0^\infty(\mathrm{R}_+)$ be such that $\eta(r) = 1$ for $r \leq \varepsilon_3$ and $\eta(r) = 0$ for $r \geq 2\varepsilon_3$, where $\varepsilon_3 > \varepsilon_-$, and put $\chi(x, y) = \eta\Bigl(\rho\bigl(\psi_-(x), y\bigr)\Bigr)$. This is a smooth function satisfying von Neumann boundary conditions in x and being equal to 1 in a neighborhood of the diagonal. We assume that ε_- and hence ε_3 can be chosen so small that $\operatorname{supp}_x \chi(x, y) \subset \Omega(y)$ for all $y \in \Omega$. Then $E_{\Omega(y)}(t, x, y)$ is smooth on the support of $\chi(x, y)$, and $\chi(x, y)$ vanishes for $\rho(x, y) \geq \delta$. If we now define $H_\Omega(t, x, y) = \chi(x, y)E_{\Omega(y)}(t, x, y)$, this is a smooth section of the pull-back of $E \boxtimes E^*$ to $\mathrm{R}_+ \times \Omega \times \Omega$ satisfying the required support condition. If $(x, y) \in \operatorname{supp} \chi \cap \partial\Omega \times \Omega$, then $x \in \partial\Omega(y)$ by our assumption. Hence H_Ω satisfies von Neumann boundary conditions in x at $\partial\Omega$. Since χ equals one in a neighborhood of the diagonal, we see that H_Ω is the required parametrix.

It is clear that the assumptions made during the construction are satisfied locally. Since by hypothesis the whole structure is S-invariant, the same ε_*'s work for all points in one S-orbit. The compactness of $S \setminus \Omega$ guarantees that they can be chosen globally. Moreover, since $C\bigl(\Omega(sy)\bigr) = C\bigl(\Omega(y)\bigr)$ for all $s \in S$, $y \in \Omega$, we see that H_Ω satisfies the required estimates with a constant C independent of y. ∎

2 Heat Kernels on the Universal Cover

In the present section we first recall a partition of the symmetric space X into domains $X(P)$ indexed by the parabolic Q-subgroups of $\mathcal{G}$. Actually, we are after their projections on the

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

quotient $\Gamma \setminus X$. Among those projections, only $\Gamma \setminus X(G)$ is compact, and the other ones are sometimes simply called cusps. For the time being, however, we stay with X. We smooth out the edges of $X(P)$ and prove estimates for the heat kernel with von Neumann boundary conditions on the resulting domain $\tilde{X}(P)$.

First we fix some notation, which will be used in the rest of the paper. We shall generally denote linear algebraic $\mathbb{Q}$-groups by calligraphic letters, their sets of $\mathbb{R}$-rational points by the corresponding roman letters and the Lie algebras by the corresponding lower-case gothic letters. For any linear algebraic $\mathbb{Q}$-group $\mathcal{H}$, let H^1 be the common kernel in $H = \mathcal{H}(\mathbb{R})$ of the absolute values of all $\mathbb{Q}$-rational characters of $\mathcal{H}$ and put $A_H = H/H^1$. If $\mathcal{M}$ is a Levi subgroup of $\mathcal{G}$, we can identify A_M with $\mathcal{A}_M(\mathbb{R})^0$, where $\mathcal{A}_M$ is the maximal $\mathbb{Q}$-split torus in the center of $\mathcal{M}$, and then $M = M^1 \times A_M$.

For our fixed connected reductive $\mathbb{Q}$-group $\mathcal{G}$, we set $X = G^1/K$. Let $\mathcal{P}$ be a parabolic $\mathbb{Q}$-subgroup of $\mathcal{G}$ with unipotent radical $\mathcal{N}$ and $\mathcal{M}$ any Levi component of $\mathcal{P}$ defined over $\mathbb{Q}$. Then $\mathcal{A}_M$ is also called a $\mathbb{Q}$-split component of $\mathcal{P}$. Let X^P be the M^1-orbit of the trivial coset in $N \setminus X$. This is a symmetric space isomorphic to M^1/K^P, where K^P is the projection of $K \cap P$ to M. One has $P^1 = M^1 N$, hence $A_M \cong A_P$. If $\mathcal{P}' \subset \mathcal{P}$, we define $A^P_{P'} = A_{P'} \cap M^1$. Then $\mathfrak{a}_{P'} = \mathfrak{a}^P_{P'} \oplus \mathfrak{a}_P$, and we write the projection $\mathfrak{a}_{P'} \to \mathfrak{a}_P$ as $H \mapsto H_P$. Let $\mathfrak{a}^{P+}_{P'}$ be the subset of $\mathfrak{a}^P_{P'}$ where all the roots of $\mathcal{A}_{M'}$ in $\mathfrak{n}' \cap \mathfrak{m}$ are positive, and define

$$^+\mathfrak{a}^P_{P'} = \{H \in \mathfrak{a}^P_{P'} \mid (H, H') \geq 0 \quad \forall H' \in \mathfrak{a}^{P+}_{P'}\},$$

where we have chosen Euclidean structures on all the $\mathfrak{a}_P$ compatible with $\mathcal{G}(\mathbb{Q})$-conjugation and with the decompositions $\mathfrak{a}_{P'} = \mathfrak{a}^P_{P'} \oplus \mathfrak{a}_P$.

Now we describe the decomposition of X connected with the truncation operator Λ^T (*cf.* [1] or [12] for details). Given K and an arithmetic subgroup Γ_0 of $\mathcal{G}(\mathbb{Q})$, one can think of the truncation parameter T as of a family of points $T_P \in \mathfrak{a}^G_P$ indexed by the parabolic $\mathbb{Q}$-subgroups such that

- $\gamma \cdot P^1 \exp(T_P)K = P'^1 \exp(T_{P'})K$ for $\gamma \in \Gamma_0$ and $\mathcal{P}' = \gamma\mathcal{P}\gamma^{-1}$,
- $T_P = (T_{P'})_P$ for $\mathcal{P}' \subset \mathcal{P}$.

Any such T is determined by the values T_P for P in a set of representatives of Γ_0-conjugacy classes of maximal parabolic $\mathbb{Q}$-subgroups. Thereby the set of truncation parameters becomes an affine space with the partial ordering given by $T < T'$ iff $T'_P - T_P \in \mathfrak{a}^{G+}_P$ for all $\mathcal{P}$. For $\mathcal{P} = \mathcal{M}\mathcal{N}$ defined over $\mathbb{Q}$, the projection of $\Gamma_0 \cap P$ on M is an arithmetic subgroup Γ^P_0 of $\mathcal{M}$, and we get a truncation parameter T^P for Γ^P_0 by setting $T^P_{Q \cap M} = T_Q$ for each $\mathcal{Q} \subset \mathcal{P}$.

Let $X(G, T)$ be the set of all $x \in X$ with the property that for any proper (equivalently: any maximal) parabolic $\mathbb{Q}$-subgroup $\mathcal{P}$ of $\mathcal{G}$ the image of x in $N \setminus X$ does not belong to the set $\exp(^+\mathfrak{a}^G_P + T_P)X^P$. Then $X(G, T)$ is Γ_0-invariant, and $\Gamma \setminus X(G, T)$ is compact for any arithmetic subgroup $\Gamma \subset \Gamma_0$. More generally, define $X(P, T)$ to be the set of all $x \in X$ whose image in $N \setminus X$ belongs to

$$\exp(\mathfrak{a}^{G+}_P + T_P)X^P(M, T^P).$$

For T large enough, X is the disjoint union of the subsets $X(P, T)$, where $\mathcal{P}$ runs through the parabolic $\mathbb{Q}$-subgroups of $\mathcal{G}$ including $\mathcal{G}$ itself (see [1, Lemma 6.4], [12, Theorem 3.4]). We shall usually fix such T and simply write $X(P)$ for $X(P, T)$.

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

Since we now have the necessary notation, let us recall the definition of the cut-off Laplacian. As E_τ is a homogeneous bundle, there is a (right) action r of G^1 on its sections. We can view sections of E_τ^Γ as Γ-invariant sections of E_τ, thereby making the Laplacian Δ^Γ meaningful even for Γ with torsion (*cf.* [5]). Let $L^2_T(\Gamma \setminus X, E_\tau^\Gamma)$ be the set of all $f \in L^2(\Gamma \setminus X, E_\tau^\Gamma)$ which have the following property for all proper (equivalently: all maximal) parabolic $\mathbb{Q}$-subgroups $\mathcal{P}$ of $\mathcal{G}$: Whenever the image of x in $N \setminus X$ belongs to $\exp({}^+\mathfrak{a}_P^G + T_P)X^P$, then the constant term

$$\int_{\Gamma\cap N\setminus N} \bigl(r(n)f\bigr)(x)\,dn$$

of f along $\mathcal{P}$ vanishes. Note that, for T large enough, the restriction of Λ^T to $L^2(\Gamma \setminus X, E_\tau^\Gamma)$ is the orthoprojector onto $L^2_T(\Gamma \setminus X, E_\tau^\Gamma)$ (as follows from [2, Lemma 1.1], or [12, p. 39]) and $\Lambda^T(1)$ is the characteristic function of $\Gamma\setminus X(G,T)$. Now Δ_T^Γ is defined as the selfadjoint operator in $L^2_T(\Gamma \setminus X, E_\tau^\Gamma)$ associated to the quadratic form $\|\nabla f\|^2$ on the intersection of the Sobolev space $H^1(\Gamma \setminus X, E_\tau^\Gamma)$ with $L^2_T(\Gamma \setminus X, E_\tau^\Gamma)$. This cut-off Laplacian differs from that defined in [10] by an additive constant depending on τ.

The domains $X(P)$ have non-smooth boundary, and it is difficult to study the heat kernels for boundary value problems on them. Therefore we shall now define modified domains $\tilde{X}(P)$. If we are given a Euclidean space V, there is a standard procedure to smooth out any convex polytope C, say, with nonempty interior in V. Namely, we choose $\varepsilon > 0$ and $\eta \in C^\infty(\mathbb{R})$ with $\eta(x) > 0$, $\eta'(x) > 0$ for $x < 0$ and $\eta(x) = 1$ for $x \geq 0$. There is a unique minimal set Φ of affine functionals on V with slope one such that

$$C = \{v \in V \mid \varphi(v) \geq 0 \quad \forall \varphi \in \Phi\}.$$

Now we put

$$\tilde{C} = \Bigl\{v \in V \mid \prod_{\varphi\in\Phi} \eta\bigl(\varphi(v)\bigr) \geq \eta(-\varepsilon)\Bigr\}.$$

Then $\tilde{C}$ has a smooth boundary, and $C \subset \tilde{C} \subset \{v \in V \mid \varphi(v) \geq -\varepsilon \quad \forall \varphi \in \Phi\}$. We fix η and ε once and for all. Then $\tilde{C}$ is determined by C and will be called its smooth hull.

To apply this to $X(P)$, recall that there is a synthetic description of these domains. Choose T_0 small enough such that X is the union of the sets $X(P_0, T_0)$ over all minimal parabolic $\mathbb{Q}$-subgroups $\mathcal{P}_0$. A special case of Langlands' Combinatorial Lemma ([1, Lemma 6.3], [12, p. 321]) states that, for each $\mathcal{P}_0$, the vector space $\mathfrak{a}_{P_0}^G$ is the disjoint union of the cones

$$\mathfrak{a}_{P_0}^G(P,T) := {}^-\mathfrak{a}_{P_0}^P + \mathfrak{a}_P^{G+} + T_{P_0}$$

over all $\mathcal{P}$ containing $\mathcal{P}_0$, where ${}^-\mathfrak{a}_{P_0}^P = -{}^+\mathfrak{a}_{P_0}^P$. We denote by $X_{P_0,T_0}(P,T)$ the set of all $x \in X(P_0,T_0)$ whose image in $N_0 \setminus X$ belongs to $\exp\bigl(\mathfrak{a}_{P_0}^G(P,T)\bigr)X^{P_0}$. If T is large enough depending on T_0, then $X(P,T)$ is the disjoint union of the sets $X_{P_0,T_0}(P,T)$ over all minimal $\mathcal{P}_0$ contained in $\mathcal{P}$. If we now replace each $\mathfrak{a}_{P_0}^G(P,T)$ by its smooth hull $\tilde{\mathfrak{a}}_{P_0}^G(P,T)$, we get sets $\tilde{X}_{P_0,T_0}(P,T)$, whose union over $\mathcal{P}_0 \subset \mathcal{P}$ we denote by $\tilde{X}(P,T)$ and call the smooth hull of $X(P,T)$. One can check that, for T large enough, $\tilde{X}(P,T)$ is a domain in X with smooth boundary and that $X(P,T) \subset \tilde{X}(P,T) \subset X\big(P, T(P)\big)$, where $T - T(P)$ can be made arbitrarily small by the choice of ε. Moreover, $\gamma\tilde{X}(P,T) = \tilde{X}(\gamma P\gamma^{-1}, T)$ for each $\gamma \in \Gamma_0$.

Let us fix a unitary representation τ of K and consider the Bochner-Laplace operator Δ in the bundle E_τ over X. For each $\mathcal{P}$, let Δ_P be the self-adjoint extension of $\Delta|_{\tilde{X}(P)}$ with respect to the von Neumann boundary conditions. These boundary conditions amount to the vanishing of the covariant derivative in the normal direction.

Proposition 2 *Let $E_P(t,x,y)$ be the heat kernel on $\tilde{X}(P)$ with von Neumann boundary conditions and coefficients in E_τ. Then for any $c > 4$ and $t_0 > 0$ there is a constant C with*

$$|E_P(t,x,y)| \le Ct^{-d/2}e^{-\rho(x,y)^2/ct},$$

$$|\nabla_x E_P(t,x,y)| \le Ct^{-(d+1)/2}e^{-\rho(x,y)^2/ct}$$

for $0 < t \le t_0$, where ρ denotes the geodesic distance in X.

Since Δ_P is self-adjoint, we have $E_P(t,y,x) = E_P(t,x,y)^*$ for each $x, y \in \Omega$, so actually it does not matter to which argument the covariant derivation applies. Using the boundedness of $x^N e^{-x}$ on $[0,\infty[$ for $N \ge 0$, one may deduce a bound for the derivative in the form $Ct^{-d/2}\rho(x,y)^{-1}e^{-\rho(x,y)^2/ct}$. First we prove:

Lemma 3 *Given $\delta > 0$, there exists a right parametrix $H_P(t,x,y)$ for the heat equation on $\tilde{X}(P)$ with von Neumann boundary conditions and coefficients in E_τ vanishing for $\rho(x,y) \ge \delta$ and satisfying the estimates in Lemma 1.*

Proof In general, there exists no subgroup S of G leaving $\tilde{X}(P,T)$ invariant and such that $S \setminus \tilde{X}(P,T)$ is compact, so Lemma 1 is not immediately applicable. We shall decompose $\tilde{X}(P,T)$ into pieces such that each piece is part of a domain to which Lemma 1 applies. The restrictions of the parametrices provided by that lemma will match together and form the parametrix for $\tilde{X}(P,T)$.

To get an idea, consider first the analogous case of a convex polytope C in a Euclidean space V. If Φ is the set of affine functionals defining C, any $v \in C$ determines $\Phi(v) = \{\varphi \in \Phi \mid \varphi(v) = 0\}$, and the set of all v for which $\Phi(v)$ equals a given subset of Φ is called a face of C. Each face F of C determines a piece $\tilde{C}_F$ of the smooth hull consisting of all $v \in \tilde{C}$ whose closest point in C belongs to F. Now $\tilde{C}$ is the disjoint union of the pieces $\tilde{C}_F$ over all faces (including the interior of C). This decomposition is not yet good enough, because $\tilde{C}_F \cap \partial\tilde{C}$ need not be parallel to F. Therefore one has to decompose $\tilde{C}$ with respect to a polytope smaller than C obtained by subtracting a constant from each $\varphi \in \Phi$.

To understand the boundary of the domain $X(P)$, recall that it is fibered by horospheres over a domain in $N \setminus X$ which is isomorphic to the direct product of the simplicial cone $\mathfrak{a}_P^{G+}$ and the manifold with corners $X^P(M)$. The faces of $\mathfrak{a}_P^{G+}$ are parametrized by the parabolic $\mathbb{Q}$-subgroups $\mathcal{P}''$ containing $\mathcal{P}$, while the faces of $X^P(M)$ are parametrized by the parabolic $\mathbb{Q}$-subgroups $\mathcal{P}'$ contained in $\mathcal{P}$. Thus the faces of $X(P)$ are parametrized by pairs $(\mathcal{P}', \mathcal{P}'')$ sandwiching $\mathcal{P}$.

Now let us define the pieces of $\tilde{X}(P,T)$ precisely. For parabolic $\mathbb{Q}$-subgroups $\mathcal{P}' \subset \mathcal{P} \subset \mathcal{P}''$ and truncation parameters $T' < T < T''$ we let $\tilde{X}_{P',T'}^{P'',T''}(P,T)$ be the set of all elements of $\tilde{X}(P,T)$ whose projection on $N \setminus X$ lies in $\exp\bigl(\mathfrak{a}_P^G(P'',T'')\bigr)X^P(P' \cap M, T')$. If T' is large enough, then $\tilde{X}(P,T)$ is the disjoint union of the pieces $\tilde{X}_{P',T'}^{P'',T''}(P,T)$ over all $\mathcal{P}'$ contained in $\mathcal{P}$ and all $\mathcal{P}''$ containing $\mathcal{P}$. Clearly, for each $\gamma \in \Gamma_0$, the γ-translate of $\tilde{X}_{P',T'}^{P'',T''}(P,T)$ is obtained by replacing $\mathcal{P}'$, $\mathcal{P}$ and $\mathcal{P}''$ by their respective γ-conjugates.

We can give a synthetic description of these pieces as we did for $\tilde{X}(P,T)$. In explicit terms, $\tilde{X}_{P_0,T_0}(P,T)$ was defined as the set of all $x \in X$ whose image in $N_0 \setminus X$ belongs to

$$\exp\bigl(\mathfrak{a}_{P_0}^G(P_0,T_0) \cap \tilde{\mathfrak{a}}_{P_0}^G(P,T)\bigr)X^{P_0}.$$

Here we may replace $\mathfrak{a}_{P_0}^G(P_0,T_0)$ by $\mathfrak{a}_{P_0}^P(P_0,T_0) + \mathfrak{a}_P^G$, because both sets have the same intersection with $\tilde{\mathfrak{a}}_{P_0}^G(P,T)$ for $T \gg T_0$. Now one sees that $\tilde{X}_{P',T'}^{P'',T''}(P,T)$ is the union over all minimal $\mathcal{P}_0$ contained in $\mathcal{P}'$ of the preimages of the sets $\exp\bigl(C(P_0,T_0,P',T',P,T,P'',T'')\bigr)X^{P_0}$ under the map $X \to N_0 \setminus X$, where $C(P_0,T_0,P',T',P,T,P'',T'')$ stands for

$$\Bigl(\bigl(\mathfrak{a}_{P_0}^P(P_0,T_0) \cap \mathfrak{a}_{P_0}^P(P',T')\bigr) + \mathfrak{a}_P^G(P'',T'')\Bigr) \cap \tilde{\mathfrak{a}}_{P_0}^G(P,T).$$

For $T' \ll T \ll T''$ this set has a decomposition compatible with $\mathfrak{a}_{P_0}^G = \mathfrak{a}_{P_0}^{P'} \oplus \mathfrak{a}_{P'}^{P''} \oplus \mathfrak{a}_{P''}^G$. Firstly, the set $\mathfrak{a}_{P_0}^P(P',T')$ has the same intersection with $\mathfrak{a}_{P_0}^P(P_0,T_0)$ as with $\mathfrak{a}_{P_0}^{P'}(P_0,T_0) + \mathfrak{a}_{P'}^P$, and secondly, the set $\mathfrak{a}_{P_0}^{P'}(P',T') + \mathfrak{a}_{P'}^{P''} + \mathfrak{a}_{P''}^G(P'',T'')$ has the same intersection with $\tilde{\mathfrak{a}}_{P_0}^G(P,T)$ as with $\mathfrak{a}_{P_0}^{P'} + \tilde{\mathfrak{a}}_{P'}^{P''}(P,T) + \mathfrak{a}_{P''}^G$. Therefore $C(P_0,T_0,P',T',P,T,P'',T'')$ equals

$$\begin{aligned}&\bigl(\mathfrak{a}_{P_0}^{P'}(P_0,T_0) \cap \mathfrak{a}_{P_0}^{P'}(P',T')\bigr)\\ &\qquad + \Bigl(\bigl(\mathfrak{a}_{P'}^P(P',T') + \mathfrak{a}_P^{P''}(P'',T'')\bigr) \cap \tilde{\mathfrak{a}}_{P'}^{P''}(P,T)\Bigr) + \mathfrak{a}_{P''}^G(P'',T'').\end{aligned}$$

The component in $\mathfrak{a}_{P'}^{P''}$ is a partially smoothed-out compact polytope. Let us denote by $D(P_0,T_0,P',T',P,T,P'',T'')$ the following larger set obtained by smoothing out this component completely and enlarging the other components:

$$\mathfrak{a}_{P_0}^{P'}(P_0,T_0) + \Bigl(\bigl(\mathfrak{a}_{P'}^P(P',T') + \mathfrak{a}_P^{P''}(P'',T'')\bigr) \cap \mathfrak{a}_{P'}^{P''}(P,T)\Bigr)^{\sim} + \mathfrak{a}_{P''}^G,$$

Taking the union over all minimal $\mathcal{P}_0$ in $\mathcal{P}'$ of the preimages under the map $X \to N_0 \setminus X$ of the sets $\exp\bigl(D(P_0,T_0,P',T',P,T,P'',T'')\bigr)X^{P_0}$, we get the domain

$$\tilde{\tilde{X}}_{P',T'}^{P'',T''}(P,T) = P'^1 A_{P''}^G \exp\Bigl(\bigl(\mathfrak{a}_{P'}^P(P',T') + \mathfrak{a}_P^{P''}(P'',T'')\bigr) \cap \mathfrak{a}_{P'}^{P''}(P,T)\Bigr)^{\sim} K/K.$$

This domain in X is invariant under $P'^1 A_{P''}^G$, compact modulo this action, and has a smooth boundary. Thus Lemma 1 applies to it and yields a parametrix satisfying uniform estimates.

The subset $\mathfrak{a}_{P'}^P(P',T') + \mathfrak{a}_P^{P''}(P'',T'')$ of $\mathfrak{a}_{P'}^{P''}$ intersects the boundary of $\tilde{\mathfrak{a}}_{P'}^{P''}(P,T)$ in the same set in which it intersects the boundary of

$$\Bigl(\bigl(\mathfrak{a}_{P'}^P(P',T') + \mathfrak{a}_P^{P''}(P'',T'')\bigr) \cap \mathfrak{a}_{P'}^{P''}(P,T)\Bigr)^{\sim}.$$

Thus

$$\tilde{X}^{P'',T''}_{P',T'}(P,T) \cap \partial\tilde{X}(P,T) = \tilde{X}^{P'',T''}_{P',T'}(P,T) \cap \partial\tilde{\tilde{X}}^{P'',T''}_{P',T'}(P,T).$$

If $T' \ll U' \ll T \ll U'' \ll T''$, then the 2δ-neighborhood of $\tilde{X}^{P'',U''}_{P',U'}(P,T)$ in $\tilde{X}(P,T)$ is contained in $\tilde{X}^{P'',T''}_{P',T'}(P,T)$. Hence the parametrix construction is applicable to $\tilde{X}(P,T)$ as long as y remains in $\tilde{X}^{P'',U''}_{P',U'}(P,T)$, and on this subset the resulting function $H_P(t,x,y)$ coincides with the parametrix on $\tilde{\tilde{X}}^{P'',T''}_{P',T'}(P,T)$. This means that the parametrix construction works for $\tilde{X}(P,T)$, because this domain is the union of the sets $\tilde{X}^{P'',U''}_{P',U'}(P,T)$ over all the (finitely many) $\mathcal{P}''$ containing $\mathcal{P}$ and all $\mathcal{P}'$ contained in $\mathcal{P}$. Note that there are only finitely many such $\mathcal{P}'$ up to $(\Gamma_0 \cap P)$-conjugacy, and since the construction is equivariant, one can choose a uniform constant in the estimates for all of them. Hence we have these estimates for H_P, too. ■

Proof of the Proposition Let $B_P(t,x,y) := (\Delta_x + \frac{\partial}{\partial t})H_P(t,x,y)$ and define recursively

$$E^0_P(t,x,y) = H_P(t,x,y),$$
$$E^{i+1}_P(t,x,y) = \int_0^t \int_{\tilde{X}(P)} E^i_P(t-t',x,z)B_P(t',z,y)\,dz\,dt'.$$

Note that the inner integral is always over a compact subset. Thus we also get a recursive formula for $\nabla_x E^i_P(t,x,y)$ by differentiating under the integral sign. If we extend H_P, B_P to $X \times X$ by setting them equal to zero if x or y is outside $\tilde{X}(P)$, we may extend the inner integral to all of X. Trivially, the estimates of the lemma remain valid on the open dense subset $X \setminus \partial\tilde{X}(P)$. Thus the method of estimating the convolutions used in [5] applies literally and yields the asserted estimates for

$$E_P(t,x,y) = \sum_{i=0}^{\infty} (-1)^i E^i_P(t,x,y)$$

and its covariant derivative. Since H_P satisfies von Neumann boundary conditions in x, so does E_P. And since H_P is a parametrix, we obtain

$$\left(\Delta_x + \frac{\partial}{\partial t}\right) E_P(t,x,y) = 0, \quad \lim_{t\to 0} \int_{\tilde{X}(P)} E_P(t,x,y) f(y)\,dy = f(x).$$

These facts identify E_P as the heat kernel sought for. ■

3 Lattices and Injectivity Radius

In this section we provide two prerequisites for the proof of majorants for heat kernels on quotients. First we introduce a more restrictive notion of bounded depth than that defined in the introduction and derive some consequences we shall need. In the proof of the main result, Theorem 16, it will then be easy to drop this restriction. Secondly, we will estimate the injectivity radius on the quotient as a function of the lattice. To begin with the first objective, remember that $\Gamma_n(N) \subset \mathrm{GL}_n(\mathbb{Q})$ denotes the principal congruence subgroup of level N.

Definition A family $\mathcal{T}$ of subgroups of $\mathcal{G}(\mathbb{Q})$ will be called a *family of strictly bounded depth* in $\mathcal{G}(\mathbb{Q})$ if there exists a faithful $\mathbb{Q}$-rational representation $\eta\colon \mathcal{G} \to \mathrm{GL}_n$, a natural number D and, for each $\Gamma \in \mathcal{T}$, a natural number $N(\Gamma)$ such that

$$\Gamma_n\bigl(DN(\Gamma)\bigr) \cap \eta\bigl(\mathcal{G}(\mathbb{Q})\bigr) \subset \eta(\Gamma) \subset \Gamma_n\bigl(N(\Gamma)\bigr) \cap \eta\bigl(\mathcal{G}(\mathbb{Q})\bigr)$$

for all $\Gamma \in \mathcal{T}$.

It is easy to see that $[\Gamma_n(1) : \Gamma_n(N)]$ is a multiplicative arithmetic function of N, which for a prime power p^k, $k > 0$, equals $p^{kn^2} \prod_{l=1}^{n}(1 - p^{-l})$. Consequently, $[\Gamma_n(N) : \Gamma_n(DN)] \leq D^{n^2}$. This shows that any family of strictly bounded depth is a family of bounded depth. Using the representation η, we identify $\mathcal{G}$ with a $\mathbb{Q}$-subgroup of $\mathrm{GL}(n)$. Note that all $\Gamma \in \mathcal{T}$ are contained in the lattice $\Gamma_0 := \Gamma_n(1) \cap \mathcal{G}(\mathbb{Q})$.

Lemma 4 *Let $\mathcal{T}$ be a family of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ and $\mathcal{P}$ a parabolic $\mathbb{Q}$-subgroup with unipotent radical $\mathcal{N}$. Then there exist a lattice Γ_{n} in the Lie algebra n of N and a natural number D_{n} such that, for all $\Gamma \in \mathcal{T}$,*

$$\exp\bigl(D_{\mathrm{n}} N(\Gamma)\Gamma_{\mathrm{n}}\bigr) \subset \Gamma \cap N \subset \exp\bigl(N(\Gamma)\Gamma_{\mathrm{n}}\bigr).$$

If $\mathcal{M}$ is a Levi component of $\mathcal{P}$ defined over $\mathbb{Q}$, then the projections Γ^P of $\Gamma \cap P$ on M for all $\Gamma \in \mathcal{T}$ make up a family of strictly bounded depth in $\mathcal{M}$, for which one can choose $N(\Gamma^P) = N(\Gamma)$.

Proof The first assertion can be easily shown using the following fact. There exists a natural number D_n such that for all natural numbers N we have:

- If $x \in \mathrm{Mat}_n(D_n N\mathbb{Z})$ is nilpotent, then $\exp x \in \Gamma_n(N)$,
- if $g \in \Gamma_n(D_n N)$ is unipotent, then $\log g \in \mathrm{Mat}_n(N\mathbb{Z})$.

Indeed, $\exp x - 1$ and $\log g$ are given by universal polynomials in x resp. $g - 1$ of degree n with coefficients in $\mathbb{Q}$ and vanishing constant term.

Let $\mathcal{A}$ be the maximal $\mathbb{Q}$-split torus in the center of $\mathcal{M}$, Φ the set of roots of $\mathcal{A}$ in n and Π the set of weights of $\mathcal{A}$ in $V = \mathbb{Q}^n$. Then there exists an ordering on a^* such that the elements of Φ are positive. Let $\varpi_1, \dots, \varpi_r$ be the elements of Π in increasing order, V_i the weight space corresponding to ϖ_i and $W_i = \sum_{j \geq i} V_j$. Then the V_i are $\mathcal{M}(\mathbb{Q})$-stable, the flag $\{W_i\}$ is $\mathcal{P}(\mathbb{Q})$-stable, and $\mathcal{N}(\mathbb{Q})$ acts trivially in W_i/W_{i+1}.

To prove the second assertion, we use more abstract notation. Namely, if L is a lattice in a $\mathbb{Q}$-vector space V, we write $\Gamma_L(N) = \{x \in \mathrm{GL}_{\mathbb{Q}}(V) \mid (x - 1)L \subset NL\}$. Take a family $\mathcal{T}$ of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ and write L for the lattice $\mathbb{Z}^n$ in the definition. Let L_P be the direct sum of the lattices $(L + W_{i+1}) \cap V_i$. Then there are natural numbers D_P', D_P'' such that $D_P' L \subset L_P$, $D_P'' L_P \subset L$, and we put $D_P = D_P' D_P'' D$.

Let $p \in \mathcal{P}(\mathrm{Q})$ and denote by m its projection on $\mathcal{M}(\mathrm{Q})$. Then $m|_{V_i} = p|_{W_i/W_{i+1}}$. If $p \in \Gamma \cap \mathcal{P}(\mathrm{Q})$, then $(p-1)L \subset N(\Gamma)L$, hence $(m-1)L_P \subset N(\Gamma)L_P$. Conversely, if $(m-1)L_P \subset D_P N(\Gamma)L_P$, then

$$(m-1)L \subset D_P'^{-1}(m-1)L_P \subset D_P'' D N(\Gamma)L_P \subset DN(\Gamma)L.$$

We have proved

$$\Gamma_{L_P}\big(D_P N(\Gamma)\big) \cap \mathcal{M}(\mathrm{Q}) \subset \Gamma^P \subset \Gamma_{L_P}\big(N(\Gamma)\big) \cap \mathcal{M}(\mathrm{Q}),$$

as desired. ∎

Now we come to the estimate of the injectivity radius. Sometimes it is convenient to replace the distance function ρ on X by another function. Choose a K-invariant scalar product b on R^n with respect to which G is self-adjoint in the sense of [9]. Let $\|x\|_2$ denote the Hilbert-Schmidt norm of $x \in \mathrm{Mat}_n(\mathrm{R})$ with respect to b and put, for $g \in \mathrm{GL}_n(\mathrm{R})$, $\|g\| = \|g\|_2 + \|g^{-1}\|_2$. Then we have $\|gh\| \leq \|g\|\,\|h\|$ and $\|g\| \geq 2$. Therefore the function $(g,h) \mapsto \log\|g^{-1}h\|$ on $\mathrm{GL}_n(\mathrm{R}) \times \mathrm{GL}_n(\mathrm{R})$ is left-invariant under the diagonal subgroup and satisfies the triangle inequality. If we pull ρ back to G, there exists $C > 1$ such that

$$C^{-1}(\log\|g\| - 1) \leq \rho(1,g) \leq C(\log\|g\| + 1)$$

for all $g \in G$. It suffices to check this for b being the standard scalar product. By the K-biinvariance of both sides we may also suppose that $g = \mathrm{diag}(e^{t_1}, \dots, e^{t_n})$. Now $\|g\| = 2\sum_{i=1}^n \cosh t_i \leq 2\prod_{i=1}^n (e^{|t_i|} + 1) \leq 2^{n+1}\exp\sum_{i=1}^n |t_i|$. On the other hand, $\prod_{i=1}^n e^{|t_i|/2} \leq (\sum_{i=1}^n e^{|t_i|})^n \leq 2^{n-1}\|g\|^n$. Now note that the Cartan involution of $\mathrm{GL}_n(\mathrm{R})$ determined by the choice of b restricts to the Cartan involution corresponding to K, hence the two polar decompositions are compatible.

Lemma 5 *Suppose that B is a subset of G/K invariant under $\Gamma_n(N_0) \cap G$ for some natural number N_0 and such that $\Gamma_n(N_0) \cap G \setminus B$ is compact. Then there exist positive constants c, d such that for any $N \geq N_0$, any $x \in B$ and any nontrivial $\gamma \in \Gamma_n(N) \cap G$ we have $\rho(x, \gamma x) \geq c\log(N-d)$.*

Proof Since $\mathrm{Mat}_n(\mathrm{Z}) \setminus \{0\}$ is discrete in $\mathrm{Mat}_n(\mathrm{R})$, the set

$$\{(x, x^{-1}yx) \mid x \in \Gamma_n(N_0) \cap G \setminus B,\ y \in \mathrm{Mat}_n(\mathrm{Z}) \setminus \{0\}\}$$

is closed in $\big(\Gamma_n(N_0) \cap G \setminus B\big) \times \big(\mathrm{Mat}_n(\mathrm{R})\big)$ by the continuity of multiplication. But $\Gamma_n(N_0) \cap G \setminus B$ is compact, so the projection of this subset on the second component is closed. As it does not contain 0, there is some $c_1 > 0$ such that $\|x^{-1}yx\|_2 \geq c_1$ for all $x \in B$ and $y \in \mathrm{Mat}_n(\mathrm{Z}) \setminus \{0\}$. This implies $\|1 + Nx^{-1}yx\|_2 \geq c_1 N - \sqrt{n}$ for all $N \in \mathrm{N}$. Thus, for $x \in B$ and $\gamma \in \Gamma_n(N)$ we have $\|x^{-1}\gamma x\| \geq c_1 N - \sqrt{n}$. It remains to pass from $\|\,.\,\|$ to ρ. ∎

4 Metrics on Horospheres

The aim of this section is to provide estimates for the diameters and pinching constants of horospheres in $\Gamma \setminus X$ which will be needed in the next section. To begin with, we consider metrics on horospheres in the universal covering X. The restriction of the Riemannian metric of X to a smooth submanifold Y gives rise to some inner metric ρ_Y, say. If some Lie subgroup U of G acts freely on such submanifold, the latter has the structure of a principal U-bundle. In this bundle we can define a standard connection by declaring the orthogonal complements of the vertical subspaces to be the horizontal ones.

Lemma 6 *Let P be a parabolic subgroup of G with unipotent radical N. Let $U \subset R$ be normal subgroups of P, of which U is unipotent. Given $x \in X$, the standard connection in the principal U-bundle Rx has zero curvature tensor. If we denote the horizontal submanifold through x by $Y(x) = Y_R^U(x)$, then for $y_1, y_2 \in Y(x)$ and $u_1, u_2 \in U$ we have*

$$\rho_{Rx}(u_1 y_1, u_2 y_2) \geq \rho_{Rx}(y_1, y_2) = \rho_{Y(x)}(y_1, y_2)$$

with equality only for $u_1 = u_2$. Moreover, if S is a unipotent subgroup of P with $U \subset S \subset R$, then $Y_R^U(x)$ is the union of all $Y_S^U(y)$ with y running through $Y_R^S(x)$.

Thus we see that ρ_{Rx} induces a metric ρ_{Rx}^U on $U \setminus Rx$ such that the natural map $Y(x) \to U \setminus Rx$ is an isometry. We have chosen our counterintuitive placement of sub- and super-scripts in order to be in accordance with the usual notation $\mathfrak{a}_P^{P'}$, *etc.*, for objects labeled by pairs $\mathcal{P} \subset \mathcal{P}'$ of parabolic subgroups. If $\mathcal{N} \supset \mathcal{N}'$ are the corresponding unipotent radicals, we can consider $Y_N^{N'}(x)$.

Proof If we show that the standard connection in the U-bundle Px has zero curvature, the same is true for any Rx by restriction, and the remaining properties follow easily. We shall prove the vanishing of the curvature in the U-bundle Px by simply exhibiting the horizontal submanifolds $Y(x)$. Due to the transitivity of P on X, it suffices to do so for $x = eK$, the trivial coset in $X = G^1/K$.

Let θ be the Cartan involution corresponding to K and M the θ-stable Levi component of P. The Riemannian metric on G/K is determined by the restriction of some Ad-invariant θ-invariant bilinear form to the orthogonal complement $\mathfrak{s}$ of $\mathfrak{k}$ in $\mathfrak{g}$. By definition, the projections of $\mathfrak{n}$ and $\mathfrak{m}$ on $\mathfrak{s}$ are orthogonal. Let $\mathfrak{v}$ be that complement of $\mathfrak{u}$ in $\mathfrak{n}$ whose projection is orthogonal to that of $\mathfrak{u}$. Let $(X, Y) \in \mathfrak{u} \times \mathfrak{v}$ and $m \in M$. From the nilpotency of $\mathfrak{n}$ we get $\langle \mathrm{Ad}(m)X, Y \rangle = 0$ and thus

$$\langle X - \theta X, \mathrm{Ad}(m)Y - \theta\,\mathrm{Ad}(m)Y \rangle = \langle \mathrm{Ad}(\theta m^{-1})X - \theta\,\mathrm{Ad}(\theta m^{-1})X, Y - \theta Y \rangle = 0,$$

which shows that $\mathfrak{v}$ is M-invariant.

Let $V = \exp \mathfrak{v}$. The differential of the product map $U \times V \times M \to G$ at (u, v, m) maps the left translate of $(X, Y, Z) \in \mathfrak{u} \times \mathfrak{v} \times \mathfrak{m}$ to the left translate of $\mathrm{Ad}(vm)^{-1}X + \mathrm{Ad}(m)^{-1}Y + Z$. (Here, left translation on V is meant with respect to the identification $V \cong U \setminus N$.) This shows that the Jacobian determinant of the product map is nonzero, whence $Y(eK) := VMK/K$ is a submanifold, and that $Y(eK)$ intersects each U-orbit orthogonally. ■

Lemma 7 *There exists a constant $C > 1$ with the following property. If $U \subset R$ are unipotent subgroups of a parabolic subgroup P of G, then for all $x \in X$ and $y \in Y(x)$ we have*

$$C^{-1}\rho(x,y) \le \log\bigl(1 + \rho_{Y(x)}(x,y)\bigr) \le C\rho(x,y).$$

Proof Since there are only finitely many parabolics P up to G^1-conjugacy, it suffices to prove the lemma for fixed P. Due to the transitivity of P on X, we may again suppose that x is the trivial coset. As we saw in the proof of Lemma 6, $Y(eK) = \exp(\mathfrak{v})K/K$, where $\mathfrak{v}$ is a complementary subspace of $\mathfrak{u}$ in $\mathfrak{r}$. Since $U \setminus R$ is abelian, the exponential map $\mathfrak{v} \to U \setminus R$ is an isomorphism. Thus

$$\rho_{Y(eK)}\bigl(eK, \exp(X)K\bigr) = |X|$$

for $X \in \mathfrak{v}$, where $|\,.\,|$ is the Euclidean norm on $\mathfrak{n}$ coming from the pull-back of the bilinear form on $\mathfrak{s}$ considered in the proof of Lemma 6. Since $\exp X$ for $X \in \mathfrak{n}$ and $\log x$ for $x \in N$ are given by polynomials of degree at most n, we have

$$C_1^{-1}|X|^{1/n} \le \|\exp X\| \le C_1(1 + |X|)^n$$

for some $C_1 > 1$. Passing from $\|\,.\,\|$ to ρ, we obtain the assertion for $\rho(eK, y) \ge C_2$. However, in a compact neighborhood of the trivial element in $\mathfrak{n} \cong N$, the Euclidean metric on $\mathfrak{n}$ is equivalent to the restriction of ρ to N. ■

Corollary 8 *In the same situation, we have*

$$C^{-1}\rho^U(x,y) \le \log\bigl(1 + \rho^U_{Rx}(x,y)\bigr) \le C\rho^U(x,y).$$

for $y \in U \setminus Rx$.

Indeed, we may choose for y the representative in $Y(x) = Y^U_R(x) \subset Y^U(x)$, and then $\rho^U_{Rx}(x,y) = \rho_{Y(x)}(x,y)$ and $\rho^U(x,y) = \rho(x,y)$ by Lemma 6.

Before considering metrics on horospheres in $\Gamma \setminus X$, we introduce some notation. If X is a metric space and U, Γ are groups of isometries of X such that Γ normalizes U, we define

$$\operatorname{pinch}(U \setminus U\Gamma x) = \inf\{\rho(x, u\gamma x) \mid u \in U, \gamma \in \Gamma, \gamma \notin U\}.$$

This is the minimal distance between the Γ-translates of the coset of x in $U \setminus X$. If ρ comes from a Riemannian metric and $U \setminus X \to U\Gamma \setminus X$ is a universal covering of a manifold, then $\operatorname{pinch}(U \setminus U\Gamma x)$ is the minimal length of a non-contractible loop in $U\Gamma \setminus X$ with basepoint x.

Let us return to our previous notation. If $\mathcal{P}$ is a parabolic $\mathbb{Q}$-subgroup of $\mathcal{G}$, we define, as usual, $H_P \colon X \to \mathfrak{a}_P^G$ by requiring the equality $\exp H_P(pK) = P^1 p$ in A_P for all $p \in P$. Due to the definition of a truncation parameter, we have $H_{P'}(\gamma x) - T_{P'} = \operatorname{Ad}(\gamma)\bigl(H_P(x) - T_P\bigr)$ for $\gamma \in \Gamma_0$ and $\mathcal{P}' = \gamma\mathcal{P}\gamma^{-1}$.

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

Lemma 9 *Given a family of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$, there exists a constant $C > 1$ with the following property. Let $\mathcal{P}$ be any parabolic $\mathbb{Q}$-subgoup of $\mathcal{G}$ with $\mathbb{Q}$-split component $\mathcal{A}$, and let $\mathcal{U} \subset \mathcal{R}$ be normal unipotent $\mathbb{Q}$-subgroups of $\mathcal{P}$ such that A acts on $\mathrm{u} \setminus \mathrm{r}$ by a root α. Then for all $\Gamma \in \mathcal{T}$ and $x \in X(P)$ we have, with respect to the metric ρ^U_{Rx},*

$$\operatorname{diam}^U_{Rx}\bigl(U(\Gamma \cap R) \setminus Rx\bigr) \le C N(\Gamma) e^{-\alpha(H_P(x)-T_P)},$$
$$\operatorname{pinch}^U_{Rx}\bigl(U \setminus U(\Gamma \cap R)x\bigr) \ge C^{-1} N(\Gamma) e^{-\alpha(H_P(x)-T_P)}.$$

Proof First we consider fixed $\mathcal{P}$ only. Write $x = pK$ with $p \in P$. As we have seen in the proof of Lemma 6, there is a complementary subspace v of u in r stable under some Levi component M of P such that $Y(x) = p\exp(\mathrm{v})K/K$. Since A acts on $\mathrm{u} \setminus \mathrm{r}$ by scalars, $U \setminus R$ is abelian, and the map $\varphi\colon \mathrm{v} \to U \setminus Rx$ given by $\varphi(X) = p\exp(X)K$ is an isometry for some Euclidean metric on v. Moreover, $\exp(X_1)\varphi(X_2) = \varphi\bigl(\mathrm{Ad}(m^{-1})X_1 + X_2\bigr)$ for $X_1, X_2 \in \mathrm{v}$, where m denotes the projection of p on M. By Lemma 4,

$$D_{\mathrm{n}} N(\Gamma)\,\mathrm{Ad}(m^{-1})\Gamma_{\mathrm{v}} \subset \varphi^{-1}\bigl(U \setminus U(\Gamma \cap R)\bigr) \subset N(\Gamma)\,\mathrm{Ad}(m^{-1})\Gamma_{\mathrm{v}},$$

where $\Gamma_{\mathrm{v}} = (\Gamma_{\mathrm{n}} + \mathrm{u}) \cap \mathrm{r}$. After conjugating M and v, if necessary, by an element of P, we may suppose that $M = \mathcal{M}(\mathbb{R})$, where $\mathcal{M}$ is the centralizer of $\mathcal{A}$ and, in particular, defined over $\mathbb{Q}$. As A acts on v via α, it remains to show that

$$\sup_{\substack{m \in M^1\\ mK^P \in X^P(M)}} \operatorname{diam}\bigl(\mathrm{Ad}(m^{-1})\Gamma_{\mathrm{v}} \setminus \mathrm{v}\bigr) < \infty, \qquad \inf_{\substack{m \in M^1\\ mK^P \in X^P(M)}} \operatorname{pinch}\bigl(\mathrm{Ad}(m^{-1})\Gamma_{\mathrm{v}}\bigr) > 0.$$

with respect to our Euclidean metric on v. This metric is stable under K^P, so the functions of which we take sup resp. inf really depend only on $mK^P \in X^P \cong M^1/K^P$. These functions are continuous, nonvanishing and invariant under the stabilizer of Γ_{n} in M. Since Γ_{n} lies in the set of $\mathbb{Q}$-rational points of n, it is easy to find an arithmetic subgroup Γ_M of $\mathcal{M}(\mathbb{Q})$ contained in Γ^P_0 and stabilizing Γ_{n}. Then $\Gamma_M \setminus X^P(M)$ is compact, and the assertion for fixed $\mathcal{P}$ follows.

Note that the preceding argument could have been reduced to the case of $\Gamma = \Gamma_n(N) \cap \mathcal{G}(\mathbb{Q})$. This is a normal subgroup of Γ_0, for which both sides of the asserted inequalities are unchanged if we replace $\mathcal{P}$ by one of its Γ_0-conjugates. Since there are only finitely many Γ_0-conjugacy classes of parabolic $\mathbb{Q}$-subgroups, C can be chosen independently of $\mathcal{P}$. ■

5 Majorants on Quotients

In this section we shall prove the main technical estimate, which will provide majorants for the heat kernels on the quotient spaces $\Gamma \cap P \setminus \tilde{X}(P, T)$. For this one has, in particular, to sum the inverse of the distance function from a fixed point $x \in X$ over the lattice points in the intersection of the horosphere Nx with a ball B in X. Basically we follow the usual method of bounding such a sum by the integral over the intersection of Nx with a larger ball B'. Naively, one would replace the values in the lattice points by the integrals over disjoint neighborhoods divided by their common volume. To make the estimates strong enough, we would like to maximize the volume of these neighborhoods by setting them equal to the translates of a fundamental domain for the action of $\Gamma \cap N$. However, such fundamental domains depend on Γ, hence so does B' in an uncontrollable way. We overcome this difficulty (and other ones) by decomposing $\Gamma \cap N$ according to a filtration of N depending on x and by adapting the fundamental domains to this filtration.

To begin with, we prepare the notation for defining the fundamental domains. Let $\mathcal{P}$ be a parabolic $\mathbb{Q}$-subgroup of $\mathcal{G}$, $\mathcal{N}$ its unipotent radical and $\mathcal{A}$ a $\mathbb{Q}$-split component. Given $x \in X(P,T)$, we write the set Φ of roots of $\mathcal{A}$ in $\mathfrak{n}$ as $\{\alpha_1, \ldots, \alpha_l\}$ in such a way that $\alpha_i\big(H_P(x) - T_P\big) \leq \alpha_{i+1}\big(H_P(x) - T_P\big)$. Since $H_P(x) - T_P \in \mathfrak{a}_P^+$, the sum of the root subspaces for the roots $\alpha_i, \ldots, \alpha_l$ is a normal Lie subalgebra for each i. Let $\mathcal{N}_i$ be the corresponding normal unipotent subgroup of $\mathcal{P}$. For unification, we write $\mathcal{N}_0 = \mathcal{P}$, $\mathcal{N}_{l+1} = \{1\}$. The restriction of the Riemannian metric of X to $N_i x$ defines some inner metric, which we denote by ρ_i. In particular, $\rho_0 = \rho$. If $y \in N_i x$ and $j > i$, let $y_j(x)$ be the element in $N_j y$ of minimal ρ_i-distance from x. This element is unique by Lemma 6, because it is the result of the horizontal transport of x from the fiber $N_j x$ to the fiber $N_j y$ with respect to the standard connection on the N_j-bundle $N_i x$. The last assertion of Lemma 6 shows that, if we replace i by any $k < i$, we get the same result. That is why we have not included i in the notation $y_j(x)$. Another consequence is that $y_j\big(y_i(x)\big) = y_j(x)$ for $j > i$. The metric ρ_i^j induced by ρ_i on $N_j \setminus N_i x$ is characterized by $\rho_i^j(x,y) = \rho_i\big(x, y_j(x)\big)$.

For each i, we set

$$\tilde{F}_i^\Gamma(x) = \{y \in N_{i+1} \setminus N_i x \mid \rho_i^{i+1}(y,x) < \rho_i^{i+1}(y,\gamma x) \qquad \forall\, \gamma \in \Gamma \cap N_{i+1} \setminus \Gamma \cap N_i \text{ with } \gamma \neq 1\}.$$

This is a fundamental domain for the action of $\Gamma \cap N_{i+1} \setminus \Gamma \cap N_i$ on $N_{i+1} \setminus N_i x$. Now we define recursively $F_{l+1}^\Gamma(x) = \{x\}$,

$$F_i^\Gamma(x) = \bigcup_{y \in \tilde{F}_i^\Gamma(x)} F_{i+1}^\Gamma\big(y_{i+1}(x)\big).$$

for $1 \leq i \leq l$. Clearly, $\gamma F_i^\Gamma(x) = F_i^\Gamma(\gamma x)$ for $\gamma \in \Gamma \cap N_i$. Finally, we put $F_N^\Gamma(x) = F_1^\Gamma(x)$. Of course, this construction depends on the choice of the filtration $\{\mathcal{N}_i\}$, which is not unique for some x.

Lemma 10 *The set $F_i^\Gamma(x)$ is a fundamental domain for the action of $\Gamma \cap N_i$ on $N_i x$. Moreover, there exists $C > 0$ with the following properties for all parabolic $\mathbb{Q}$-subgroups $\mathcal{P}$, all $x \in X(P,T)$, $i = 1, \ldots, l$ and $\Gamma \in \mathcal{T}$:*

(i) If $y \in N_{i+1} \setminus N_i x$ is in the closure of $\tilde{F}_i^\Gamma(x)$ and $\gamma \in \Gamma \cap N_{i+1} \setminus \Gamma \cap N_i$ is nontrivial, then

$$\rho^{i+1}(x, \gamma y) \geq C^{-1} \log\big(1 + N(\Gamma) e^{-\alpha_i(H_P(x) - T_P)}\big).$$

(ii) With respect to the metric ρ we have

$$\operatorname{diam}\big(F_i^\Gamma(x)\big) \leq C \log\big(1 + N(\Gamma) e^{-\alpha_i(H_P(x) - T_P)}\big).$$

(iii) Any two elements of the closure of $F_i^\Gamma(x)$ can be connected in this set by a piecewise smooth path whose length is at most $C N(\Gamma) e^{-\alpha_i(H_P(x) - T_P)}$.

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

Proof If $\gamma \in \Gamma \cap N_i$ and $y \in F_i^\Gamma(x) \cap \gamma F_i^\Gamma(x)$, we see by projecting on $N_{i+1} \setminus N_i x$ that $\tilde{F}_i^\Gamma(x) \cap \gamma \tilde{F}_i^\Gamma(x) \neq \varnothing$, hence $\gamma \in \Gamma \cap N_{i+1}$. Now $y_{i+1}(\gamma x) = \gamma y_{i+1}(x)$, and we get $F_{i+1}^\Gamma\big(y_{i+1}(x)\big) \cap F_{i+1}^\Gamma\big(\gamma y_{i+1}(x)\big) \neq \varnothing$ by intersecting with $N_{i+1} y$. This shows by induction that the $(\Gamma \cap N_i)$-translates of $F_i^\Gamma(x)$ are disjoint.

Suppose we have already shown that the $(\Gamma \cap N_{i+1})$-translates of the closure of $F_{i+1}^\Gamma(x)$ cover $N_{i+1}x$ for any x. Let $y \in N_i x$. Then it is easy to see that there exists $\gamma \in \Gamma \cap N_i$ such that $\gamma y_{i+1}(x) = (\gamma y)_{i+1}(x)$ is in the closure of $\tilde{F}_i^\Gamma(x)$ modulo N_{i+1}. By the induction hypothesis there exists $\gamma' \in \Gamma \cap N_{i+1}$ such that $\gamma'\gamma y$ is in the closure of $F_{i+1}^\Gamma\big(\gamma y_{i+1}(x)\big)$. This means that $\gamma'\gamma y$ is in the closure of $F_i^\Gamma(x)$.

(i) For y and γ as in the statement we have by definition

$$\rho_i^{i+1}(x, \gamma x) \leq \rho_i^{i+1}(x, \gamma y) + \rho_i^{i+1}(\gamma x, \gamma y) \leq 2\rho_i^{i+1}(x, \gamma y).$$

The left-hand side is bounded from below by $\operatorname{pinch}_i^{i+1}\big(N_{i+1} \setminus N_{i+1}(\Gamma \cap N_i)x\big)$, for which we have the lower bound from Lemma 9. It remains to use the resulting inequality together with Corollary 8.

(ii) It suffices to bound $\rho(x, y)$ for $y \in F_i^\Gamma(x)$. It is clear that $x = y_i(x)$ and $y = y_{l+1}(x)$. Since $y_{j+1}(x) \in \tilde{F}_j^\Gamma\big(y_j(x)\big)$ for $j = i, \dots, l$, we have

$$\begin{aligned} \rho_j\big(y_{j+1}(x), y_j(x)\big) &= \rho_j^{j+1}\big(y_{j+1}(x), y_j(x)\big) \\ &= \inf_{\gamma \in \Gamma \cap N_j} \rho_j^{j+1}\big(y_{j+1}(x), \gamma y_j(x)\big) \\ &\leq \operatorname{diam}_j^{j+1}\big(N_{j+1}(\Gamma \cap N_j) \setminus N_j y_j(x)\big). \end{aligned}$$

Lemma 9 provides a bound for this diameter, and Lemma 7 shows that

$$\rho\big(y_{j+1}(x), y_j(x)\big) \leq C \log\Big(1 + \rho_j\big(y_{j+1}(x), y_j(x)\big)\Big).$$

Now the assertion follows from the triangle inequality and the choice of our filtration.

(iii) It suffices to connect x with any $y \in F_i^\Gamma(x)$ by a path with the required properties. For $j \geq i$, let $Y_j^{j+1}(x)$ be the horizontal submanifold through x in the N_{j+1}-bundle $N_j x$. This is a Euclidean space projecting isometrically on $N_{j+1} \setminus N_j x$. By definition, $y_{j+1}(x)$ lies in the inverse image of $\tilde{F}_j^\Gamma\big(y_j(x)\big)$ in $Y_j^{j+1}\big(y_j(x)\big)$. Since this is an intersection of half-spaces in $Y_j^{j+1}\big(y_j(x)\big)$, it contains the straight line connecting $y_j(x)$ with $y_{j+1}(x)$, whose length equals $\rho_j^{j+1}\big(y_j(x), y_{j+1}(x)\big)$. Applying the bound on $\operatorname{diam}_j^{j+1}\big(N_{j+1}(\Gamma \cap N_j) \setminus N_j y_j(x)\big)$ from Lemma 9, we get the assertion. ■

Our next result contains an analogue of Lemma 4.2 in [5] adapted to our situation as well as a generalization of Lemma 3.3 in [6] which takes care of the dependence on the lattice.

Proposition 11 *Given a family $\mathcal{T}$ of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ and a positive number r', there exist positive constants C, c such that for all $\Gamma \in \mathcal{T}$ with $N(\Gamma)$ large enough, all proper* *parabolic* $\mathbb{Q}$*-subgoup* $\mathcal{P}$*, all* $x \in X(P)$*,* y *in the closure of* $F_N^\Gamma(x)$ *and* $r \geq 1$ *we have*

$$\#\{\gamma \in \Gamma \cap P \mid \gamma \neq 1, \rho(x, \gamma y) \leq r\} \leq CN(\Gamma)^{-1} e^{2\rho_P(H_P(x) - T_P)} e^{cr},$$

$$\sum_{\substack{\gamma \in \Gamma \cap N \\ \gamma \neq 1 \\ \rho(x,\gamma y) \leq r'}} \rho(x, \gamma y)^{-1} \leq CN(\Gamma)^{-1} e^{2\rho_P(H_P(x) - T_P)} \alpha_P(x).$$

Here we denote, as usual, by ρ_P *the half-sum of roots with multiplicities. Moreover, we set* $\alpha_P(x) = \min_{\alpha \in \Phi}\Big(1 + \alpha\big(H_P(x) - T_P\big)\Big)$*, where* Φ *is the set of roots of* $\mathfrak{a}$ *in* $\mathfrak{n}$*. (This factor may be omitted if the minimum is obtained at* α *with* $\dim \mathfrak{n}_\alpha > 1$*.) If* $\mathcal{P} = \mathcal{G}$*, then the first estimate is still true if one omits the factor* $N(\Gamma)^{-1}$*.*

Proof Let us fix a filtration $\{\mathcal{N}_1, \ldots, \mathcal{N}_l\}$ adapted to x as above and let Γ_i be the set-theoretic difference of $\Gamma \cap N_i$ and $\Gamma \cap N_{i+1}$. We shall prove the estimates for each Γ_i and for the remaining subset of $\Gamma \cap P$ separately.

Fix $1 \leq i \leq l$ and let y be in the closure of $F_N^\Gamma(x)$, $\gamma \in \Gamma_i$. We may apply Lemma 10(i) to $N_{i+1}y$, which is in the closure of $\tilde{F}_i^\Gamma\big(y_i(x)\big)$, concluding that there exists $\varepsilon > 0$ such that

$$\rho^{i+1}\big(y_i(x), \gamma y\big) \geq \varepsilon \log\big(1 + N(\Gamma) e^{-\alpha_i(H_P(x) - T_P)}\big) =: p_i^\Gamma(x).$$

In the case that $\rho^{i+1}\big(x, y_i(x)\big) \geq \frac{1}{2} p_i^\Gamma(x)$ we get

$$\rho^{i+1}(x, \gamma y) \geq \rho^{i+1}\big(x, y_i(x)\big) \geq \frac{1}{2} p_i^\Gamma(x),$$

while in the case that $\rho^{i+1}\big(x, y_i(x)\big) \leq \frac{1}{2} p_i^\Gamma(x)$ we get

$$\rho^{i+1}(x, \gamma y) \geq \rho^{i+1}\big(y_i(x), \gamma y\big) - \rho^{i+1}\big(y_i(x), x\big) \geq p_i^\Gamma(x) - \frac{1}{2} p_i^\Gamma(x).$$

Therefore, in any case we have

$$p_i^\Gamma(x) \leq 2\rho^{i+1}(x, \gamma y) \leq 2\rho(x, \gamma y).$$

Now let $z \in F_i^\Gamma\big(\gamma y_i(x)\big) = \gamma F_i^\Gamma\big(y_i(x)\big) = \gamma F_N^\Gamma(x) \cap N_i y$. Then

$$\rho\big(y_i(x), z\big) \leq \rho\big(y_i(x), x\big) + \rho(x, \gamma y) + \rho(\gamma y, z).$$

The first term equals $\rho^i(\gamma y, x)$ and is bounded by $\rho(x, \gamma y)$, while the last term is bounded by $\operatorname{diam}\Big(F_i^\Gamma\big(\gamma y_i(x)\big)\Big)$. It follows from Lemma 10(ii) that there is a constant $C > 0$ independent of $i > 1$, $x \in X(P)$, the parabolic $\mathcal{P}$ and $\Gamma \in \mathcal{T}$ such that

$$\operatorname{diam}\big(F_i^\Gamma(x)\big) \leq C p_i^\Gamma(x).$$

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

Applying this and the bound on $p_i^\Gamma(x)$ just derived, we get

$$\rho\big(y_i(x), z\big) \le 2(1+C)\rho(x, \gamma y).$$

On the one hand this shows that

$$\bigcup_{\substack{\gamma\in\Gamma_i \\ \rho(x,\gamma y)\le r}} F_i^\Gamma\big(\gamma y_i(x)\big) \subset B_i\big(y_i(x), cr\big) \tag{1}$$

(disjoint union by Lemma 10), where $c = 2(1+C)$ and $B_i(y, r)$ denotes the ball of radius r around y in $N_i y$ with respect to the restriction of ρ. On the other hand, it implies that

$$\sum_{\substack{\gamma\in\Gamma_i \\ \rho(x,\gamma y)\le r}} \mathrm{vol}_i\Big(F_i^\Gamma\big(\gamma y_i(x)\big)\Big)\rho(x,\gamma y)^{-1} \le c \int_{B_i(y_i(x),cr)} \rho\big(y_i(x), z\big)^{-1}\, dz, \tag{2}$$

where dz is the Riemannian measure on $N_i y$ coming from ρ_i and vol_i the corresponding volume.

For fixed $r = r'$, one can prove more. Lemma 10(i) can be applied to the N_{i+1}-orbit of $z \in F_i^\Gamma\big(\gamma y_i(x)\big)$ for $\gamma \in \Gamma_i$ and gives

$$\rho\big(y_i(x), z\big) \ge \rho^{i+1}\big(y_i(x), z\big) \ge p_i^\Gamma(x).$$

If $\rho(x, \gamma y) \le r'$, we may replace the logarithm in $p_i^\Gamma(x)$ by a linear function and get

$$\rho\big(y_i(x), z\big) \ge \varepsilon' N(\Gamma) e^{-\alpha_i\left(H_P(x) - T_P\right)}$$

for some $\varepsilon' > 0$ independent of i, x, Γ and $\mathcal{P}$. Indeed, by the above,

$$p_i^\Gamma(x) \le \rho\big(y_i(x), z\big) \le c\rho(x, \gamma y) \le cr',$$

which was fixed. The upshot is that we may actually replace the domain of integration in (2) by the set-theoretic difference of the balls $B_i\big(y_i(x), cr'\big)$ and $B_i\big(y_i(x), \varepsilon' N(\Gamma) e^{-\alpha_i(H_P(x))}\big)$.

So we have to estimate integrals of the type

$$\int_{B_i(x,r')\setminus B_i(x,\varepsilon')} \rho(x, z)^{-1}\, dz$$

for fixed r' and varying $\varepsilon' > 0$, i, x and P. Such an integral remains unchanged if we replace x and N_i by gx and gN_ig^{-1}, resp., for some $g \in G^1$. Thus we may fix P and suppose that $x = eK$. If we pull ρ back under the embedding $\exp\colon \mathfrak{n}_i \to N_i \to N_iK/K \subset X$, its restriction to a compact subset is equivalent to the Euclidean metric on $\mathfrak{n}_i$. Hence the integral is bounded independently of ε' for $\dim N_i > 1$. Now suppose that $\dim N_i = 1$. After passing to the Euclidean metric, the integral evaluates as $c' \max\big(1 + \log(r'/\varepsilon'), 0\big)$ with $c' > 0$. Using this and the fact that $N(\Gamma) > 1$, the right-hand side of (2) for fixed

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

$r = r'$ can be estimated by $C'\Big(1 + \alpha_i\big(H_P(x) - T_P\big)\Big)$ for some $C' > 0$ independent of i, x and P. If $i > 1$, this term will be absorbed by the terms with smaller i.

In the case $i = 0$, we lack a bound for the diameter of a fundamental domain for Γ^P in $X^P(M)$. Therefore, we consider

$$F_0^\Gamma(x) = \bigcup_{y \in B^0(x,r_0)} F_1^\Gamma\big(y_1(x)\big)$$

instead, where $B^0(x,r_0)$ denotes the ball of radius r_0 around x in $N \setminus X$. We may replace $\mathcal{T}$ by the subset of those Γ for which $N(\Gamma)$ is greater than a given number. Then Lemmas 4 and 5 yield a constant $r_0 > 0$ such that $\operatorname{pinch}(\Gamma^P x_0) \geq r_0\big(1 + \log N(\Gamma)\big)$ for all $x_0 \in X^P(M)$ and all $\Gamma \in \mathcal{T}$. For this choice of r_0, it is easy to show as in the proof of Lemma 10 that the sets $F_0^\Gamma(\gamma x)$ for $\gamma \in \Gamma \cap P$ are pairwise disjoint. Moreover, it is clear that there exists $C_0 > 0$ such that $\operatorname{diam}\big(F_0^\Gamma(x)\big) \leq C_0\big(1 + \log N(\Gamma)\big)$. Replacing Γ by $\Gamma_n\big(N(\Gamma)\big) \cap \mathcal{G}(\mathbb{Q})$ decreases r_0 while replacing it by $\Gamma_n\big(DN(\Gamma)\big) \cap \mathcal{G}(\mathbb{Q})$ increases C_0. Since the principal congruence subgroups are normal in Γ_0, we see as above that r_0 and C_0 can be chosen independently of $\mathcal{P}$.

If $\gamma \in \Gamma \cap P$, $\gamma \notin \Gamma \cap N$ and $y \in Nx$ with $\rho(x, \gamma y) \leq r$, then $r \geq r_0\big(1 + \log N(\Gamma)\big)$. If, moreover, $z \in F_0^\Gamma(\gamma y)$, then $\rho(x,z) \leq \rho(x,\gamma y) + \rho(\gamma y, z) \leq r + \operatorname{diam}\big(F_0^\Gamma(y)\big) \leq c_0 r$ for $c_0 = 1 + C_0/r_0$. Therefore

$$\bigcup_{\substack{\gamma \in \Gamma \cap P \\ \gamma \notin \Gamma \cap N \\ \rho(x,\gamma y) \leq r}} F_0^\Gamma(\gamma y) \subset B_0(x, c_0 r), \tag{3}$$

where $B_0(x,r)$ denotes the ball of radius r around x in X with respect to ρ.

To obtain the asserted inequalities from (1), (2) and (3), it remains to calculate the Riemannian volume of $F_i^\Gamma(x)$ and $B_i(x,r)$ with respect to ρ_i. The push-forward of the Haar measure on N_i under the map $N_i \to N_i x$ is proportional to this Riemannian measure, since both are invariant. If we pull the Riemannian metric g back under $P \to G/K = X$, we have $g_{px} = g_{x^{-1}px}$. The inner automorphism $p \to x^{-1}px$ of P scales the invariant measure on N_i by $e^{-2\rho_{N_i}(H_P(x))}$, where $\rho_{N_i}\big(H_P(x) - T_P\big) \leq \rho_P\big(H_P(x) - T_P\big)$, and leaves that on $N \backslash P$ unchanged. Thus we get

$$\operatorname{vol}_i\big(F_i^\Gamma(x)\big) = \operatorname{vol}(\Gamma \cap N_i \setminus N_i x) = C' \operatorname{vol}(\Gamma \cap N_i \setminus N_i) e^{-2\rho_{N_i}(H_P(x) - T_P)}$$

for $i \geq 1$. We see that $\operatorname{vol}\big(F_i^\Gamma(x)\big)$ is left N_i-invariant. Thus

$$\operatorname{vol}\big(F_0^\Gamma(x)\big) = C_0' \operatorname{vol}(\Gamma \cap N \setminus N) \operatorname{vol}\big(B^0(x_0, r_0)\big) e^{-2\rho_P(H_P(x) - T_P)}.$$

Note that by Lemma 4 the volume of $\Gamma \cap N_i \setminus N_i$ grows at least linearly in $N(\Gamma)$ (except in the case $\mathcal{P} = \mathcal{G}$). The usual conjugacy argument shows that the lower bound we get is uniform in $\mathcal{P}$. Finally, we have to show that $\operatorname{vol}\big(B_i(x,r)\big)$ grows at most exponentially in r. Using Lemma 6, one reduces this to the analogous assertions about balls in M/K^P and $N_{i+1} \setminus N_i$, which are well known. ■

Corollary 12 *Given a family $\mathcal{T}$ of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ and a constant $c > 0$, there exists a constant $C > 0$ such that for all $\Gamma \in \mathcal{T}$ with $N(\Gamma)$ large enough, all parabolic $\mathbb{Q}$-subgoups $\mathcal{P}$, all $x \in X(P)$ and y in the closure of $F_N^\Gamma(x)$*

$$\sum_{\substack{\gamma \in \Gamma \cap P \\ \gamma \neq 1}} \big(\rho(x, \gamma y)^{-1} + 1\big) e^{-c\rho(x,\gamma y)^2} \leq C N(\Gamma)^{-1} e^{2\rho_P(H_P(x) - T_P)} \alpha_P(x).$$

Indeed, by Lemmas 4 and 5 there exists $r_0 > 0$ such that $\operatorname{pinch}\big(N \setminus N(\Gamma \cap P)x\big) \geq r_0$ for all $\Gamma \in \mathcal{T}$ with $N(\Gamma)$ large enough. As we just remarked in the proof, r_0 can be chosen independently of $\mathcal{P}$. Now we follow [6], p. 246: In the terms with $\rho(x, \gamma y) < r_0$ we replace the exponential factor by 1 and apply the second estimate, while in the remaining terms we replace $\rho(x, \gamma y)^{-1}$ by r_0^{-1} and apply the first estimate and the argument from [5], p. 491.

6 Heat Kernels on Quotients

We are now in a position to construct and estimate the heat kernels on the quotients $\Gamma \cap P \setminus \tilde{X}(P)$, which will allow us to prove the main result, *viz.* the spectral estimates for the cut-off Laplacian.

Since E_τ is a homogeneous bundle, G^1 also acts on its sections. However, we now prefer to view sections of E_τ as functions $f\colon G^1 \to V_\tau$ satisfying $f(xk) = \tau(k^{-1}) f(x)$ for all $x \in G^1$ and $k \in K$. Then the aforementioned action is simply given by left translation. Let $\tilde{G}(P)$ be the inverse image of $\tilde{X}(P)$ in G^1. Then we may view $E_P(t, x, y)$ as an $\operatorname{End}(V_\tau)$-valued function on $\mathbb{R}_+ \times \tilde{G}(P) \times \tilde{G}(P)$. For $x, y \in \tilde{G}(P)$ and $t > 0$, put

$$E_P^\Gamma(t, x, y) = \sum_{\gamma \in \Gamma \cap P} E_P(t, x, \gamma y).$$

By Proposition 2 and the results of [5], this sum converges absolutely on compact subsets of $\mathbb{R}_+ \times \tilde{G}(P) \times \tilde{G}(P)$, and the same is true after termwise covariant differentiation. E_P^Γ is the heat kernel on $\Gamma \cap P \setminus \tilde{X}(P)$ with coefficients in the bundle $E_\tau^{\Gamma \cap P}$ and von Neumann boundary conditions, since it has all its characteristic properties.

Lemma 13 *For any family $\mathcal{T}$ of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ there exists a constant $C > 0$ with the following property. Let $\mathcal{P} \subset \mathcal{P}'$ be parabolic $\mathbb{Q}$-subgoups, of which $\mathcal{P}'$ is maximal, with unipotent radicals $\mathcal{N} \supset \mathcal{N}'$. Let $\mathcal{A}$ be a $\mathbb{Q}$-split component of $\mathcal{P}$. Then for all $\Gamma \in \mathcal{T}$ with $N(\Gamma)$ large enough, all $x \in \tilde{G}(P)$, all $y, z \in N'x$ and $0 < t \leq 1$ we have*

$$|E_P^\Gamma(t, x, y) - E_P^\Gamma(t, x, z)| \leq C t^{-d/2} \big(1 + e^{(2\rho_P - \alpha)(H_P(x) - T_P)} \alpha_P(x)\big),$$

where α_P is as in Proposition 11 and α is the only fundamental root of $\mathcal{A}$ in $\mathfrak{n}$ which does not vanish on $\mathfrak{a}_{P'}$.

Proof Choose an ordering $\{\alpha_1, \ldots, \alpha_l\}$ of the roots of $\mathcal{A}$ in $\mathfrak{n}$ adapted to xK, which gives rise to a filtration $\{\mathcal{N}_i\}$ of $\mathcal{N}$ and the fundamental domains $F_i^\Gamma(x)$ of $\Gamma \cap N_i$ in $N_i x K/K$. If i is that number with $\alpha_i = \alpha$, then the other roots of $\mathcal{A}$ in $\mathfrak{n}'$ are of the form α_j with $j > i$,

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

in other words, $\mathcal{N}' \subset \mathcal{N}_i$. Thus $y, z \in N_i x$, and since $E_P^\Gamma(t,x,y)$ is $(\Gamma \cap N_i)$-invariant, we can choose yK, zK in the closure of $F_i^\Gamma(x)$. By Lemma 10(iii), we can connect them inside this closure by a piecewise smooth path c whose length is bounded by a constant times $N(\Gamma)e^{-\alpha(H_P(x)-T_P)}$. Let $\tilde{c}$ be the lift of this path which connects y and z in $N_i x \subset \tilde{G}(P)$. We have

$$E_P^\Gamma(t,x,y) - E_P^\Gamma(t,x,z) = \sum_{\gamma\in\Gamma\cap P} \big(E_P(t,x,\gamma y) - E_P(t,x,\gamma z)\big).$$

Due to the first estimate in Proposition 2, the term with $\gamma = 1$ is bounded by $C_1 t^{-d/2}$ for some $C_1 > 0$. We write the other terms as

$$E_P(t,x,\gamma y) - E_P(t,x,\gamma z) = \int_{\tilde{c}} \tilde{X}_u E_P(t,x,\gamma u)\, du,$$

where the measure du is defined by the normal parametrization of the path $c \subset X$ transferred to $\tilde{c}$ and $\tilde{X}_u$ is the tangent vector of $\tilde{c}$ at the point u according to this parametrization.

For any smooth function $f\colon \tilde{G}(P) \to V_\tau$ satisfying $f(xk) = \tau(k^{-1})f(x)$ we have

$$\tilde{X}_u f(u) = \frac{d}{dt} f\big(\exp(tX)u\big)\Big|_{t=0} = \frac{d}{dt} f\Big(u\exp\big(t\,\mathrm{Ad}(u^{-1})X\big)\Big)\Big|_{t=0}$$

for some $X \in \mathfrak{n}$. If we decompose $\mathrm{Ad}(u^{-1})X = Y = Y^{\mathfrak{k}} + Y^{\mathfrak{s}}$ according to the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}$, then the component of $\tilde{X}_u$ corresponding to $Y^{\mathfrak{s}}$ is just the tangent vector X_u of c, and

$$X_u f(u) = \nabla_{X_u} f(u) - \tau(Y^{\mathfrak{k}}) f(u),$$

where f is interpreted as a section of E_τ in the first term. Suppose for a moment that $u \in P$; then $Y \in \mathfrak{n}$. Since the projection $\mathfrak{n} \to \mathfrak{s}$ is injective, there exists $C_2 > 0$ such that $|Y^{\mathfrak{k}}| \le C_2 |Y^{\mathfrak{s}}|$ for all $Y \in \mathfrak{n}$ and a fixed K-invariant norm on $\mathfrak{g}$. By K-invariance, the same is true for any u. Moreover, as there are only finitely many parabolic subgroups up to K-conjugacy, C_2 is independent of P. Since X_u is a unit vector for the Riemannian metric on X, we get

$$|\tilde{X}_u f(u)| \le |\nabla f(u)| + C_3 |f(u)|$$

for some $C_3 > 0$. Now it follows from Proposition 2 that the integrand is bounded by

$$C_4 t^{-d/2} \big(\rho(x,\gamma u)^{-1} + 1\big) e^{-\rho(x,\gamma u)^2/c}$$

for some $C_4 > 0$. Applying the preceding corollary, we obtain our assertion. ∎

We again consider sections of E_τ as V_τ-valued functions on G^1. Let $L^2_{\mathrm{cus}}\big(\Gamma\cap P \setminus \tilde{X}(P), E_\tau^{\Gamma\cap P}\big)$ be the subspace of $L^2\big(\Gamma \cap P \setminus \tilde{X}(P), E_\tau^{\Gamma\cap P}\big)$ consisting of all functions whose constant terms along proper parabolic $\mathbb{Q}$-subgroups containing $\mathcal{P}$ vanish. Since $\tilde{X}(P)$ is invariant under N, this makes sense. Let Λ^P be the orthoprojector of $L^2\big(\Gamma\cap P\setminus\tilde{X}(P), E_\tau^{\Gamma\cap P}\big)$ onto that subspace. Then

$$\Lambda^P f(x) = \sum_{\mathcal{P}'\supset\mathcal{P}} (-1)^{\dim \mathfrak{a}_{P'}^G} \int_{\Gamma\cap N'\setminus N'} f(nx)\,dn,$$

where the sum is over all parabolic $\mathbb{Q}$-subgroups $\mathcal{P}'$ containing $\mathcal{P}$ and the Haar measures on their unipotent radicals N' are normalized so that the quotients appearing have total mass one.

Let Δ_P^Γ be the Laplacian with von Neumann boundary conditions on $\Gamma\cap P\setminus\tilde{X}(P)$ with values in the bundle $E_\tau^{\Gamma\cap P}$. Since Λ^P commutes with both Δ and the normal covariant derivation at the boundary, it reduces Δ_P^Γ and defines a restriction $\Delta_{P,\mathrm{cus}}^\Gamma$. Therefore the heat kernel for the von Neumann problem in $L^2_{\mathrm{cus}}\big(\Gamma\cap P\setminus\tilde{X}(P), E_\tau^{\Gamma\cap P}\big)$, *i.e.*, the kernel of $\exp(t\Delta_{P,\mathrm{cus}}^\Gamma)$, equals

$$\bar{E}_P^\Gamma(t,x,y) = \Lambda^P E_P^\Gamma(t,x,y),$$

where it is immaterial to which of the spatial variables Λ^P applies because E_P is N'-invariant and N' is normalized by $\Gamma\cap P$ for each $\mathcal{P}'$.

Proposition 14 *Given a family $\mathcal{T}$ of strictly bounded depth in $\mathcal{G}(\mathbb{Q})$ and a parabolic $\mathbb{Q}$-subgoup $\mathcal{P}$, there exist constants $C>0$, $\varepsilon>0$ such that for $0<t\le 1$, all $\Gamma\in\mathcal{T}$ with $N(\Gamma)$ large enough and all $x\in\tilde{X}(P)$*

$$|\bar{E}_P^\Gamma(t,x,x)| \le Ct^{-d/2}e^{(2-\varepsilon)\rho_P(H_P(x)-T_P)}.$$

Proof In the case $\mathcal{P}=\mathcal{G}$ we have $\bar{E}_G^\Gamma = E_G^\Gamma$, $\rho_G=0$, and by the argument from [5, p. 491], the assertion follows immediately from Proposition 2 and the first estimate in Proposition 11.

Now we consider the case $\mathcal{P}\neq\mathcal{G}$. Let $\mathcal{A}$ be a $\mathbb{Q}$-split component of $\mathcal{P}$ and Ψ the set of its fundamental roots in $\mathfrak{n}$. For each subset F of Ψ, let $\mathfrak{n}_F$ be the sum of the root subspaces in $\mathfrak{n}$ corresponding to those roots which cannot be written as sums of elements of F alone. By assigning to each such F the normalizer of $N_F := \exp\mathfrak{n}_F$ in G, we put the subsets of Ψ in bijection with the $\mathbb{Q}$-parabolics containing $\mathcal{P}$. For any locally integrable $(\Gamma\cap N)$-invariant function f on $\tilde{G}(P)$ with values in a finite-dimensional Hilbert space V, we write

$$f_F(x) = \int_{\Gamma\cap N_F\setminus N_F} f(nx)\,dn.$$

Let us fix $\alpha\in\Psi$ and write $N^\alpha = N_{\Psi\setminus\{\alpha\}}$. Then $N_{F\setminus\{\alpha\}} = N_F N^\alpha$, and $\Lambda^P f(x)$ equals

$$\sum_{F\subset\Psi}(-1)^{\#(F)} f_F(x) = \sum_{\substack{F\subset\Psi\\ F\ni\alpha}}(-1)^{\#(F)}\bigg(f_F(x) - \int_{\Gamma\cap N^\alpha\setminus N^\alpha} f_F(nx)\,dn\bigg).$$

The value of the integral is contained in the convex hull of the set $\{f_F(nx) \mid n\in\Gamma\cap N^\alpha\setminus N^\alpha\}$, and so is $f_F(x)$. Hence the difference is bounded by the diameter of this compact set in

V, which is realized by two of its elements $f_F(x'_{F,\alpha})$ and $f_F(x''_{F,\alpha})$, say, where $x'_{F,\alpha}, x''_{F,\alpha} \in N^\alpha x$. Thus

$$|\Lambda^P f(x)| \le \sum_{\substack{F\subset\Psi\\ F\ni\alpha}} |f_F(x'_{F,\alpha}) - f_F(x''_{F,\alpha})|.$$

Let us apply this to $E_P^\Gamma(t,x,y)$ as a function of its last argument. We obtain a sum of integrals over $\Gamma \cap N_F \setminus N_F$, and for each integrand the preceding lemma provides a bound independent of the variable of integration. Hence there exists $C > 0$ such that

$$|\bar{E}_P^\Gamma(t,x,x)| \le Ct^{-d/2}\bigl(1 + e^{(2\rho_P-\alpha)(H_P(x)-T_P)}\alpha_P(x)\bigr).$$

Since $\alpha \in \Psi$ was arbitrary, we may now take the minimum over all α and use the fact that $\max_{\alpha\in\Psi}\alpha(H) \ge \varepsilon\rho_P(H)$ for all $H \in \mathfrak{a}_P^+$ and some $\varepsilon > 0$. ■

Lemma 15 *If $\lambda_1 \le \lambda_2 \le \cdots$ and $\sum_{n=1}^\infty e^{-t\lambda_n} \le Ct^{-m}$ for $0 < t \le 1$, then for $\lambda \ge 0$ we have $N(\lambda) := \max\{n \mid \lambda_n \le \lambda\} \le Ce(1+\lambda)^m$.*

Proof We have $ne^{-t\lambda_n} \le \sum_{k=1}^n e^{-t\lambda_k} \le Ct^{-m}$. If $\lambda \ge 0$, we may put $t = (1+\lambda)^{-1}$, which is in the required interval. For $\lambda_n \le \lambda$ we then obtain $\frac{n}{e} \le ne^{-\lambda_n/(1+\lambda)} \le C(1+\lambda)^m$, as claimed. ■

Let $\tilde{\Delta}_{\mathrm{cus}}^\Gamma$ denote the direct sum of the operators $\Delta_{P,\mathrm{cus}}^\Gamma$ over all parabolic $\mathbb{Q}$-subgroups $\mathcal{P}$ (including $\mathcal{G}$ itself) up to Γ-conjugacy and $\tilde{N}^\Gamma(\lambda)$ its spectral counting function.

Theorem 16 *If $\mathcal{T}$ is a family of bounded depth in Γ_0, then there exists a constant $C > 0$ such that for all $\Gamma \in \mathcal{T}$ and all $\lambda \ge 0$ we have*

$$\tilde{N}^\Gamma(\lambda) \le C[\Gamma_0 : \Gamma](1+\lambda)^{d/2}.$$

Proof First suppose that $\mathcal{T}$ is a family of strictly bounded depth. Since there are only finitely many Γ_0-conjugacy classes of parabolic $\mathbb{Q}$-subgroups, we may fix one of them, say $\mathcal{C}$, and consider only the subsum $\Delta_{\mathcal{C},\mathrm{cus}}^\Gamma$ of $\Delta_{P,\mathrm{cus}}^\Gamma$ over $\mathcal{P} \in \mathcal{C}$ up to Γ-conjugacy. Now $\tilde{X}_{\mathcal{C}} := \bigcup_{P\in\mathcal{C}} \tilde{X}(P)$ is a disjoint union, and $\Gamma \setminus \tilde{X}_{\mathcal{C}}$ is isometric the disjoint union of the manifolds $\Gamma \cap P \setminus \tilde{X}(P)$ over all $\mathcal{P} \in \mathcal{C}$ up to Γ-conjugacy. The cuspidal heat kernel $\bar{E}_{\mathcal{C}}^\Gamma(t,x,y)$ of the von Neumann problem on this union equals $\bar{E}_P^\Gamma(t,x,y)$ if $x, y \in \Gamma \cap P \setminus \tilde{X}(P)$ and vanishes if x, y are in different components. We have to estimate

$$\int_{\Gamma\setminus\tilde{X}_{\mathcal{C}}} \operatorname{tr}_\tau \bar{E}_{\mathcal{C}}^\Gamma(t,x,x)\,dx.$$

If $N(\Gamma)$ is large enough, Proposition 14 provides a bound for the integrand, which is obviously Γ_0-invariant. After inserting this bound, the integral becomes $[\Gamma_0 : \Gamma]$ times the corresponding integral over $\Gamma_0 \setminus \tilde{X}_{\mathcal{C}}$, *i.e.*, a constant times

$$[\Gamma_0 : \Gamma] \sum_P \int_{\Gamma_0\cap P\setminus\tilde{X}(P)} e^{(2-\varepsilon)\rho_P(H_P(x)-T_P)}\,dx,$$

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press

where the sum is now taken over $\mathcal{P} \in \mathcal{C}$ up to Γ_0-conjugacy. It is well known that the latter integral is finite. This shows that $\operatorname{tr} \exp(t\Delta^{\Gamma}_{\mathcal{C},\mathrm{cus}}) \leq C_{\mathcal{C}}[\Gamma_0 : \Gamma]t^{-d/2}$ for some $C_{\mathcal{C}} > 0$ and all $0 < t \leq 1$, and it remains to apply the preceding lemma.

Now suppose that $\mathcal{T}$ is a family of bounded depth in Γ_0. Then there exists a family $\mathcal{T}'$ of strictly bounded depth for which the theorem is already proved and such that for each $\Gamma \in \mathcal{T}$ there is a $\Gamma' \in \mathcal{T}'$ with $\Gamma' \subset \Gamma$ and $[\Gamma : \Gamma'] \leq D$. *E.g.*, one may take the family of principal congruence subgroups $\Gamma_n(N) \cap \mathcal{G}(\mathbb{Q})$ with N large enough. It follows that

$$\tilde{N}^{\Gamma}(\lambda) \leq \tilde{N}^{\Gamma'}(\lambda) \leq C[\Gamma_0 : \Gamma'](1+\lambda)^{d/2} \leq CD[\Gamma_0 : \Gamma](1+\lambda)^{d/2}. \qquad \blacksquare$$

Corollary 17 *Let Δ^{Γ}_T be the cut-off Laplacian with coefficients in the bundle E^{Γ}_{τ} and $N^{\Gamma}_T(\lambda)$ its spectral counting function. In the situation of the theorem there exists a constant $C > 0$ such that for all $\Gamma \in \mathcal{T}$ and all $\lambda \geq 0$ we have*

$$N^{\Gamma}_T(\lambda) \leq C[\Gamma_0 : \Gamma](1+\lambda)^{d/2}.$$

Indeed, there is $T' > T$ (depending on the choice of ε in the definition of a smooth hull) such that the image of $\tilde{X}(P', T')$ in $N \setminus X$ is contained in $\exp({}^{+}\mathfrak{a}^{G}_{P} + T_P)X^P$ for all pairs $\mathcal{P}' \subset \mathcal{P}$. Thus we get an embedding

$$L^2_T(\Gamma \setminus X, E^{\Gamma}_{\tau}) \to \bigoplus_{\mathcal{P}'} L^2_{\mathrm{cus}}\big(\Gamma \cap P' \setminus \tilde{X}(P', T'), E^{\Gamma \cap P'}_{\tau}\big),$$

and the corollary can be proved just as Theorem 3.23 of [10] using von Neumann bracketing. By restriction we immediately obtain

Corollary 18 *Let $N^{\Gamma}(\lambda)$ be the spectral counting function for the restriction of the Bochner-Laplace operator Δ^{Γ} to $L^2_{\mathrm{cus}}(\Gamma \setminus X, E^{\Gamma}_{\tau})$. If $\mathcal{T}$ is a family of bounded depth in $\mathcal{G}(\mathbb{Q})$, then there exists a constant $C > 0$ such that for all $\Gamma \in \mathcal{T}$ and all $\lambda \geq 0$ we have*

$$N^{\Gamma}(\lambda) \leq C[\Gamma_0 : \Gamma](1+\lambda)^{d/2}.$$

7 An Adelic Version

In this section we shall give an adelic version of our spectral estimate for the cut-off Laplacian. We consider only groups defined over $\mathbb{Q}$, since the general case of a number field can be reduced to this one by restriction of scalars. Denote the ring of adeles of $\mathbb{Q}$ by $\mathbb{A}$ and the subring of finite adeles by $\mathbb{A}_{\mathrm{f}}$. The latter has the maximal compact subring $\hat{\mathbb{Z}} = \prod_p \mathbb{Z}_p$. Recall that the principal congruence subgroup of level N in $\mathrm{GL}_n(\mathbb{A}_{\mathrm{f}})$ is defined as the kernel $K_n(N)$ of the residue map $\mathrm{GL}(n, \hat{\mathbb{Z}}) \to \mathrm{GL}_n(\hat{\mathbb{Z}}/N\hat{\mathbb{Z}})$. Let $\mathcal{G}$ be, as before, a connected reductive linear algebraic $\mathbb{Q}$-group.

Definition A family $\mathcal{K}$ of subgroups of $\mathcal{G}(\mathbb{A}_{\mathrm{f}})$ will be called a *family of bounded depth* in $\mathcal{G}(\mathbb{A}_{\mathrm{f}})$ (with respect to a faithful $\mathbb{Q}$-rational representation $\eta\colon \mathcal{G} \to \mathrm{GL}_n$) if there exists a natural number D with the following property: For each $K \in \mathcal{K}$ there is a natural number N such that $K_n(N) \cap \eta\big(\mathcal{G}(\mathbb{A}_{\mathrm{f}})\big)$ is a subgroup of $\eta(K)$ of index at most D.

Again it is easy to see that this notion is independent of the choice of η.

Now we define the bundles to be considered. Given any linear algebraic Q-group $\mathcal{H}$, we write $\mathcal{H}(\mathrm{A})^1$ for the subgroup of $\mathcal{H}(\mathrm{A})$ consisting of all elements on which every Q-rational character of $\mathcal{H}$ takes a value of idele norm one. Note that we may identify A_H with $\mathcal{H}(\mathrm{A})^1 \setminus \mathcal{H}(\mathrm{A})$. Again we fix a maximal compact subgroup of $G = \mathcal{G}(\mathrm{R})$, which we now denote by K_∞, and a unitary representation τ of K_∞ on a finite-dimensional Hilbert space V_τ. For each open compact subgroup K of $\mathcal{G}(\mathrm{A_f})$, we consider the $\mathcal{G}(\mathrm{A})^1$-homogeneous hermitian vector bundle $\mathrm{E}_\tau^K := \big(\mathcal{G}(\mathrm{A})^1 \times_{K_\infty} V_\tau\big)/K$ over $\mathrm{X}^K := \mathcal{G}(\mathrm{A})^1/K_\infty K$. Since $\mathcal{G}(\mathrm{A})^1$ is the direct product of G^1 and $\mathcal{G}(\mathrm{A_f})^\natural := \big(A_G \times \mathcal{G}(\mathrm{A_f})\big) \cap \mathcal{G}(\mathrm{A})^1 \cong \mathcal{G}(\mathrm{A_f})$, E_τ^K is a smooth bundle isomorphic to $E_\tau \times \mathcal{G}(\mathrm{A_f})^\natural/K$, which carries a canonical G^1-invariant connection ∇. Embedding, as usual, $\mathcal{G}(\mathrm{Q})$ into $\mathcal{G}(\mathrm{A})$ diagonally, we get a hermitian bundle $E_\tau^K := \mathcal{G}(\mathrm{Q}) \setminus \mathrm{E}_\tau^K$ over $X^K := \mathcal{G}(\mathrm{Q}) \setminus \mathrm{X}^K$, whose sections we usually identify with $\mathcal{G}(\mathrm{Q})$-invariant sections of E_τ^K. We define the constant term of a locally integrable section f along the parabolic Q-subgroup $\mathcal{P}$ of $\mathcal{G}$ as

$$f_P(x) = \int_{\mathcal{N}(\mathrm{Q})\setminus\mathcal{N}(\mathrm{A})} \big(r(n)f\big)(x)\,dn,$$

where $\mathcal{N}$ is the unipotent radical of $\mathcal{P}$ and r the right action of $\mathcal{G}(\mathrm{A})^1$ on sections.

Next we turn to truncation. We fix a maximal open subgroup K_f of $\mathcal{G}(\mathrm{A_f})$ each local component of which is special. Writing $K_{\max} = K_\infty K_\mathrm{f}$, we then have $\mathcal{G}(\mathrm{A}) = \mathcal{P}(\mathrm{A})K_{\max}$ for each parabolic Q-subgroup $\mathcal{P}$ of $\mathcal{G}$. A truncation parameter T for $\mathcal{G}$ and $K_{\max}$ is a family of points $T_P \in \mathfrak{a}_P^G$ indexed by the parabolic Q-subgroups such that

- $\gamma \cdot \mathcal{P}(\mathrm{A})^1 \exp(T_P)K_{\max} = \mathcal{P}'(\mathrm{A})^1 \exp(T_{P'})K_{\max}$ for $\gamma \in \mathcal{G}(\mathrm{Q})$ and $\mathcal{P}' = \gamma\mathcal{P}\gamma^{-1}$,
- $T_P = (T_{P'})_P$ for $\mathcal{P}' \subset \mathcal{P}$.

Any such T is determined by the value T_P for a fixed minimal parabolic Q-subgroup. Given an open subgroup K of K_f, let $L_T^2(X^K, E_\tau^K)$ be the subspace of $L^2(X^K, E_\tau^K)$ consisting of all sections f with the following property for each proper (equivalently: each maximal) parabolic Q-subgroup $\mathcal{P}$: If x belongs to $\mathcal{P}(\mathrm{A})^1 \exp({}^+\mathfrak{a}_P^G + T_P)K_{\max}/K_\infty K$, then $f_P(x) = 0$. Now Δ_T^K is defined as the selfadjoint operator in $L_T^2(X^K, E_\tau^K)$ associated to the quadratic form $\|\nabla f\|^2$ on the intersection of the Sobolev space $H^1(X^K, E_\tau^K)$ with $L_T^2(X^K, E_\tau^K)$.

Theorem 19 *If $\mathcal{K}$ is a family of bounded depth in $\mathcal{G}(\mathrm{A}_f)$ and T is a sufficiently large truncation parameter for $\mathcal{G}$ and $K_{\max}$, then there exists a constant $C > 0$ such that for all $K \in \mathcal{K}$ contained in K_f and all $\lambda \geq 0$ the spectral counting function $N_T^K(\lambda)$ of Δ_T^K satisfies*

$$N_T^K(\lambda) \leq C \operatorname{vol}(X^K)(1+\lambda)^{d/2}.$$

This theorem has the obvious corollary concerning the spectral counting function of the Laplacian in the cuspidal subspace of $L^2(X^K, E_\tau^K)$. We shall now deduce it from its non-adelic counterpart. The group G^1 acts on $\mathcal{G}(\mathrm{Q}) \setminus \mathcal{G}(\mathrm{A})^1/K$ from the right, and each orbit has a representative $\xi \in \mathcal{G}(\mathrm{A_f})^\natural$. The stabilizer of ξ is that subgroup $\Gamma_{K,\xi}$ of $\mathcal{G}(\mathrm{Q})$ which is mapped on $\mathcal{G}(\mathrm{Q}) \cap \xi K \xi^{-1}$ under the diagonal embedding into $\mathcal{G}(\mathrm{A_f})$. Thus we have an isomorphism of right G^1-spaces

$$\mathcal{G}(\mathrm{Q}) \setminus \mathcal{G}(\mathrm{A})^1/K \cong \coprod_{\xi \in \mathcal{G}(\mathrm{Q})\setminus\mathcal{G}(\mathrm{A_f})^\natural/K} \Gamma_{K,\xi} \setminus G^1,$$

where $\#\big(\mathcal{G}(\mathrm{Q}) \setminus \mathcal{G}(\mathrm{A_f})^\natural/K\big) < \infty$ by [3]. Taking the fibered product with V_τ over K_∞, we get

$$E_\tau^K \cong \bigsqcup_{\xi \in \mathcal{G}(\mathrm{Q})\setminus\mathcal{G}(\mathrm{A_f})^\natural/K} E_\tau^{\Gamma_{K,\xi}},$$

which yields the obvious isomorphism between the spaces of L^2-sections.

Now we translate the adelic truncation to the non-adelic picture. Let $\mathcal{P}$ be a parabolic Q-subgroup with unipotent radical $\mathcal{N}$. If $f \in L^2(X^K, E_\tau^K)$, the integral defining $f_P(x\xi)$ for $x \in G^1$ and $\xi \in \mathcal{G}(\mathrm{A_f})^\natural$ can be taken over $\mathcal{N}(\mathrm{Q}) \setminus \mathcal{N}(\mathrm{A})/\mathcal{N}(\mathrm{A_f}) \cap \xi K \xi^{-1}$, which is N-isomorphic to $\Gamma_{K,\xi} \cap N \setminus N$ by the additive approximation theorem. We can define a map $H_P \colon \mathcal{G}(\mathrm{A}) \to \mathfrak{a}_P$ by the requirement $\exp H_P(pk) = \mathcal{P}(\mathrm{A})^1 p$ for $p \in \mathcal{P}(\mathrm{A})$, $k \in K_{\max}$, thus extending the previously defined map $H_P \colon X \to \mathfrak{a}_P$ pulled back to G^1. Since $H_P(x\xi) = H_P(x) + H_P(\xi)$ for x, ξ as just before, one easily checks that

$$L_T^2(X^K, E_\tau^K) \cong \bigoplus_{\xi \in \mathcal{G}(\mathrm{Q})\setminus\mathcal{G}(\mathrm{A_f})^\natural/K} L_{T_\xi}^2\big(\Gamma_{K,\xi}, E_\tau^{\Gamma_{K,\xi}}\big)$$

provided we define a truncation parameter T_ξ for $\Gamma_{K_\mathrm{f},\xi}$ and K_∞ by $T_{\xi,P} = T_P - H_P(\xi)$. This isomorphism is, of course, compatible with the cut-off Laplacians, whence for $\lambda > 0$ we have

$$N_T^K(\lambda) = \sum_{\xi \in \mathcal{G}(\mathrm{Q})\setminus\mathcal{G}(\mathrm{A_f})^\natural/K} N_{T_\xi}^{\Gamma_{K,\xi}}(\lambda).$$

Let us identify $\mathcal{G}$ via η with a subgroup of GL_n and put $K_0 = K_n(1) \cap \mathcal{G}(\mathrm{A_f})$. Choose a (finite) set $Z \subset \mathcal{G}(\mathrm{A_f})$ of representatives for $\mathcal{G}(\mathrm{Q}) \setminus \mathcal{G}(\mathrm{A_f})/K_0$. Then there exist natural numbers D_Z, C_Z such that $K_n(D_Z N) \subset \zeta K_n(N)\zeta^{-1}$ and $[\zeta K_n(N)\zeta^{-1} : K_n(D_Z N)] \le C_Z$ for all $\zeta \in Z$ and all natural numbers N. Of course, $K_n(N) \cap \mathrm{GL}(n, \mathrm{Q}) = \Gamma_n(N)$. If $\zeta \in Z$ and $\xi \in \zeta K_0$, then $\xi K_n(N)\xi^{-1} = \zeta K_n(N)\zeta^{-1}$. For $K_n(N) \cap \mathcal{G}(\mathrm{A_f}) \subset K$ this implies

$$\begin{aligned} \big[\Gamma_{K,\xi} : \Gamma_n(D_Z N) \cap \mathcal{G}(\mathrm{Q})\big] &\le \big[\xi K \xi^{-1} : K_n(D_Z N) \cap \mathcal{G}(\mathrm{A_f})\big] \\ &\le C_Z\big[K : K_n(N) \cap \mathcal{G}(\mathrm{A_f})\big]. \end{aligned}$$

Thus, if $\mathcal{K}$ is a family of bounded depth in $\mathcal{G}(\mathrm{A_f})$, then the set of all $\Gamma_{K,\xi}$ with $K \in \mathcal{K}$, $\xi \in ZK_0$ is a family of bounded depth in $\mathcal{G}(\mathrm{Q})$. Since T_ξ depends only on the right K_f-coset of ξ, we can now deduce from Corollary 17 that there exists a constant $C > 0$ such that

$$N_T^K(\lambda) \le C \sum_{\xi \in \mathcal{G}(\mathrm{Q})\setminus\mathcal{G}(\mathrm{A_f})^\natural/K} \mathrm{vol}(\Gamma_{K,\xi} \setminus X)(1+\lambda)^{d/2} = C\,\mathrm{vol}(X^K)(1+\lambda)^{d/2}$$

for all $K \in \mathcal{K}$ with $K \subset K_\mathrm{f}$.

References

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

[2] ———, *A trace formula for reductive groups II: Applications of a truncation operator*. Comp. Math. **40**(1980), 87–121.

[3] A. Borel, *Some finiteness properties of adele groups over number fields.* Inst. Hautes Études Sci. Publ. Math. **16**(1963), 5–30.

[4] P. Chernoff, *Essential self-adjointness of powers of generators of hyperbolic equations*. J. Funct. Anal. **12**(1973), 401–414.

[5] H. Donnelly, *Asymptotic expansions for the compact quotients of properly discontinuous group actions.* Illinois J. Math. **23**(1979), 485–496.

[6] ———, *On the cuspidal spectrum for finite volume symmetric spaces.* J. Differential Geom. **17**(1982), 239–253.

[7] A. Friedman, *Partial differential equations of parabolic type.* Prentice-Hall, Englewood Cliffs, 1964.

[8] L. Ji, *The trace class conjecture for arithmetic groups*. J. Differential Geom. **48**(1998), 165–203.

[9] G. D. Mostow, *Self-adjoint groups*. Ann. of Math. **62**(1955), 44–55.

[10] W. Müller, *The trace class conjecture in the theory of automorphic forms*. Ann. Math. **130**(1989), 473–529.

[11] ———, *The trace class conjecture in the theory of automorphic forms. II*. Geom. Funct. Anal. **8**(1998), 315–355.

[12] M. S. Osborne and G. Warner, *The Selberg trace formula II: Partition, reduction, truncation*. Pacific J. Math. **106**(1983), 307–496.

[13] D. B. Ray and I. M. Singer, *R-torsion and the Laplacian on Riemannian manifolds*. Advances in Math. **7**(1971), 145–210.

Mathematisches Institut
Im Neuenheimer Feld 288
69126 Heidelberg
Germany
email: anton@mathi.uni-heidelberg.de

Humboldt-Universität zu Berlin
Institut für Mathematik
Jägerstr. 10/11
10117 Berlin
Germany
email: hoffmann@mathematik.hu-berlin.de

https://doi.org/10.4153/CJM-1999-014-3 Published online by Cambridge University Press