Restrictions of $SL_3$ Maass Forms to Maximal Flat Subspaces

Simon Marshall
Department of Mathematics, University of Wisconsin–Madison, 480 Lincoln Drive, Madison, WI 53706, USA

Correspondence to be sent to: e-mail: marshall@math.wisc.edu

Let $\psi$ be a Hecke–Maass cusp form on a cubic central simple algebra over $\mathbb{Q}$. We apply arithmetic amplification to improve the local bound for the $L^2$-norm of $\psi$ restricted to maximal flat subspaces.

1 Introduction

Let $S$ be the globally symmetric space $SL_3(\mathbb{R})/SO(3)$. Let $\Gamma \subset SL_3(\mathbb{R})$ be an arithmetic congruence lattice arising from a central simple algebra over $\mathbb{Q}$, and let $X = \Gamma \backslash S$ (see Section 2 for definitions). $X$ is a Riemannian orbifold of finite volume, and a manifold if $\Gamma$ is torsion free. Let $\psi$ be a Hecke–Maass cusp form on $X$, that is to say, a cuspidal eigenfunction of the full ring of invariant differential operators and the Hecke operators. We assume that $\|\psi\|_2 = 1$. We also assume that the spectral parameter of $\psi$ is of the form $t\lambda$, where $t > 1$ and $\lambda \in B^*$, and $B^*$ is a fixed compact regular subset of $a^*$. Let $\Omega \subseteq X$ be a compact set. Let $E \subset \Omega$ be a ball of radius 1 inside a maximal flat subspace of $S$. It was proved in [14, Theorem 1.2] that $\|\psi|_E\|_2 \ll_{B^*} t^{3/4}$, and moreover that this bound is sharp on the compact globally symmetric space $SU(3)/SO(3)$ which is dual to $S$. Note that [6, Theorem 3] and the $L^\infty$ bound of [16] also provide bounds of this form with exponents of 1 and $3/2$, respectively. In this paper, we apply arithmetic amplification to improve this exponent further.
Theorem 1.1. For any $0 < \delta < \frac{1}{12000}$, there is $C = C(B^*, \Omega, \delta) > 0$ such that $\|\psi\|_E \leq Ct^{3/4 - \delta}$. □

We have chosen this particular restriction problem because it is one of only two cases in which we can restrict to a maximal flat subspace and observe a regime change, or kink point, in the local bound for the $L^p$-norm of the restricted eigenfunction. The other case is a geodesic on a surface, treated in [13]; see [14, Theorem 1.2] for the proof of this classification. Both of these features simplify the problem, as the flatness prevents us from having to use the nonabelian Fourier transform on the subspace, and the presence of a regime change lies behind the strong bound we are able to prove for the “off-diagonal” oscillatory integrals in Proposition 4.3. It would be interesting to see whether Theorem 1.1 could be used to prove a power saving for the global $L^p$-norms of $\psi$ for small $p$ as in [1, 5, 17].

1.1 Outline of the proof

We will prove Theorem 1.1 in the same way as the main theorem of [13]. We shall use standard notation defined in Section 2. Choose a compact set $\Omega \subset \text{SL}_3(\mathbb{R})$. If we fix a real-valued function $b \in C_0^\infty(a)$, it suffices to estimate the norm of $b\psi = b(H)\psi(g \exp(H)) \in L^2(a)$, uniformly for $g \in \Omega$. We shall do this by estimating $\langle b\psi, \phi \rangle$ for $\phi \in L^2(a)$ of norm one.

To prove such an estimate, we shall construct an $SO(3)$-bi-invariant function $k_t \in C_0^\infty(\text{SL}_3(\mathbb{R}))$ whose Harish–Chandra transform $h_t$ is nonnegative on the spectral parameters of unitary representations, satisfies $h_t(t\lambda) \geq 1$, and is concentrated near $Wt\lambda$. Form the kernel function

$$K(x, y) = \sum_{\gamma \in \Gamma} k_t(x^{-1}\gamma y)$$

on $\Gamma\backslash \text{SL}_3(\mathbb{R})$. The integral operator on $L^2(X)$ with kernel $K(x, y)$ is then an approximate spectral projector on to eigenfunctions with spectral parameter $t\lambda$. We write the spectral decomposition of $L^2(X)$ as

$$L^2(X) = \int \mathcal{V}_\pi \, d\pi,$$

where each $\mathcal{V}_\pi$ is a 1D space spanned by an eigenform $\psi_\pi$ (which may be an Eisenstein series or an iterated residue thereof). We assume that $\psi \in \{\psi_\pi\}$. If we let the spectral
parameter of $\psi_\pi$ be $\lambda_\pi$, the kernel $K(x, y)$ has the spectral expansion

$$K(x, y) = \int h_t(\lambda_\pi) \psi_\pi(x) \overline{\psi_\pi(y)} \, d\pi.$$  

If we form the integral

$$\int_a b_\phi(H_1) b_\phi(H_2) K(g \exp(H_1), g \exp(H_2)) \, dH_1 \, dH_2,$$

substituting the spectral and geometric expansions of $K(x, y)$ gives

$$\sum_{\gamma \in \Gamma} \int_a b_\phi(H_1) b_\phi(H_2) \overline{k_t}(\exp(-H_1) g^{-1} g \exp(H_2)) \, dH_1 \, dH_2 = \int h_t(\lambda_\pi) |\langle \psi_\pi, b_\phi \rangle|^2 \, d\pi$$

$$\geq |\langle \psi, b_\phi \rangle|^2, \quad (1)$$

which is analogous to a relative trace formula for the subgroup $A$.

We shall need to apply the identity (1) in two different ways. We define the Fourier transform on $L^2(a)$ by

$$\hat{f}(\mu) = \int_a f(H) e^{-i\mu(H)} \, dH.$$  

Let $\beta$ be a parameter satisfying $1 \leq \beta \leq t^{1/2}$. For $\mu \in a^*$, define $H(\mu, \beta) \subset L^2(a)$ to be the space of functions whose Fourier support lies in the ball of radius $\beta$ around $\mu$ with respect to the norm $\| \cdot \|$ defined in Section 2.2. Define

$$H_\beta = \bigoplus_{w \in W} H(w t\lambda, \beta).$$  

If we choose $k_t$ to be sufficiently small that only the identity contributes to (1), and let $\phi \in H_\beta^1$ have norm 1, analyzing the resulting integral allows us to prove the bound

$$\langle b\psi, \phi \rangle \ll_{\epsilon} t^{3/4} \beta^{-1/4+\epsilon}. \quad (2)$$

This shows that the main contribution to $\|b\psi\|_2$ comes from those frequencies near $Wt\lambda$. The exact statement we prove is Proposition 5.1. The bound (2) is purely local, and does not use the arithmeticity of $\Gamma$ or $\psi$. However, it does require asymptotics for the spherical functions on $SL_3$ that are strongly uniform in the group variable, which are taken from [14, Theorems 1.3 and 1.4].
We bound $\langle b\psi, \phi \rangle$ for $\phi \in H(wt\lambda, \beta)$ of norm 1 by applying a technique known as arithmetic amplification, which amounts to introducing a Hecke operator in the identity (1). This method was introduced as a way of bounding $L^\infty$-norms of Maass forms on GL$_2$ by Iwaniec and Sarnak [11]. It has since been used by many authors to bound $L^\infty$-norms of Maass forms on other groups (see for instance [2–4, 10]), which may be thought of as bounding the $L^2$-norms of their restrictions to points. This paper is the second time the method has been applied to restrictions to submanifolds of positive dimension, with the first being [13].

We shall give an outline of the amplification method. We let $T$ be a Hecke operator, and apply $TT^*$ to $K(x, y)$ in the first variable. The identity corresponding to (1) is now

$$
\sum_{\gamma \in TT^* \Gamma} C(\gamma) \int_a b\phi(H_1) b\phi(H_2) \tilde{k}(\exp(-H_1) g^{-1} \gamma g \exp(H_2)) \, dH_1 \, dH_2 = \int h_\ast(\lambda_{\pi}) |\langle T\psi_{\pi}, b\phi \rangle|^2 \, d\pi \geq |\langle T\psi, b\phi \rangle|^2, \tag{3}
$$

where $TT^* \Gamma$ is the set of isometries appearing in $TT^*$, and $C(\gamma)$ is the coefficient of $\gamma \in TT^* \Gamma$. We choose $T$ so that its eigenvalue on $\psi$ is large, and it should have small eigenvalues on the remaining $\psi_{\pi}$ by the orthogonality of systems of Hecke eigenvalues. The term “amplification” comes from this way in which $T$ picks out $\psi$ from the collection $\{\psi_{\pi}\}$. We prove in Section 6 that the oscillatory integrals appearing in (3) are small unless $g^{-1} \gamma g$ is very close to $A$. In Section 4, we use a diophantine argument to show that there are few $\gamma$ such that this happens. This argument is taken from an unpublished paper of Lior Silberman and Akshay Venkatesh, and we thank the authors for permission to reproduce it here. Combining these gives the bound

$$
\langle b\psi, \phi \rangle \ll t^{3/4-\delta} \beta^{9/4}, \tag{4}
$$

where we may take any $0 < \delta < \frac{1}{1200}$. The exact statement we prove is Proposition 4.1. Combining (2) and (4) with $\beta = t^{2\delta/5}$ gives Theorem 1.1.

2 Notation

Throughout the paper, the notation $A \ll B$ will mean that there is a positive constant $C$ such that $|A| \leq CB$, and $A \sim B$ will mean that there are positive constants $C_1$ and $C_2$ such that $C_1 B \leq A \leq C_2 B$. 

2.1 Division algebras and adelic groups

We let $\mathbb{A}$ be the adeles of $\mathbb{Q}$, and $\mathbb{A}_f$ the finite adeles. Let $D$ be a cubic central simple algebra over $\mathbb{Q}$. We denote the reduced norm on $D$ by $nr$, and denote the kernel of $nr$ by $D^1$. We let $D^\times$ and $D^1$ be the algebraic groups over $\mathbb{Q}$ such that $D^\times(\mathbb{Q}) = D^\times$ and $D^1(\mathbb{Q}) = D^1$. We denote the center of $D^\times$ by $Z$, and define $G = D/Z$. We denote $D \otimes \mathbb{Q}_v$ by $D_v$. We denote $D^\times(\mathbb{Q}_v)$ by $D^\times_v$, and likewise for the other groups we have introduced. Let $S_f$ be a finite set of finite places containing all places at which $D$ is ramified. We choose an isomorphism $\phi_v : D_v \simeq M_3(\mathbb{Q}_v)$ for every $v \notin S_f$. We shall implicitly identify $D^\times_v$ with $GL_3(\mathbb{Q}_v)$ via $\phi_v$ for $v \notin S_f$.

Let $R \subset D$ be a maximal order. We define $R_p = R \otimes \mathbb{Z}_p \subset D_p$ for every prime $p$. $R_p$ is a maximal order in $D_p$ for all $p$, and for $p \notin S_f$ we choose $\phi_p$ so that $\phi_p(R_p) = M_3(\mathbb{Z}_p)$. We let $K_f = \otimes_p K_p$ be a maximal compact subgroup of $D^\times(\mathbb{A}_f)$. We assume that $K_p \subseteq R_p^\times$ for all $p$, and that $K_p = R_p^\times$ for $p \notin S_f$. We define $K_\infty = \phi_\infty^1(SO(3))$, and let $K = K_\infty \otimes K_f$.

For each $p$, we define $dg_p$ to be the Haar measure on $D_p^\times$ that assigns mass 1 to $R_p^\times$. We define $dz_p$ to be the Haar measure on $Z_p$ that assigns mass 1 to $Z_p^\times$, and let $d\tilde{g}_p$ be the quotient measure on $G_p$. We choose Haar measures $dg_\infty$, $dz_\infty$, and $d\tilde{g}_\infty$ on $D^\times_\infty$, $Z_\infty$, and $G_\infty$, and denote the product measures by $dg$, $dz$, and $d\tilde{g}$.

Define $X = D^\times(\mathbb{Q}) \setminus D^\times(\mathbb{A})/KZ(\mathbb{A})$. $X$ is compact iff $D \not\simeq M_3(\mathbb{Q})$. If $X$ is an orbifold, we resolve any difficulties in talking about differential operators on $X$ by passing to a cover that is a manifold.

2.2 Lie groups and algebras

We define $A$ to be the subgroup of $SL_3(\mathbb{R})$ consisting of diagonal matrices with positive entries, and let $Z_A$ be the centralizer of $A$ in $M_3(\mathbb{R})$. We define $N$ to be the subgroup of strictly upper triangular matrices. We denote the Lie algebras of $SL_3(\mathbb{R})$, $N$, $A$, and $SO(3)$ by $\mathfrak{g}$, $\mathfrak{n}$, $\mathfrak{a}$, and $\mathfrak{k}$, respectively. We denote the roots of $\mathfrak{a}$ in $\mathfrak{g}$ by $\Delta$, and the set of positive roots corresponding to $\mathfrak{n}$ by $\Delta^+$. We denote the set of regular and singular points in $\mathfrak{a}$ and $\mathfrak{a}^*$ by $\mathfrak{a}_r$, $\mathfrak{a}_s$, etc. We let $M$ and $M'$ be the centralizer and normalizer of $\mathfrak{a}$ in $SO(3)$, and define the Weyl group $W = M'/M$.

We equip $M_3(\mathbb{R})$ with the standard Euclidean norm as a 9D vector space, which we denote by $\| \cdot \|$. We obtain a positive definite norm on $\mathfrak{g}$ from $\| \cdot \|$ under the natural restriction, as well as a norm on $\mathfrak{a}^*$ by duality. We shall also denote these norms by $\| \cdot \|$, the particular one we are using will be clear from the context. We let $B_n$, $B_\alpha$, and $B_\varepsilon$ be the unit balls in $\mathfrak{n}$, $\mathfrak{a}$, and $\mathfrak{k}$ with respect to $\| \cdot \|$. We let $d(\cdot, \cdot)$ be the left-invariant metric
on $\text{SL}_3(\mathbb{R})$ associated to $\| \cdot \|$. The Killing form on $\mathfrak{g}$ will always be denoted by $(,)$, and we give $S$ the Riemannian structure determined by $(,)$. The identification $\phi_{\infty}$ gives an identification of $G_{\infty}$ with $\text{SL}_3(\mathbb{R})$, and we implicitly transfer all the definitions we have made here to $G_{\infty}$.

2.3 Hecke algebras

If $p \notin S_f$, we let $\mathcal{H}_p$ denote the space of functions in $C_{0}^{\infty}(\text{GL}_3(\mathbb{Q}_p))$ that are bi-invariant under $K_p$. It is an algebra under convolution with respect to the measure $dg_p$. If $\varphi \in \mathcal{H}_p$, we denote its adjoint operator by $\varphi^*$, which satisfies $\varphi^*(g) = \overline{\varphi(g^{-1})}$. If $(a, b, c) \in \mathbb{Z}^3$, we define $K_p(a, b, c) \subset \text{GL}_3(\mathbb{Q}_p)$ to be the double coset

$$K_p(a, b, c) = K_p \left( \begin{array}{ccc} p^a & & \\ & p^b & \\ & & p^c \end{array} \right) K_p,$$

and let $\Phi_p(a, b, c) \in \mathcal{H}_p$ be the characteristic function of $K_p(a, b, c)$.

We let $\mathcal{H}$ be the space of functions in $C_{0}^{\infty}(D^\times(\mathbb{A}_f))$ that are bi-invariant under $K_f$. We shall frequently identify $\mathcal{H}_p$ with a subalgebra of $\mathcal{H}$ in the natural way. If $(a, b, c) \in (\mathbb{Q}^\times)^3$, define $K(a, b, c) \subset D^\times(\mathbb{A}_f)$ to be

$$K(a, b, c) = \bigotimes_{p \in S_f} K_p \otimes \bigotimes_{p \notin S_f} K_p(\text{ord}_p(a), \text{ord}_p(b), \text{ord}_p(c)),$$

and let $\Phi(a, b, c) \in \mathcal{H}$ be the characteristic function of $K(a, b, c)$. We shall sometimes implicitly identify $\Phi(a, b, c)$ and $K(a, b, c)$ with their images in $G$ under central integration and projection. The action of $\varphi \in \mathcal{H}$ on a function $f$ on $D^\times(\mathbb{Q}) \backslash D^\times(\mathbb{A})$ is given by the usual formula

$$\varphi \cdot f(x) = f \ast \varphi^\vee(x) = \int_{D^\times(\mathbb{A}_f)} f(xg) \varphi(g) \, dg,$$

where $\varphi^\vee(g) = \overline{\varphi(g^{-1})}$.

2.4 Spherical functions

If $\mu \in \mathfrak{a}^*$, we define $\varphi_{\mu}$ to be the corresponding spherical function on $G_{\infty}$. If $k \in C_{0}^{\infty}(G_{\infty})$, we define its Harish–Chandra transform by

$$\hat{k}(\mu) = \int_{G_{\infty}} k(g) \varphi_{-\mu}(g) \, d\tilde{g}_{\infty}.$$
2.5 Maass forms

Let $\psi$ be a cuspidal Hecke–Maass form on $X$, that is to say, an eigenfunction of the ring of invariant differential operators and of the Hecke algebras $\mathcal{H}_p$ for $p \not\in S_f$. We assume that $\|\psi\|_2 = 1$. We let the spectral parameter of $\psi$ be $t \lambda$, where $t > 0$ and $\lambda \in B^*$, and $B^*$ is a fixed compact subset of $a^*_r$.

3 Constructing an Amplifier

Let $p \not\in S_f$ be a prime. In this section, we construct an element $T_p \in \mathcal{H}_p$ that will form part of the amplifier $T$. We begin with the following relation in $\mathcal{H}_p$.

Lemma 3.1. We have

$$\Phi_p(1, 0, 0)^* \Phi_p(1, 1, 0) = \Phi_p(2, 1, 0) + (p^2 + p + 1) \Phi_p(1, 1, 1).$$

Proof. As $\Phi_p(1, 0, 0)^* \Phi_p(1, 1, 0)$ must be supported on those double cosets $K_p(a, b, c)$ with $a + b + c = 3$ and $a, b, c \geq 0$, we must have

$$\Phi_p(1, 0, 0)^* \Phi_p(1, 1, 0) = a \Phi(1, 1, 1) + b \Phi_p(2, 1, 0) + c \Phi_p(3, 0, 0)$$

(5)

for some $a, b,$ and $c \in \mathbb{R}$. Taking adjoints gives

$$\Phi_p(-1, 0, 0)^* \Phi_p(-1, -1, 0) = a \Phi_p(-1, -1, -1) + b \Phi_p(-2, -1, 0) + c \Phi_p(-3, 0, 0),$$

and it follows that $c = 0$ as there can be no matrices whose entries have denominator $p^3$ in the support of $\Phi_p(-1, 0, 0)^* \Phi_p(-1, -1, 0)$.

If we translate (5) by $p^{-1} I$ we obtain

$$\Phi_p(1, 0, 0)^* \Phi_p(-1, 0, 0) = a \Phi(0, 0, 0) + b \Phi_p(1, 0, -1),$$

(6)

and it may be easily shown that

$$\text{d}g_p(K_p(1, 0, 0)) = p^2 + p + 1, \quad \text{d}g_p(K_p(1, 0, -1)) = (p^2 + p)(p^2 + p + 1).$$

Evaluating (6) at the identity gives

$$a = \Phi_p(1, 0, 0)^* \Phi_p(-1, 0, 0)(e) = \|\Phi_p(1, 0, 0)\|_2^2 = p^2 + p + 1.$$
Integrating (6) over $\text{GL}_3(\mathbb{Q}_p)$ gives

$$dg_p(K_p(1, 0, 0))^2 = a + b \, dg_p(K_p(1, 0, -1)),$$

and it follows that $b = 1$. ■

Lemma 3.1 implies that if we define $a(\psi, p)$ and $b(\psi, p)$ by

$$\Phi_p(1, 0, 0)\psi = a(\psi, p)p\psi, \quad \Phi_p(2, 1, 0)\psi = b(\psi, p)p^2\psi,$$

then we cannot have both $|a(\psi, p)| \leq \frac{1}{2}$ and $|b(\psi, p)| \leq \frac{1}{2}$. We define

$$T_p = \begin{cases} 
\Phi_p(1, 0, 0)/a(\psi, p)p & \text{if } |a(\psi, p)| > \frac{1}{2}, \\
\Phi_p(2, 1, 0)/b(\psi, p)p^2 & \text{otherwise.} 
\end{cases}$$

(7)

It follows that $T_p\psi = \psi$ for all $p \notin S_f$. We shall need the following bound for the coefficients in the expansion of $T_pT_p^*$.

**Lemma 3.2.** Write

$$T_pT_p^* = \sum_{a \geq b \geq c} \alpha(a, b, c)\Phi_p(a, b, c).$$

If $\alpha(a, b, c) \neq 0$, then we have

$$\alpha(a, b, c) \ll p^{c-a}$$

(8)

where the implied constant is absolute. Moreover, one of the three pairs of inequalities

$$-2 \leq a, b, c \leq 2 \quad \text{and} \quad a + b + c = 2,$$

$$-2 \leq a, b, c \leq 2 \quad \text{and} \quad a + b + c = 0,$$

$$-2 \leq a, b, c \leq 1 \quad \text{and} \quad a + b + c = -2,$$

(9)

holds. ■

**Proof.** If we define

$$\Phi_p = p^{-1}\Phi_p(1, 0, 0) + p^{-2}\Phi_p(2, 1, 0)$$

and write

$$\Phi_p \ast \Phi_p^* = \sum_{a \geq b \geq c} \beta(a, b, c)\Phi_p(a, b, c),$$

(10)
then we have $\beta(a, b, c) \geq 0$ and $|\alpha(a, b, c)| \leq 2\beta(a, b, c)$. The bounds (9) on $a$, $b$, and $c$ may be proven on a case by case basis by taking determinants, and taking adjoints and considering denominators as in Lemma 3.1.

We shall bound $\beta(a, b, c)$ using the spherical transform. Define $A_p : \text{GL}_3(\mathbb{Q}_p) \to \mathbb{Z}^3$ to be the $p$-adic Iwasawa $A$ coordinate with respect to the standard collection of subgroups, and let

$$
\rho : \mathbb{Z}^3 \to \mathbb{Z} \\
(a, b, c) \mapsto a - c
$$

be the half sum of the positive roots. We define the function $\varphi_0$ by

$$
\varphi_0(g) = \int_{K_p} p^{\rho(A_p(g))} \, dK_p,
$$

so that $\varphi_0$ is the spherical function with trivial Satake parameter. It follows from [7, Proposition 7.2] that

$$
\Phi_p(1, 0, 0) \ast \varphi_0 = \Phi_p(1, 1, 0) \ast \varphi_0 = 3p\varphi_0.
$$

Combining this with Lemma 3.1 gives $\Phi_p(2, 1, 0) \ast \varphi_0 = (8p^2 - p - 1)\varphi_0$, and so

$$
(\Phi_p \ast \Phi_p^*) \ast \varphi_0 = (11 - p^{-1} - p^{-2})^2 \varphi_0.
$$

Taking only one term in the expansion (10) and evaluating at the identity gives

$$
\beta(a, b, c)(\Phi_p(a, b, c) \ast \varphi_0)(e) \leq (11 - p^{-1} - p^{-2})^2.
$$

We have

$$
(\Phi_p(a, b, c) \ast \varphi_0)(e) = \int_{K_p(a, b, c)} \varphi_0(g^{-1}) \, dg_p \\
= \int_{K_p(a, b, c)} p^{\rho(A(g^{-1}))} \, dg_p \\
\geq p^{a-c},
$$

and (8) now follows. \qed
4 Amplification of Periods Along Flats

Fix a real-valued function \( b \in C_0^\infty(a) \) with \( \text{supp}(b) \subseteq B_a \). Let \( \Omega \subset G_\infty \) be a compact set, let \( g_0 \in \Omega \), and let \( b \psi \) denote the function \( b(H)\psi(g_0 \exp(H)) \in L^2(a) \). This means that we are restricting ourselves to flat subspaces contained in the identity component of \( X \). This entails no loss of generality, as we may treat the other connected components by first translating \( \psi \) by an element of \( G(\mathbb{A}_f) \) (in fact, the Hasse–Schilling norm theorem (see [15, Theorem 33.15]) and the fact that \( D_1 \) satisfies strong approximation imply that an element of \( (R \otimes \mathbb{Z} \mathbb{A}_f)^* \) will suffice.) If we let \( w \in W \) and \( \phi \in H(wt\lambda, \beta) \) satisfy \( \| \phi \|_2 = 1 \), where \( H(wt\lambda, \beta) \) is as in Section 1.1, we will prove the following bound for \( \langle b \psi, \phi \rangle \).

**Proposition 4.1.** For any \( \epsilon > 0 \) there is \( C = C(B^*, \Omega, \epsilon) > 0 \) such that

\[
\langle b \psi, \phi \rangle \leq C t^{3/4 - 1/1200 + \epsilon} \beta^{9/4}. \]

To prove Proposition 4.1, we may assume without loss of generality that \( w = 1 \) as the other cases are identical. Fix a real-valued nonnegative function \( h \in C_\infty(a^*) \) of Paley–Wiener type that satisfies \( h(0) = 1 \). Define \( \tilde{h}_t^0 \) by

\[
\tilde{h}_t^0(\mu) = \sum_{w \in W} h(w \mu - t\lambda),
\]

and let \( k_t^0 \) be the \( K \)-bi-invariant function on \( G_\infty \) with Harish–Chandra transform \( \tilde{h}_t^0 \). The Paley–Wiener theorem of Gangolli [9] implies that \( k_t^0 \) is of compact support that is uniform in \( t\lambda \), and may be chosen to be arbitrarily small. We define \( k_t = k_t^0 \ast (k_t^0)^* \) and \( h_t = \hat{k}_t \). Assume that \( \text{supp}(k_t(\exp(H))) \subseteq B_a \).

Let \( 1 \leq N \leq t \) be an integer to be chosen later, and define \( T \) to be the Hecke operator

\[
T = \sum_{1 \leq p \leq N} T_p
\]

where \( T_p \) is as in (7). If we define \( T(\psi) \) to be the scalar by which \( T \) acts on \( \psi \), the equations \( T_p \psi = \psi \) for \( p \notin S_f \) imply that \( T(\psi) \gg \rho \), \( N^{1-\epsilon} \). We shall estimate \( \langle b \psi, \phi \rangle \) by estimating \( \langle T \psi, b\phi \rangle \). If \( g \in G_\infty \), we define the integral

\[
I(t, \phi, g) = \int_a b\phi(H_1)\overline{b\phi(H_2)} k_t(\exp(-H_1)g \exp(H_2)), \, dH_1 \, dH_2.
\]

We suppose that \( b \) and \( h_t \) are chosen so that \( I(t, \phi, g) = 0 \) unless \( d(g, e) \leq 1 \). The main amplification inequality that we shall use is the following lemma.
Lemma 4.2. We have

$$|⟨Tψ, bφ⟩|^2 \leq \sum_{γ \in G(Q)} |(TT^∗)(γ)I(t, φ, g_0^{-1}γg_0)|.$$ (11)

□

Proof. Consider the function

$$K(x, y) = \sum_{γ \in G(Q)} \tilde{k}_γTT^∗(x^{-1}γy)$$

on $G(\mathbb{A}) \times G(\mathbb{A})$. We have

$$\int_{\mathcal{A}} b\phi(H_1)b\phi(H_2)K(g_0 \exp(H_1), g_0 \exp(H_2)) \, dH_1 \, dH_2 = \sum_{γ \in G(Q)} (TT^∗)(γ)\overline{I(t, φ, g_0^{-1}γg_0)}.$$ (11)

We write the spectral decomposition of $L^2(X)$ as

$$L^2(X) = \int V_π \, dπ,$$

where each $V_π$ is a 1D space spanned by an eigenform $ψ_π$ (which may be an Eisenstein series or an iterated residue thereof). We assume that $ψ ∈ \{ψ_π\}$. If we let the spectral parameter of $ψ_π$ be $λ_π$, the identity $k_t = k'_t$ implies that

$$\pi(\tilde{k}_t)ψ_π = \psi_π \int_{G_∞} \tilde{k}_t(g)φ_{λ_π}(g) \, d\tilde{g}_∞$$

$$= \psi_π \int_{G_∞} k_t(g)φ_{λ_π}(g^{-1}) \, d\tilde{g}_∞$$

$$= \psi_π \int_{G_∞} k_t(g)φ_{-λ_π}(g) \, d\tilde{g}_∞$$

$$= h_t(λ_π)ψ_π.$$ (12)

The spectral expansion of $K(x, y)$ is therefore

$$K(x, y) = \int h_t(λ_π)(Tψ_π)(x)(\overline{Tψ_π})(y) \, dπ,$$

which implies that

$$\int_{\mathcal{A}} b\phi(H_1)b\phi(H_2)K(g_0 \exp(H_1), g_0 \exp(H_2)) \, dH_1 \, dH_2 = \int h_t(λ_π)|⟨Tψ_π, bφ⟩|^2 \, dπ.$$ (12)
Unitarity implies that all spectral parameters in the integral satisfy $h_t(\lambda \pi) = |h^{0}_{t}(\lambda \pi)|^2 \geq 0$. Dropping all terms in (12) but $\psi$ and using the inequality $h_t(t\lambda) \geq 1$ gives the lemma.

We shall estimate (11) with the aid of two propositions. The first says that the only significant contribution to (11) comes from those $\gamma \in G(\mathbb{Q}) \cap \text{supp}(TT^*)$ for which $g_0^{-1} \gamma g_0$ is near $MA$.

**Proposition 4.3.** We have

$$|I(t, \phi, g)| \ll t^{3/2+\epsilon} \beta^{9/2}$$

for all $g$ and any $\epsilon > 0$. If $d(g, MA) \geq t^{-1/2+\epsilon} \beta^{1/2}$ for some $\epsilon > 0$, we have

$$|I(t, \phi, g)| \ll t^{-A}.$$  

The implied constants depend on $B^*$, $\epsilon$, and $A$ where applicable.

The second proposition bounds the number of isometries that map $E$ close to itself. If $x \in R$ is nonzero, denote its image in $G(\mathbb{Q})$ by $\bar{x}$. If $g \in G_\infty$, $a, b, c \in \mathbb{Q}^\times$, $n \in \mathbb{Z}_{>0}$, and $\kappa > 0$, we define

$$\mathcal{M}(g, a, b, c, \kappa) = \{ \gamma \in G(\mathbb{Q}) \cap K(a, b, c) | d(g^{-1} \gamma g, e) \leq 1, d(g^{-1} \gamma g, MA) \leq \kappa \},$$

$$M(g, a, b, c, \kappa) = |\mathcal{M}(g, a, b, c, \kappa)|,$$

and

$$\mathcal{L}(g, n, \kappa) = \{ x \in R | \|nr(x)\| = n, d(g^{-1} \bar{x}g, e) \leq 1, d(g^{-1} \bar{x}g, MA) \leq \kappa \},$$

$$L(g, n, \kappa) = |\mathcal{L}(g, n, \kappa)|.$$

The following three results allow us to bound $M(g, a, b, c, \kappa)$.

**Lemma 4.4.** If $a, b, c \in \mathbb{Z}_{>0}$ are relatively prime to $S_f$ and satisfy $(a, b, c) = 1$, we have

$$M(g, a, b, c, \kappa) \leq L(g, abc, \kappa).$$
Proof. Suppose $\gamma \in \mathcal{M}(g, a, b, c, \kappa)$. If $x \in R$ is a primitive element that projects to $\gamma$, then $x_p$ must be primitive in $R_p$ for all $p$. We have

$$x_p \in K_p Z_p \subseteq R_p^* Z_p, \quad p \in S_f,$$

$$x_p \in K_p (\text{ord}_p(a), \text{ord}_p(b), \text{ord}_p(c)) Z_p, \quad p \notin S_f.$$

The primitivity of $x_p$ and the condition $(a, b, c) = 1$ imply that

$$x_p \in R_p, \quad p \in S_f,$$

$$x_p \in K_p (\text{ord}_p(a), \text{ord}_p(b), \text{ord}_p(c)), \quad p \notin S_f,$$

so that $|\text{nr}(x)| = abc$. It follows that $x \in \mathcal{L}(g, abc, \kappa)$. Because distinct elements of $\mathcal{M}(g, a, b, c, \kappa)$ have distinct elements of $\mathcal{L}(g, abc, \kappa)$ assigned to them, the result follows.

Proposition 4.5. There is $C = C(\Omega) > 0$ such that if $g \in \Omega$ and $\kappa \leq C n^{-24}$, then $L(g, n, \kappa) \ll_{\Omega, \epsilon} n^\epsilon$.

Proof. Suppose that $x \in \mathcal{L}(g, n, \kappa)$. The assumptions that $d(g^{-1} \bar{x}g, e) \leq 1$ and $g \in \Omega$ imply that $\bar{x}$ lies in a compact set $\Omega' \subset G_\infty$ depending on $\Omega$. Combined with the assumption that $|\text{nr}(x)| = n$, this gives $\|x\| \ll \Omega n^{1/3}$.

The assumption $n(g^{-1} \bar{x}g, MA) \leq \kappa$ gives

$$\inf \{\|x - z\| \mid z \in gZ_A g^{-1}\} \leq C' \kappa n^{1/3} \quad (15)$$

for some $C'$ depending only on $\Omega$. Lemma 4.6 implies that there is $C > 0$ depending on $\Omega$ such that if $\kappa \leq C n^{-24}$, then $\mathcal{L}(g, n, \kappa)$ must be contained in a cubic subfield $F \subset D$.

As $x \in R \cap F$, $x$ must lie in the ring of integers $O_F$ of $F$, and as the reduced norm $\text{nr}$ on $D$ agrees with the norm $N$ on $F$ we must have $|N(x)| = n$. The condition $\|x\| \ll \Omega n^{1/3}$ implies that the image of $x$ under all archimedean embeddings of $F$ must be $\ll_{\Omega} n^{1/3}$, and the number of algebraic integers in $F$ satisfying these conditions may easily be seen to be $\ll_{\Omega, \epsilon} n^\epsilon$, uniformly in $F$.

The following lemma is taken from an unpublished paper of Lior Silberman and Akshay Venkatesh. We thank the authors for permission to reproduce it here.
Lemma 4.6. Let \( S \subset D_\infty \) be an \( \mathbb{R} \)-subalgebra isomorphic to \( \mathbb{R}^3 \). For \( C > 0 \) sufficiently small depending on \( R \), and any \( X > 1 \) and \( 0 < \epsilon < 1 \), the set of \( x \in R \) satisfying

\[
\|x\| \leq X, \quad \inf\{\|x - s\| \mid s \in S\} \leq \epsilon
\]

is contained in a subalgebra \( F \subset D \) of dimension at most 3 as long as

\[
\epsilon X^{71} < C. \tag{17}
\]

Proof. Fix an integral basis \( \{v_i\}_{i=1}^9 \) of \( R \). If \( (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in D^4 \), define \( P(\alpha_1, \ldots, \alpha_4) \) by expressing each \( \alpha_i \) in the basis \( \{v_i\} \), combining the resulting vectors into a \( 4 \times 9 \) matrix, and taking the sums of the squares of the \( 4 \times 4 \) minors. This makes \( P : D^4 \to \mathbb{Q} \) a polynomial map of degree 8, with integral coefficients with respect to \( R \), such that \( P(\alpha_1, \ldots, \alpha_4) = 0 \) exactly when \( \alpha_1, \ldots, \alpha_4 \) span a linear subspace of dimension \( \leq 3 \).

Let \( F \) be the \( \mathbb{Q} \)-algebra generated by those \( x \in R \) satisfying (16). It is clear that \( F \) is in fact spanned by monomials in such \( x \) of length at most \( 9 = \dim_\mathbb{Q} D \). Each such monomial \( y \) lies in \( R \), and satisfies the bounds

\[
\|y\| \ll X^9, \quad \inf\{\|y - s\| \mid s \in S\} \ll X^8 \epsilon. \tag{18}
\]

Take \( y_1, \ldots, y_4 \in R \) satisfying (18). There are \( z_1, \ldots, z_4 \in S \) such that \( \|y_i - z_i\| \ll X^8 \epsilon \), and so \( P(y_1, \ldots, y_4) \ll (X^9)^7 X^8 \epsilon = X^{71} \epsilon \). On the other hand, if \( P(y_1, \ldots, y_4) \neq 0 \), then we must have \( |P(y_1, \ldots, y_4)| \geq 1 \) by the condition that \( P \) had integral coefficients with respect to \( R \). It follows that if a condition of the form (17) holds for \( C \) as stated, then \( y_1, \ldots, y_4 \) span a \( \mathbb{Q} \)-linear space of dimension at most 3. It follows that \( F \) has dimension at most 3, which completes the proof. \( \square \)

With these results, we may now estimate (11). In the calculation occupying the rest of this section, all implied constants will depend on \( B^*, \Omega \), and \( \epsilon \). We have

\[
TT^* = \sum_{p \neq q \leq N} T_p T_q^* + \sum_{p \leq N} T_p T_p^*.
\]

We may bound

\[
\sum_{\gamma \in G(\mathbb{Q})} \sum_{p \neq q \leq N} |T_p T_q^*(\gamma) I(t, \phi, g_0^{-1} \gamma g_0)|
\]

...
by the sum of four expressions, the first of which is

$$\sum_{p \neq q \leq N} \sum_{\gamma \in G(\mathbb{Q})} \frac{1}{pq} \Phi(pq, p, 1)(\gamma)|I(t, \phi, g_0^{-1}\gamma g_0)|.$$

The bound (14) and our assumption that $N \leq t$ allow us to restrict the sum over $\gamma$ to $\mathcal{M}(g_0, pq, p, 1, t^{-1/2+\epsilon} \rho^{1/2})$. If we assume that $N \leq t^{1/600}$, we may combine Lemma 4.4, Proposition 4.5, and the bound (13) to obtain

$$\sum_{\gamma \in G(\mathbb{Q})} \Phi(pq, p, 1)(\gamma)|I(t, \phi, g_0^{-1}\gamma g_0)| \ll t^{3/2+\epsilon} \rho^{9/2}.$$ 

We therefore have

$$\sum_{p \neq q \leq N} \sum_{\gamma \in G(\mathbb{Q})} \frac{1}{pq} \Phi(pq, p, 1)(\gamma)|I(t, \phi, g_0^{-1}\gamma g_0)| \ll t^{3/2+\epsilon} \rho^{9/2}. \quad (19)$$

The other three expressions we must consider are similar, but with $\Phi(pq, p, 1)/pq$ replaced with $\Phi(p^2q, p, 1)/p^2q$, $\Phi(pq^2, pq, 1)/pq^2$, and $\Phi(p^2q^2, pq, 1)/p^2q^2$, respectively. We may bound them in the same way using our assumption that $N \leq t^{1/600}$.

The analysis of the $T_p T_p^*$ terms is similar. Lemma 3.2 implies that we must consider terms of the form $\Phi(p^a, p^b, 1)/p^a$ with $a \geq b \geq 0$, and the inequalities (9) imply that $a + b + c \leq 6$. Arguing as above using the bound $N \leq t^{1/600}$ gives

$$\sum_{\gamma \in G(\mathbb{Q})} |T_p T_p^*(\gamma)|I(t, \phi, g_0^{-1}\gamma g_0)| \ll t^{3/2+\epsilon} \rho^{9/2}.$$ 

If we sum over $p$ and combine this with (19), we obtain

$$\sum_{\gamma \in G(\mathbb{Q})} |(TT^*)(\gamma)|I(t, \phi, g_0^{-1}\gamma g_0)| \ll N t^{3/2+\epsilon} \rho^{9/2},$$

and Lemma 4.2 gives

$$\langle T \psi, b \phi \rangle \ll N^{1/2} t^{3/4+\epsilon} \rho^{9/4}.\quad$$

The bound $T(\psi) \gg_{\epsilon} N^{1-\epsilon}$ gives $\langle \psi, b \phi \rangle \ll N^{-1/2} t^{3/4+\epsilon} \rho^{9/4}$. Choosing $N = t^{1/600}$ completes the proof of Proposition 4.1.
5 Bounds Away from the Spectrum

Let $H_\beta$ be as in Section 1.1, and let $\phi \in H_\beta^1$ satisfy $\|\phi\|_2 = 1$. In this section, we prove the following bound for $\langle b\psi, \phi \rangle$.

**Proposition 5.1.** For any $\epsilon > 0$ there is $C = C(B^*, \Omega, \epsilon) > 0$ such that $\langle b\psi, \phi \rangle \leq C t^{3/4} \beta^{-1/4 + \epsilon}$. \hfill \Box

**Proof.** If we apply Lemma 4.2 with $T$ chosen to be the characteristic function of a sufficiently small open subgroup of $G(A_f)$, depending on $\Omega$, then only $\gamma = 1$ will make a nonzero contribution to the sum. This gives

$$|\langle \psi, b\phi \rangle|^2 \ll \Omega \left| \int a b\phi(H_1) b\phi(H_2) k_t(\exp(H_2) - H_1) \, dH_1 \, dH_2 \right|.$$ 

If we define $p_t(H) = k_t(\exp(H))$, we may rewrite this as $|\langle \psi, b\phi \rangle|^2 \ll |\langle p_t * b\phi, b\phi \rangle|$.

For $v \in a^*$ and $C \geq 0$, we define $B(Wv, C)$ to be the union of the balls of radius $C$ around the points in $Wv$ with respect to $\|\cdot\|$. Write $b\phi = \phi_1 + \phi_2$, where $\hat{\phi}_1$ is supported on $B(Wt\lambda, \beta/2)$ and $\hat{\phi}_2$ is supported on $a^* \setminus B(Wt\lambda, \beta/2)$. Because $b$ was a fixed smooth function, we have $\|\phi_1\|_2 \ll A \beta^{-A}$. We have

$$\langle p_t * b\phi, b\phi \rangle = \langle p_t * \phi_1, \phi_1 \rangle + \langle p_t * \phi_2, \phi_2 \rangle \leq OA(\beta^{-A}) \|\hat{p}_t\|_\infty + \sup_{\mu \notin B(Wt\lambda, \beta/2)} |\hat{p}_t(\mu)|.$$

The result now follows from the following lemma. \hfill \Box

**Lemma 5.2.** We have $\|\hat{p}_t\|_\infty \ll_{B^*} t^{3/2}$, and

$$\hat{p}_t(\mu) \ll_{B^*, \epsilon} t^{3/2} \beta^{-1/2 + \epsilon} \quad (20)$$

for $\mu \notin B(Wt\lambda, \beta/2)$. \hfill \Box

**Proof.** The first assertion is proved in [14, Section 3], or can be deduced easily from the bound

$$k_t(\exp(H)) \ll_{B^*} t^3 \prod_{\alpha \in A^+} (1 + t|\alpha(H)|)^{-1/2}$$

proved in [14, Lemma 2.6]. We will prove (20) with the aid of two asymptotic formulae for $\varphi_{t\nu}$ taken from [14]. Let $B_0^*$ be a compact subset of $a^*_0$ that contains an open neighborhood...
of $B^*$. In the rest of the proof, all implied constants will depend on $B_0^*$ unless otherwise stated. For $H \in a$, we define

$$
\|H\|_s = \inf \{\|H - Z\| | Z \in a_s\}.
$$

We may apply [14, Theorems 1.3 and 1.4] to deduce that if $H \in 2B_a$ and $\nu \in B_0^*$, we have

$$
\varphi_{t\nu}(\exp(H)) \ll \prod_{\alpha \in \Delta^+} (1 + t|\alpha(H)|)^{-1/2},
$$

and there are functions $f_w \in C^\infty((2B_a \cap a_r) \times B_0^* \times \mathbb{R}_{>0})$ for $w \in W$ such that

$$
\left(\frac{\partial}{\partial H}\right)^n f_w(H, \nu, t) \ll_n \|H\|^{-n}_s \prod_{\alpha \in \Delta^+} (t|\alpha(H)|)^{-1/2}
$$

and

$$
\varphi_{t\nu}(\exp(H)) = \sum_{w \in W} f_w(H, \nu, t) \exp(it\nu(wH)) + O_A((t\|H\|^{-A}) \prod_{\alpha \in \Delta^+} (t|\alpha(H)|)^{-1/2}.
$$

Let $\chi$ be the characteristic function of the set of points in $a$ that are at distance at most 2 from $a_s$. Fix a function $b_0 \in C_0^\infty(a)$ with integral 1 and supp$(b_0) \subseteq B_a$. If we define $b_1 = \chi * b_0$, we have $b_1(H) = 1$ if $\|H\|_s \leq 1$. Fix a function $b_2 \in C_0^\infty(a)$ with supp$(b_2) \subseteq 2B_a$ that is equal to 1 on $B_a$. Our assumption that supp$(k_t(\exp(H))) \subseteq B_a$ implies that

$$
\hat{\psi}_t(\mu) = \int_a b_2(H)k_t(\exp(H)) e^{-i\mu(H)} dH.
$$

We have the trivial bound

$$
\int_a b_2(H)\varphi_{t\nu}(\exp(H)) e^{-i\mu(H)} dH \ll 1,
$$

where the implied constant is independent of $B_0^*$, and (21) gives

$$
\int_a b_2(H)\varphi_{t\nu}(\exp(H)) e^{-i\mu(H)} dH \ll t^{-3/2}
$$

for $\nu \in B_0^*$. If we could prove that

$$
\int_a b_2(H)\varphi_{t\nu}(\exp(H)) e^{-it\mu(H)} dH \ll \epsilon t^{-3/2} \beta^{-1/2 + \epsilon}
$$
for $\mu \not\in B(W\lambda, \beta/2t)$ and $\nu \in B(W\lambda, \beta/4t)$, then (20) would follow by combining (24–26), inverting the Harish–Chandra transform, and using the rapid decay of $h_t$ away from the set $Wt\lambda$ as in [13, Proof of Lemma 4.4] or [14, Lemma 2.6].

We decompose the left-hand side of (26) as

$$\int_a^b b_2(H)b_1(\beta H)\varphi_t,\exp(H)) e^{-iH\mu} dH$$

$$+ \sum_{n=1}^{\infty} \int_a^b b_2(H)(b_1(2^{-n}\beta H) - b_1(2^{-n+1}\beta H)))\varphi_t,\exp(H)) e^{-iH\mu} dH.$$ 

The bound (21) implies that the first integral is $\ll t^{-3/2} \beta^{-1/2+\epsilon}$ as in [14, Section 2.5]. As $\nu \in B_0^*$ for $t$ large, we may estimate the second integral by applying (23). We shall only consider the term $w = e$, as the others are identical. We wish to estimate

$$\int_a^b b_2(H)(b_1(2^{-n}\beta H) - b_1(2^{-n+1}\beta H))) f_e(H, v, t) \exp(it(v-\mu)(H)) dH.$$ 

After replacing $H$ with $2^n\beta^{-1}H$ this becomes

$$2^{2n}\beta^{-2}\int_a^b b_2(2^n\beta^{-1}H)(b_1(H) - b_1(2H))) f_e(2^n\beta^{-1}H, v, t) \exp(it2^n\beta^{-1}(v-\mu)(H)) dH. \quad (27)$$ 

Our assumptions on $\mu$ and $\nu$ imply that $t2^n\beta^{-1}\|v-\mu\| \geq 2^{n-2}$, and (22) implies that all derivatives of the amplitude factor in (27) are

$$\ll t^{-3/2} \beta^{3/2} 2^{-3n/2} \prod_{\alpha \in \Delta^+} |\alpha(H)|^{-1/2}. \quad (28)$$

Integrating by parts once, we obtain

$$(27) \ll t^{-3/2} \beta^{-1/2} 2^{-n/2} \int_a^b b_2(2^n\beta^{-1}H)(b_1(H) - b_1(2H))) \prod_{\alpha \in \Delta^+} |\alpha(H)|^{-1/2} dH$$

$$\ll t^{-3/2} \beta^{-1/2+\epsilon} 2^{(-1/2-\epsilon)n}.$$ 

Summing over $n$ completes the proof.

6 Oscillatory Integrals

We now prove the bounds of Proposition 4.3 for $I(t, \phi, g)$. We continue to use the notation introduced earlier in the paper, with the exceptions that $B^*$ now denotes a fixed compact subset of $\mathfrak{a}_r^*$ that contains an open neighborhood of the set $B^*$ used earlier, and we now
set $K = \text{SO}(3)$. If $\nu \in a^*$, we let $H_\nu \in a$ denote the element dual to it under the Killing form. We write the Iwasawa decomposition on $\text{SL}_3(\mathbb{R})$ as

$$g = n(g) \exp(A(g))k(g) = \exp(N(g)) \exp(A(g))k(g).$$  \hspace{1cm} (28)$$

We begin by calculating a certain derivative of the function $A$.

**Lemma 6.1.** Let $g \in \text{SL}_3(\mathbb{R})$ have Iwasawa decomposition $g = nak$. If $H_1, H_2 \in a$, then

$$\frac{\partial}{\partial s} \langle H_1, A(g \exp(sH_2)) \rangle \Big|_{s=0} = \langle H_1, \text{Ad}(k)H_2 \rangle.$$  \hspace{1cm} □

**Proof.** We have

$$\frac{\partial}{\partial s} \langle H_1, A(g \exp(sH_2)) \rangle \Big|_{s=0} = \frac{\partial}{\partial s} \langle H_1, A(k \exp(sH_2)) \rangle \Big|_{s=0}$$

$$= \frac{\partial}{\partial s} \langle H_1, A(\exp(s\text{Ad}(k)H_2)) \rangle \Big|_{s=0}$$

$$= \langle H_1, \text{Ad}(k)H_2 \rangle.$$  \hspace{1cm} ■

For $g \in \text{SL}_3(\mathbb{R})$, let $\Phi_g : K \to K$ be the map sending $k$ to $k(kg)$.

**Lemma 6.2.** $\Phi_g$ is a diffeomorphism that depends smoothly on $g$.  \hspace{1cm} □

**Proof.** The smoothness of the Iwasawa decomposition implies that $\Phi_g$ is smooth, and depends smoothly on $g$. The inclusion

$$\Phi_g(k)g^{-1} \in NAk$$  \hspace{1cm} (29)

implies that $\Phi_g$ is injective. To prove surjectivity, let $k_0 \in K$ and define $k_1 \in K$ by $k_0g^{-1} \in NAk_1$. The condition (29) implies that $\Phi_g(k_1)g^{-1} \in NAk_1 = NAk_0g^{-1}$. This implies that $\Phi_g(k_1) \in NAk_0$, so $\Phi_g(k_1) = k_0$, and $\Phi_g$ is surjective. If $\Phi_g^{-1}$ is the inverse function, (29) gives $kg^{-1} \in NA\Phi_g^{-1}(k)$ which implies that $\Phi_g^{-1}$ is smooth.  \hspace{1cm} ■

If $g \in NA$ and $m \in M$, we have

$$\Phi_g(m) = m, \quad \Phi_g(mk) = m\Phi_g(k).$$  \hspace{1cm} (30)

We now estimate two integrals that comprise $I(t, \phi, g)$. 

---

Proposition 6.3. Fix $C, \epsilon > 0$, a compact set $D_B \subset NA$, and a function $b \in C_0^\infty(a)$ with $\text{supp}(b) \subseteq B_a$. If $k \in K$ and $\lambda, \nu \in B^*$ satisfy $\|\lambda - \nu\| \leq \beta/t$ and

$$k \notin M \exp(C t^{-1/2 + \epsilon} \beta^{1/2} B_t),$$

and $g \in D_B$, then we have

$$\int_a^b (H) \exp(-it\lambda(H) + itv(A(kg \exp(H)))) dH \ll_A t^{-A}. \quad (32)$$

The implied constant depends on $B^*, D_B, C, \epsilon$, and the size of the derivatives of $b$ of order at most $n$, where $n$ depends on $\epsilon$ and $A$. \hfill \square

Proof. Define

$$\psi(H, k, g) = \lambda(H) - v(A(kg \exp(H))), \quad H \in B_a, \ k \in K, \ g \in D_B,$$

to be the phase of the integral (32). Throughout the proof, the variables $g$ and $H$ will always be restricted to $D_B$ and $B_a$, respectively. We shall always let $H_v \psi$ denote the derivative of $\psi$ with respect to the variable $H$. Lemma 6.1 implies that

$$H_v \psi(H, k, g) = \langle H_v, H_\lambda \rangle - \langle H_v, \text{Ad}(\Phi^g_{\exp(H)}(k))H_v \rangle$$

$$= \langle H_v, H_\lambda - H_v \rangle + \langle H_v, (1 - \text{Ad}(\Phi^g_{\exp(H)}(k)))H_v \rangle. \quad (33)$$

It is proved in [8, Proposition 5.4] that the function $\langle H_v, \text{Ad}(k)H_v \rangle$ has a critical point at the identity, and it follows from [8, Proposition 6.5] or [14, Proposition 4.4] that the Hessian at this critical point is negative definite, uniformly for $\nu \in B^*$. It follows that, given $B^*$, we may choose an open neighborhood $0 \in U^1 \subset \mathfrak{k}$ such that for $X \in U^1$ we have

$$\langle H_v, H_v \rangle - \langle H_v, e^{ad(X)}H_v \rangle \sim_{B^*} \|X\|^2, \quad (34)$$

and $\exp$ gives a diffeomorphism $U^1 \rightarrow U_1 := \exp(U^1)$. Let $0 \in U_t \subset \mathfrak{k}$ be an open set such that $\exp$ gives a diffeomorphism $U_t \rightarrow U := \exp(U_t)$, and

$$U \subseteq \bigcap_{h \in D_B \exp(B_a)} \Phi^{-1}_h(U_1). \quad (35)$$
Assume that $k \notin MU$. Theorem 8.2 of [12] implies that if $k_0 \in K$, we have $\langle H_v, \text{Ad}(k_0)H_v \rangle \leq \langle H_v, H_v \rangle$ with equality iff $k_0 \in M$. It follows from this and (30) that there exists $\delta > 0$ depending only on $B^*$ and $D_B$ such that

$$\langle H_v, (1 - \text{Ad}(\Phi_{g\exp(H)}(k)))H_v \rangle > \delta. \quad (36)$$

The condition $\|\lambda - \nu\| \leq \beta/t$ implies that

$$|\langle H_v, H_{\lambda} - H_v \rangle| \ll_B \beta/t \leq t^{-1/2}, \quad (37)$$

and combined with (33) and (36) this implies that $H_v \psi \geq \delta/2$ for $t$ sufficiently large. The bound (32) now follows by integration by parts.

Now assume that $k \in MU$. In the rest of the proof, all implied constants will depend on $B^*$ and $D_B$. As $\psi(H, mk, g) = \psi(H, k, g)$ for $m \in M$, we may assume that $k = \exp(X) \in U$ with $X \in U_t$. The assumption (31) implies that

$$\|X\|^2 \geq C^2 t^{-1+2\epsilon} \beta. \quad (38)$$

The condition (35) on $U$ implies that $\Phi_{g\exp(H)}(k) \in U_1$, and so we may define $Y(H, k, g) \in U_1^t$ by requiring that $\Phi_{g\exp(H)}(k) = \exp(Y(H, k, g))$. Because $\Phi_{g\exp(H)}$ is a smoothly varying family of diffeomorphisms fixing $e$, we have $\|Y(H, k, g)\| \sim \|X\|$, and hence (34) gives

$$\langle H_v, (1 - \text{Ad}(\Phi_{g\exp(H)}(k)))H_v \rangle = \langle H_v, (1 - e^{\text{ad}(Y(H,g,k))})H_v \rangle \gg \|X\|^2.$$

It follows by combining this with (37) and (38) that $H_v \psi \gg \|X\|^2$. Combining (33) and (34) gives

$$H_v \psi(H, k, g) - \langle H_v, H_{\lambda} - H_v \rangle \ll \|X\|^2,$$

which implies that $H_v \psi(H, k, g) - \langle H_v, H_{\lambda} - H_v \rangle$ vanishes to second order at $k = e$. It follows that

$$H^n_v \psi(H, k, g) \ll_n \|X\|^2$$

for $n \geq 2$. As $t\|X\|^2 \geq C^2 t^{2\epsilon} \beta$, the result now follows by integration by parts with respect to $H_v$. ■
Proposition 6.4. Fix $C > 0$, $\frac{1}{10} > \epsilon > 0$, a compact set $D_B \subset NA$, and a function $b \in C_0^\infty(a)$ with $\text{supp}(b) \subseteq B_a$. If $g = na \in D_B$ and $\lambda, \nu \in B^*$ satisfy $\|\lambda - \nu\| \leq \beta/t$ and

$$n \notin \exp(Ct^{-1/2+\epsilon}b^{1/2}B_n),$$

then

$$\int_a b(H)e^{-it\lambda(H)}\varphi_{tv}(g \exp(H)) \, dH \ll_A t^{-A}. \tag{39}$$

The implied constant depends on $B^*$, $D_B$, $C$, $\epsilon$, and the size of the derivatives of $b$ of order at most $n$, where $n$ depends on $\epsilon$ and $A$. \qed

Proof. If we substitute the formula for $\varphi_{tv}$ as an integral of plane waves into (40), it becomes

$$\int_a \int_K b(H) \exp(-it\lambda(H) + (\rho + it\nu)(A(kg \exp(H)))) \, dk \, dH. \tag{40}$$

Choose a function $f \in C_0^\infty(t)$ with $\text{supp}(f) \subseteq 2B_t$ and $f(X) = 1$ for $X \in B_t$. Let $C_1 > 0$ be a constant to be chosen later, define $b_1 \in C^\infty(K)$ to be the pushforward of $f(C_1^{-1}t^{1/2-\epsilon}b^{-1/2}X)$ under $\exp$, and define $b_2$ by $b_2(k) = 1 - \sum_{m \in M} b_1(mk)$. If $t$ is sufficiently large depending on $C_1$, we have

$$\text{supp}(b_1) \subseteq \exp(2C_1t^{-1/2+\epsilon}b^{1/2}B_t),$$

$$\text{supp}(b_2) \subseteq K \setminus M \exp(C_1t^{-1/2+\epsilon}b^{1/2}B_t).$$

Proposition 6.3 and the assumption $\|\lambda - \nu\| \leq \beta/t$ imply that

$$\int_a \int_K b_2(k)b(H) \exp(-it\lambda(H) + (\rho + it\nu)(A(kg \exp(H)))) \, dk \, dH \ll_A t^{-A}.$$ 

It therefore suffices to prove that

$$\int_a \int_K b_1(mk)b(H) \exp(-it\lambda(H) + (\rho + it\nu)(A(kg \exp(H)))) \, dk \, dH \ll_A t^{-A}$$

for $m \in M$, and we assume without loss of generality that $m = e$. We shall do this by proving the bound

$$\int_K b_1(k) \exp(it\nu(A(kg))) \, dk \ll_A t^{-A} \tag{41}$$

for the integrals over $K$. Note that we have absorbed $\exp(H)$ into $g$ after enlarging $D_B$, and absorbed the factor $\exp(\rho(A(kg \exp(H))))$ into $b_1(k)$ and suppressed the dependence
on \( H \) and \( g \). The implied constants in the above three bounds have the same dependence as (40).

We define the phase functions

\[
\psi(k, g) = \nu(A(kg)), \quad k \in K, \ g \in NA,
\]

and

\[
\psi_S(x) = \nu(A(x)), \quad x \in S.
\]

If \( X \in g \), we let \( X^S \) be the vector field on \( S \) whose value at \( x \in S \) is \((\partial/\partial t) \exp(tX)|_{t=0}\). It may be shown that these vector fields satisfy \([X^S, Y^S] = -[X, Y]^S\), where the first Lie bracket is on \( S \) and the second is in \( g \). We choose \( X_\alpha \in \mathfrak{g}_\alpha \) for each \( \alpha \in \Delta \) so that the relations \( X_{-\alpha} = \theta X_\alpha \) and \( \langle X_\alpha, X_{-\alpha} \rangle = -1/2 \) are satisfied, where \( \theta \) is the Cartan involution on \( g \). If \( \alpha \in \Delta^+ \), we let \( K_\alpha = X_\alpha + X_{-\alpha} \in \mathfrak{t} \). It may be seen that the fields \( \{X^S_\alpha | \alpha \in \Delta^+\} \) form a basis for the normal bundle to \( A \) in \( S \).

Proposition 5.4 of [8] implies that the set of \( x \in S \) where the functions \( \{K^S_\alpha \psi_S | \alpha \in \Delta^+\} \) vanish simultaneously is exactly \( A \). The following lemma shows that these functions in fact form a coordinate system transversely to \( A \).

**Lemma 6.5.** If \( \alpha, \beta \in \Delta^+ \), and \( x \in A \), we have \( X^S_\alpha K^S_\beta \psi_S(x) = \delta_{\alpha\beta} \langle \alpha, \nu \rangle \psi_S(x)/2 \). \( \square \)

**Proof.** We have

\[
X^S_\alpha K^S_\beta \psi_S = K^S_\beta X^S_\alpha \psi_S + [X^S_\alpha, K^S_\beta] \psi_S.
\]

The first term vanishes, as \( X^S_\alpha \psi_S \equiv 0 \). If \( \alpha = \beta \), then

\[
[X^S_\alpha, K^S_\alpha] = -[X_\alpha, K_\alpha]^S = \frac{1}{2} H^S_\alpha,
\]

and the lemma follows. If \( \alpha \neq \beta \), then \( [X_\alpha, K_\beta] \in \mathfrak{a}^\perp \), so along \( A \) we have

\[
[X_\alpha, K_\beta]^S \in \text{span}(X^S_\gamma | \gamma \in \Delta^+).
\]

As the fields \( X^S_\gamma \) annihilate \( \psi_S \), the lemma follows. \( \blacksquare \)

Fix a compact set \( D_A \subset A \) such that \( D_B \subset ND_A \), and a second compact set \( D'_A \subset A \) that contains an open neighborhood of \( D_A \). Let \( S_n \) be the unit sphere in \( n \) with respect to \( \| \cdot \| \). We have the following corollary of Lemma 6.5.
Corollary 6.6. There exist open sets $U'_a \subset S_n$ for each $\alpha \in \Delta^+$, and $\sigma$, $\delta > 0$, such that $S_n = \bigcup U'_a$, and if $\alpha \in \Delta^+$ and $x \in \exp([0, \sigma) U'_a) D'_A \subset S$, then

$$|K^S_{\psi_S}(x)| \geq \delta \|N(x)\|,$$

where $N(x)$ is as in (28). The data $U'_a$, $\delta$, and $\sigma$ depend on $B^*$ and $D'_A$. \hfill $\square$

Suppose that $g \in \exp(\sigma B/2) D_A$. For each $\alpha \in \Delta^+$, we choose a second open set $U_\alpha \subset S_n$ such that $\bar{U}_\alpha \subset U'_a$ and we still have $S_n = \bigcup U_\alpha$. Let $\alpha \in \Delta^+$ be such that $g \in \exp([0, \sigma/2) U_\alpha) D_A$. The assumption (39) is equivalent to the bound $||N(g)|| \geq C t^{-1/2+\epsilon} \beta^{1/2}$, and we may assume that $k \in \text{supp}(b_1) \subset \exp(2C_1 t^{-1/2+\epsilon} \beta^{1/2} B_t)$. These imply that if $C_1$ is sufficiently small depending on $C$ and $D_B$, we have $kg \in \exp([0, \sigma) U'_a) D'_A$, and there is $C_2 = C_2(C, C_1, D_B) > 0$ such that $||N(kg)|| \geq C_2 t^{-1/2+\epsilon} \beta^{1/2}$. It follows from (42) that

$$|K_\alpha \psi(k, g)| = |K^S_{\psi_S}(kg)| \geq \delta ||N(kg)|| \geq \delta C_2 t^{-1/2+\epsilon} \beta^{1/2}$$

for $k \in \text{supp}(b_1)$. The bound (41) now follows from $K^n_\alpha \psi(k, g) \ll C_3(B^*, D_B, n)$, and an application of Lemma 6.7 below with $\delta = t^{-1/2+\epsilon} \beta^{1/2}$.

Suppose that $g \in D_B \setminus \exp(\sigma B/2) D_A$. The fact that the simultaneous vanishing set of the functions $K^S_{\psi_S}$ is exactly $A$ implies that there exists an $\alpha \in \Delta^+$, and $C_A(B^*, D_B, \sigma) > 0$, such that $|K_\alpha \psi(k, g)| \geq C_4$ for $k \in \text{supp}(b_1)$ and $C_1$ sufficiently small depending on $B^*$, $D_B$, and $\sigma$. The proposition again follows by integration by parts. \hfill $\blacksquare$

Lemma 6.7. If $b \in C^\infty_0(\mathbb{R})$ is a cutoff function at scale $1 \geq \delta > 0$, and $\phi \in C^\infty(\mathbb{R})$ satisfies

$$\phi'(x) \gg \delta, \quad \phi^{(n)}(x) \ll_n 1$$

for $x \in \text{supp}(b)$ and $n > 1$, then

$$\int b(x) e^{it\phi(x)} \, dx \ll \delta (t\delta^2)^{1-A}. \hfill \square$$

Proof. We have

$$\int b(x) e^{it\phi(x)} \, dx = \delta \int b(\delta x) e^{i(t\delta^2)\phi(\delta x)} \, dx.$$ 

The function $b(\delta x)$ is now a cutoff function at scale $1$, and $\delta^{-2} \phi(\delta x)$ satisfies

$$\frac{d}{dx} \delta^{-2} \phi(\delta x) \gg 1, \quad \left( \frac{d}{dx} \right)^n \delta^{-2} \phi(\delta x) \ll_n \delta^{n-2} \leq 1$$

for $x \in \text{supp}(b(\delta \cdot))$ and $n > 1$. The lemma now follows by integration by parts. \hfill $\blacksquare$
We now combine Propositions 6.3 and 6.4 to prove the following, which will imply Proposition 4.3 after inverting the various transforms.

**Proposition 6.8.** Fix functions $b_1, b_2 \in C_0^\infty(a)$ with $\text{supp}(b_1) \subseteq B_a$. If $g \in \text{SL}_3(\mathbb{R})$ and $\nu, \lambda_1, \lambda_2 \in B^*$ satisfy

\[ d(g, e) \leq 1 \quad \text{and} \quad \|\lambda_i - \nu\| \leq \beta/\nu, \]

then we have

\[ \int_{a} b_1(H_1)b_2(H_2) e^{it(\lambda_1(H_1) - \lambda_2(H_2))} \varphi_{tv}(\exp(-H_1)g \exp(H_2)) \, dH_1 \, dH_2 \ll t^{-3/2+\epsilon/2} \beta^{3/2} \quad (43) \]

for any $\epsilon > 0$. If we also have $d(g, MA) \geq t^{-1/2+\epsilon/2} \beta^{1/2}$ for some $\epsilon > 0$, then

\[ \int_{a} b_1(H_1)b_2(H_2) e^{it(\lambda_1(H_1) - \lambda_2(H_2))} \varphi_{tv}(\exp(-H_1)g \exp(H_2)) \, dH_1 \, dH_2 \ll_A t^{-A}. \quad (44) \]

The implied constants depend only on $B^*, \epsilon$, and the size of the derivatives of $b_i$ of order at most $n$, where $n$ depends on $\epsilon$ and $A$. 

**Proof.** If we unfold the integral over $K$ defining $\varphi_{tv}$ in the LHS of (43), we obtain the integral

\[ \int_{a} \int_{K} b_1(H_1)b_2(H_2) e^{it(\lambda_1(H_1) - \lambda_2(H_2))} \exp((\rho + it\nu)(A(k \exp(-H_1)g \exp(H_2)))) \, dk \, dH_1 \, dH_2. \quad (45) \]

Choose constants $C_1, \epsilon > 0$, and assume that the variable $k$ in (45) satisfies $k \notin M \exp(C_1 t^{-1/2+\epsilon} \beta^{1/2} B_t)$. After writing $g \exp(H_2) = n' a' k'$, we see Proposition 6.3 implies that the partial integral of (45) with respect to $H_1$ is $\ll_A t^{-A}$ for this fixed value of $k$ and any $H_2 \in B_a$. This implies that we may restrict the integral over $K$ in (45) to $M \exp(C_1 t^{-1/2+\epsilon} \beta^{1/2} B_t)$, and this gives the bound (43).

We now prove (44). Let $\epsilon > 0$ be given, and assume that $d(g, MA) \geq t^{-1/2+\epsilon/2} \beta^{1/2}$. We may assume without loss of generality that $\epsilon < \frac{1}{10}$. Let $g = k_1 n_1 a_1$, and let

\[ \exp(H)g = k_1(H) n_1(H) a_1(H) \]

for $H \in a$. Choose a constant $C_2 > 0$, and assume that $k_1 \notin M \exp(C_2 t^{-1/2+\epsilon} \beta^{1/2} B_t)$. Then there is an absolute constant $\kappa > 0$ such that we have

\[ k_1(H) \notin M \exp(\kappa C_2 t^{-1/2+\epsilon} \beta^{1/2} B_t) \quad (46) \]
for \( H \in B_a \). Let \( k \in K \), and consider the integral
\[
\int_a b_1(H_1)b_2(H_2)e^{i(H_1-H_2)} \exp((\rho + iv)A(k \exp(-H_1)g \exp(H_2))) \, dH_1 \, dH_2. \tag{47}
\]
The condition (46) implies that if \( C_1 \) is sufficiently small depending on \( C_2 \), one of
\[
k \not\in M \exp \left( C_1 t^{-1/2+\epsilon} \beta^{1/2} B_t \right) \tag{48}
\]
\[
kk_1(-H_1) \not\in M \exp \left( C_1 t^{-1/2+\epsilon} \beta^{1/2} B_t \right) \quad \text{for all } H_1 \in B_a \tag{49}
\]
must hold. If (48) holds, Proposition 6.3 applied to the integral of (47) over \( H_1 \) implies that (47) \( \ll_A t^{-A} \). If (49) holds, Proposition 6.3 applied to the integral over \( H_2 \) also implies (47) \( \ll_A t^{-A} \). Integrating these bounds over \( K \) gives (44).

We may therefore assume that \( k_1 \in M \exp \left( C_2 t^{-1/2+\epsilon} \beta^{1/2} B_t \right) \) for any \( C_2 > 0 \), and hence that
\[
k_1(H) \in M \exp \left( C_2 t^{-1/2+\epsilon} \beta^{1/2} B_t \right) \tag{50}
\]
for any \( C_2 > 0 \) and \( H \in B_a \). Our assumption that \( d(g, MA) \geq t^{-1/2+\epsilon} \beta^{1/2} \) implies that there is an absolute \( C_3 > 0 \) such that
\[
d(\exp(H)g, MA) \geq C_3 t^{-1/2+\epsilon} \beta^{1/2} \tag{51}
\]
for \( H \in B_a \). If \( C_2 \) is sufficiently small, (50) and (51) imply that there is an absolute \( C_4 > 0 \) such that \( d(n_1(H), e) \geq C_4 t^{-1/2+\epsilon} \beta^{1/2} \) for \( H \in B_a \). The result now follows by applying Proposition 6.4 to the integral of the LHS of (44) over \( H_2 \) for each fixed \( H_1 \).

\[\square\]

**Corollary 6.9.** Fix functions \( b_1, b_2 \in C_0^\infty(a) \) with \( \text{supp}(b_i) \subseteq B_a \). Let \( \nu \in B^* \), and choose \( \phi \in \mathcal{H}(tv, \beta) \) with \( \|\phi\|_2 = 1 \). If \( g \in \text{SL}_3(\mathbb{R}) \) satisfies \( d(g, e) \leq 1 \), we have
\[
\int_a b_1(H_1)b_2(H_2)\varphi_{tv}(\exp(-H_1)g \exp(H_2)) \, dH_1 \, dH_2 \ll t^{-3/2+\epsilon} \beta^{9/2}
\]
for any \( \epsilon > 0 \). If in addition we have \( d(g, MA) \geq t^{-1/2+\epsilon} \beta^{1/2} \) for some \( \epsilon > 0 \), then
\[
\int_a b_1(H_1)b_2(H_2)\varphi_{tv}(\exp(-H_1)g \exp(H_2)) \, dH_1 \, dH_2 \ll_A t^{-A}.
\]
The implied constants depend only on \( B^*, \epsilon \), and the size of the derivatives of \( b_i \) of order at most \( n \), where \( n \) depends on \( \epsilon \) and \( A \).
Proof. The follows immediately from Proposition 6.8 after inverting the Fourier transform and noting that \( \| \hat{\phi} \|_1 \leq \| \hat{\phi} \|_2 (4 \pi \beta^3 / 3)^{1/2} \ll \beta^{3/2} \).

Proposition 4.3 now follows by inversion of the Harish–Chandra transform as in [13, Section 6.3] or [14, Lemma 2.6].

Funding

This work was supported by the National Science Foundation [grant number DMS-1201321].

Acknowledgement

We thank Valentin Blomer, Zeev Rudnick, and Matthew Stover for helpful comments, and the referee for a careful reading of the manuscript.

References


