J. London Math. Soc. (2) 102 (2020) 143–204 doi:10.1112/jlms.12316

An effective equidistribution result for $\mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$ and application to inhomogeneous quadratic forms

Andreas Strömbergsson and Pankaj Vishe

Abstract

Let $G = \mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$ and let Γ be a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z}) \ltimes (\mathbb{Z}^2)^{\oplus k}$. We prove a polynomially effective asymptotic equidistribution result for special types of unipotent orbits in $\Gamma\backslash G$ which project to pieces of closed horocycles in $\mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R})$. As an application, we prove an effective quantitative Oppenheim-type result for the quadratic form $(m_1-\alpha)^2 + (m_2-\beta)^2 - (m_3-\alpha)^2 - (m_4-\beta)^2$, for $(\alpha,\beta) \in \mathbb{R}^2$ of Diophantine type, following the approach by Marklof [*Ann. of Math.* 158 (2003) 419–471] using theta sums.

1. *Introduction*

The results of Ratner on measure rigidity and equidistribution of orbits of a unipotent flow [**31, 32**], play a fundamental role in homogeneous dynamics. These results also have many applications outside of dynamics, ranging from problems in number theory to mathematical physics. In recent years, there has been an increased interest in obtaining *effective* versions of Ratner's results in special cases, that is, to provide an explicit rate of density or equidistribution for the orbits of a (non-horospherical) unipotent flow (cf. [**3, 6, 13, 20, 27, 30, 38**]). In particular, in [**3, 38**], effective equidistribution results were obtained for orbits of a 1-parameter unipotent flow on $\mathrm{SL}(2,\mathbb{Z}) \ltimes \mathbb{Z}^2 \backslash \mathrm{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$, using Fourier analysis and methods of from analytic number theory, and in the very recent paper [**30**], building on similar methods, effective equidistribution of diagonal translates of certain orbits in $\mathrm{SL}(3,\mathbb{Z}) \ltimes \mathbb{Z}^3 \backslash \mathrm{SL}(3,\mathbb{R}) \ltimes \mathbb{R}^3$ was established. Our purpose in the present paper is to prove results of a similar nature for homogeneous spaces of the group $G = \mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$ for $k \geqslant 2$, and to apply these to derive an effective quantitative Oppenheim-type result for a certain family of inhomogeneous quadratic forms of signature (2,2). Here, $(\mathbb{R}^2)^{\oplus k}$ denotes the direct sum of k copies of $\mathbb{R}^2$, each provided with the standard action of $\mathrm{SL}(2,\mathbb{R})$.

We now turn to a precise description of our setting. We represent vectors by column matrices. Throughout the paper, we will identify $(\mathbb{R}^2)^{\oplus k}$ with $\mathbb{R}^{2k}$, so that the action of $G' := \mathrm{SL}(2,\mathbb{R})$ is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{x}' \end{pmatrix} = \begin{pmatrix} a\boldsymbol{x} + b\boldsymbol{x}' \\ c\boldsymbol{x} + d\boldsymbol{x}' \end{pmatrix} \quad \text{for } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G',\ \boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^k.$$

The elements of

$$G = \mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$$

Received 10 December 2018; revised 20 February 2020; published online 7 April 2020.

2010 *Mathematics Subject Classification* 37A17, 37A45 (primary), 11E20, 11K60 (secondary).

A.S. was supported by the Swedish Research Council Grant 2016-03360. P.V. was partly supported by the Göran Gustafsson Foundation (KVA) at KTH and by the EPSRC programme grant EP/J018260/1.

© 2020 The Authors. *Journal of the London Mathematical Society* is copyright © London Mathematical Society. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

are then represented by pairs $(M, \boldsymbol{v}) \in G' \times \mathbb{R}^{2k}$, with a multiplication law

$$(M, \boldsymbol{v})(M', \boldsymbol{v}') = (MM', \boldsymbol{v} + M\boldsymbol{v}').$$

Let

$$a(y) = \begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix} \quad \text{and} \quad u(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \quad (y > 0,\ x \in \mathbb{R}).$$

We will always view $G' = \mathrm{SL}(2, \mathbb{R})$ as a subgroup of G through $M \mapsto (M, \mathbf{0})$; in particular, $a(y)$ and $u(x)$ are also elements of G. We set

$$\overline{\Gamma} = \mathrm{SL}(2, \mathbb{Z}) \ltimes (\mathbb{Z}^2)^{\oplus k}.$$

In our notation, this is the subgroup of all $(M, \boldsymbol{v}) \in G$ with $M \in \mathrm{SL}(2, \mathbb{Z})$ and $\boldsymbol{v} \in \mathbb{Z}^{2k}$. Given a subgroup Γ of $\overline{\Gamma}$ of finite index, we consider the homogeneous space

$$X = \Gamma \backslash G.$$

As we will detail below, this space is a torus bundle over a finite cover of the familiar 3-dimensional homogeneous space $\mathrm{SL}(2, \mathbb{Z}) \backslash \mathrm{SL}(2, \mathbb{R})$ classifying unimodular lattices in $\mathbb{R}^2$. We fix μ to be the (left and right invariant) Haar measure on G, normalized so as to induce a probability measure on X, which we also denote by μ.

The following equidistribution result is a special case of [5, Theorem 3][†]; alternatively it may be deduced (with some work) as a consequence of [36, Theorem 1.4]. Note that both [5, Theorem 3] and [36, Theorem 1.4] depend crucially on Ratner's classification of invariant measures.

For any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^k$ we denote by $\boldsymbol{ab}$ the standard scalar product, $\boldsymbol{ab} = a_1 b_1 + \cdots + a_k b_k$.

THEOREM 1.1. *Let Γ be a subgroup of $\overline{\Gamma} = \mathrm{SL}(2, \mathbb{Z}) \ltimes (\mathbb{Z}^2)^{\oplus k}$ of finite index. Fix $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{pmatrix}$ in $\mathbb{R}^{2k}$ subject to the condition that there does not exist any $\boldsymbol{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ for which both $\boldsymbol{m}\boldsymbol{\xi}_1$ and $\boldsymbol{m}\boldsymbol{\xi}_2$ are integers. Then for any Borel probability measure λ on $\mathbb{R}$ which is absolutely continuous with respect to the Lebesgue measure, and any bounded continuous function f on $X = \Gamma \backslash G$,*

$$\lim_{y \to 0^+} \int_{\mathbb{R}} f(\Gamma(1_2, \boldsymbol{\xi}) u(x) a(y))\, d\lambda(x) = \int_X f\, d\mu. \tag{1}$$

In view of the relation

$$u(x)a(y) = a(y)u(y^{-1}x),$$

the integration in the left-hand side of (1) is along an orbit of the unipotent flow

$$U^t : \Gamma g \mapsto \Gamma g u(t) \quad (t \in \mathbb{R})$$

on X. Let $D : G \to G'$ be the natural projection sending $(M, \boldsymbol{v})$ to M; then $D(\Gamma)$ is a finite index subgroup of $\mathrm{SL}(2, \mathbb{Z})$, and D induces a projection map from X to $X' := D(\Gamma) \backslash G'$, which we also call D; this realizes X as a torus bundle over the space X', which in turn is a finite cover of $\mathrm{SL}(2, \mathbb{Z}) \backslash \mathrm{SL}(2, \mathbb{R})$. The orbits which appear in (1) are exactly those orbits of the flow

[†]Apply [5, Theorem 3] with $d = 2$ and $M = 1_2$ and use the anti-automorphism $(M, \left(\begin{smallmatrix} \boldsymbol{x} \\ \boldsymbol{x}' \end{smallmatrix}\right)) \mapsto ({}^t M, {}^t M \left(\begin{smallmatrix} \boldsymbol{x}' \\ -\boldsymbol{x} \end{smallmatrix}\right))$ of G to translate from the setting with G/Γ in [5] into our setting with $X = \Gamma \backslash G$. As noted in [5, Remark 7.2], the proof of [5, Theorem 3] extends trivially to the case when Γ is an arbitrary subgroup of $\mathrm{SL}(2, \mathbb{Z}) \ltimes (\mathbb{Z}^2)^{\oplus k}$ of finite index.

U^t which project to a closed horocycle in X' around its cusp at ∞. Letting y decrease toward zero means that we are considering expanding translates of the initial orbit $x \mapsto \Gamma(1_2, \boldsymbol{\xi})u(x)$.

Let us note that the condition imposed on $\boldsymbol{\xi}$ in Theorem 1.1 cannot be weakened. Indeed, for any $\boldsymbol{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, set

$$X_{\boldsymbol{m}} := \left\{ \Gamma\left(M, \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix}\right) \ : \ M \in \mathrm{SL}(2,\mathbb{R}),\ \boldsymbol{m}\boldsymbol{v}_1 \in \mathbb{Z},\ \boldsymbol{m}\boldsymbol{v}_2 \in \mathbb{Z} \right\}.$$

This is a closed embedded submanifold of codimension 2 in X. If both $\boldsymbol{m}\boldsymbol{\xi}_1$ and $\boldsymbol{m}\boldsymbol{\xi}_2$ are integers, then

$$\Gamma(1_2, \boldsymbol{\xi})u(x)a(y) \in X_{\boldsymbol{m}} \quad \text{for all } \ x \in \mathbb{R},\ y > 0, \tag{2}$$

and therefore the curve certainly cannot become equidistributed in X, that is, (1) fails for some f. (For example, consider any bounded continuous $f \geqslant 0$ such that $f_{|X_{\boldsymbol{m}}} \equiv 0$ while $\int_X f\, d\mu > 0$.)

Marklof in [**24**, Theorem 5.7] proved Theorem 1.1 in the special case of $\boldsymbol{\xi}_1 = \mathbf{0}$, and then in [**25**, Theorem 3.1] in the special case of $\boldsymbol{\xi}_2 = \mathbf{0}$. Note that if $\boldsymbol{\xi}_1 = \mathbf{0}$, the condition on $\boldsymbol{\xi}_2$ in the theorem becomes that 1 together with the k components of $\boldsymbol{\xi}_2$ should be linearly independent over $\mathbb{Q}$ (and vice versa if $\boldsymbol{\xi}_2 = \mathbf{0}$). Our main results in the present paper are in Theorems 1.2 and 1.3, which give *effective* versions of these two special cases of Theorem 1.1, under the further requirement that Γ is a congruence subgroup of $\overline{\Gamma}$.

To prepare for the statement of the main theorems we introduce some further notation. For a positive integer N, $\Gamma(N)$ denotes the principal congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$ of level N:

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod N \right\}.$$

We will consider $X = \Gamma\backslash G$ where Γ is a subgroup of $\overline{\Gamma}$ of the form $\Gamma = \Gamma(N) \ltimes \mathbb{Z}^{2k}$. (The case of an arbitrary congruence subgroup of $\overline{\Gamma}$ can easily be reduced to the case of $\Gamma = \Gamma(N) \ltimes \mathbb{Z}^{2k}$, by using the fact that for any $q \in \mathbb{Z}^+$, the map $(M, \boldsymbol{v}) \mapsto (M, q\boldsymbol{v})$ is an automorphism of G.)

We introduce the following cuspidal height function, for $(M, \boldsymbol{v}) \in G$:

$$\mathcal{Y}(M, \boldsymbol{v}) = \mathcal{Y}(M) = \sup\{\mathrm{Im}\, \gamma M(i) \ : \ \gamma \in \mathrm{SL}(2,\mathbb{Z})\}, \tag{3}$$

where in the right-hand side, we use the standard action of $G' = \mathrm{SL}(2,\mathbb{R})$ on the Poincaré upper half plane $\mathbb{H} = \{\tau = u + iv \in \mathbb{C} \ : \ v > 0\}$. Then $\mathcal{Y}(M, \boldsymbol{v}) \geqslant \sqrt{3}/2$ for all $(M, \boldsymbol{v}) \in G$. Note that $\mathcal{Y}(M, \boldsymbol{v})$ depends only on the coset $\overline{\Gamma}(M, \boldsymbol{v})$, and in particular $\mathcal{Y}$ can be viewed as a function on X. Given $p_1, p_2, \ldots \in X$, we have $\mathcal{Y}(p_j) \to \infty$ if and only if the sequence $p_1, p_2, \ldots$ leaves all compact subsets of X.

For $m \geqslant 0$ and $a \in \mathbb{R}$, we let $\mathrm{C}^m_a(X)$ be the space of all m times continuously differentiable functions on X, all of whose derivatives up to order m are $\ll \mathcal{Y}^{-a}$ throughout X. In more precise terms, let $\mathfrak{g}$ be the Lie algebra of G, and fix a basis $X_1, \ldots, X_{2k+3}$ of $\mathfrak{g}$ (we make a definite choice of this basis; cf. (18)). Each $Y \in \mathfrak{g}$ can be realized as a left invariant differential operator on functions on G, and thus also a differential operator on $X = \Gamma\backslash G$, which we will also denote by Y. For any $f \in \mathrm{C}^m(X)$, set

$$\|f\|_{\mathrm{C}^m_a} := \sum_{\mathrm{ord}(D) \leqslant m} \sup_{p \in X} |\mathcal{Y}(p)^a (Df)(p)|, \tag{4}$$

where the sum is taken over all monomials in $X_1, \ldots, X_{3+2k}$ of degree $\leqslant m$. In particular, $\|\cdot\|_{\mathrm{C}^0_0}$ is the supremum norm. Then $\mathrm{C}^m_a(X)$ is the space of all $f \in \mathrm{C}^m(X)$ with $\|f\|_{\mathrm{C}^m_a} < \infty$.

For any integer $n \geqslant 0$ and real numbers $a \geqslant 0$ and $p \in [1, +\infty]$, we introduce the weighted Sobolev norm $S_{p,a,n}(h)$ on functions $h \in \mathrm{C}^n(\mathbb{R})$ through

$$S_{p,a,n}(h) = \sum_{j=0}^{n} \|(1+|x|)^a \, \partial^j h(x)\|_{\mathrm{L}^p}.$$

For $x \in \mathbb{R}$ let $\langle x \rangle$ denote the distance to the nearest integer; $\langle x \rangle = \min_{n \in \mathbb{Z}} |x - n|$. Given $\beta > k$, $\boldsymbol{\xi} \in \mathbb{R}^k$, $T > 0$, we define

$$\delta_{\beta,\boldsymbol{\xi}}(T) = \sum_{\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}} \|\boldsymbol{r}\|^{-\beta} \sum_{j=1}^{\infty} \frac{1 + \log^+\left(\frac{T\langle j\boldsymbol{r}\boldsymbol{\xi}\rangle}{j}\right)}{j^2 + Tj\langle j\boldsymbol{r}\boldsymbol{\xi}\rangle}, \tag{5}$$

where $\log^+(x) := \log(x)$ for $x > 1$ and $\log^+(x) := 0$ for $x \in [0,1]$. Since $\log^+(x) < x$ $(\forall x > 0)$, one has

$$\delta_{\beta,\boldsymbol{\xi}}(T) \leqslant C_{k,\beta} := \sum_{\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}} \|\boldsymbol{r}\|^{-\beta} \sum_{j=1}^{\infty} j^{-2} < \infty, \tag{6}$$

for all $\boldsymbol{\xi}$ and T.

We now state our two main theorems.

THEOREM 1.2 (Effective version of Theorem 1.1 when $\boldsymbol{\xi}_1 = \mathbf{0}$). *Let $k \geqslant 2$ and $\Gamma = \Gamma(N) \ltimes \mathbb{Z}^{2k}$. Fix $\varepsilon > 0$ and an integer $\beta \geqslant \max(8-k, 1+k)$, and set $m = 3(\beta+k+1)$ and $a = (\beta-1)/2$. Then for any $f \in \mathrm{C}_a^m(X)$, $h \in \mathrm{C}^2(\mathbb{R})$ with $S_{\infty,2+\varepsilon,2}(h) < \infty$, $\boldsymbol{\xi}_2 \in \mathbb{R}^k$ and $y > 0$, we have*

$$\begin{aligned} &\left| \int_{\mathbb{R}} f\left(\Gamma\left(1_2, \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\right) u(x) a(y)\right) h(x)\, dx - \int_X f\, d\mu \int_{\mathbb{R}} h\, dx \right| \\ &\quad \ll \|f\|_{\mathrm{C}_a^m} S_{\infty,2+\varepsilon,2}(h) \Big(\delta_{\beta,\boldsymbol{\xi}_2}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon} \Big), \end{aligned} \tag{7}$$

where the implied constant depends only on k, N, ε, β.

THEOREM 1.3 (Effective version of Theorem 1.1 when $\boldsymbol{\xi}_2 = \mathbf{0}$). *Let $k \geqslant 2$ and $\Gamma = \Gamma(N) \ltimes \mathbb{Z}^{2k}$. Fix $\varepsilon > 0$ and an integer $\beta \geqslant \max(7-k, 1+k)$, and set $m = 3(\beta+k)+2$ and $a = (\beta-1)/2$. Then for any $f \in \mathrm{C}_a^m(X)$, $h \in \mathrm{C}^2(\mathbb{R})$ with $S_{1,0,2}(h) < \infty$, $\boldsymbol{\xi}_1 \in \mathbb{R}^k$ and $y > 0$, we have*

$$\begin{aligned} &\left| \int_{\mathbb{R}} f\left(\Gamma\left(1_2, \begin{pmatrix}\boldsymbol{\xi}_1\\ \mathbf{0}\end{pmatrix}\right) u(x) a(y)\right) h(x)\, dx - \int_X f\, d\mu \int_{\mathbb{R}} h\, dx \right| \\ &\quad \ll \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h) \Big(\delta_{\beta,\boldsymbol{\xi}_1}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon} \Big), \end{aligned} \tag{8}$$

where the implied constant depends only on k, N, ε, β.

Let us make some comments on these results. First, note that for any fixed $\boldsymbol{\xi}_2 \in \mathbb{R}^k$ and $\beta > k$, one has $\delta_{\beta,\boldsymbol{\xi}_2}(y^{-1/2}) \to 0$ as $y \to 0$ if and only if $\boldsymbol{r}\boldsymbol{\xi}_2 \notin \mathbb{Q}$ for all $\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$. Hence, Theorem 1.2 indeed gives an effective version of Theorem 1.1 in the special case when Γ is a congruence subgroup of $\overline{\Gamma}$ and $\boldsymbol{\xi}_1 = \mathbf{0}$. Similarly, Theorem 1.3 gives an effective version of Theorem 1.1 when $\boldsymbol{\xi}_2 = \mathbf{0}$. Second, as we will explain in Section 3 (see especially Lemmata 3.1 and 3.3, and the relation (28)), for a sufficiently large β and $\boldsymbol{\xi}_2$ subject to a Diophantine condition, the majorant function $\delta_{\beta,\boldsymbol{\xi}_2}(T)$ has a power rate decay in T as $T \to \infty$. In particular, for any $\varepsilon > 0$, $\delta_{\beta,\boldsymbol{\xi}_2}(T) \ll T^{\varepsilon-1}$ holds for all $\boldsymbol{\xi}_2 \in \mathbb{R}^k$ outside a set of Hausdorff dimension $< k$. Note that for any such β and $\boldsymbol{\xi}_2$, the bound in Theorem 1.2 decays like $y^{\frac{1}{4}-\varepsilon}$ as $y \to 0$. An analogous statement holds for Theorem 1.3.

One should also note that the integral in Theorem 1.3 (but not the one in Theorem 1.2) runs over a *closed* orbit in X; indeed the point $\Gamma\big(1_2, \binom{\xi_1}{0}\big)u(x)$ is invariant under $x \mapsto x + N$, since $u(t)\binom{\xi_1}{0} = \binom{\xi_1}{0}$, $\forall t$, and $u(N) \in \Gamma(N)$. Hence, it is only natural that the bound obtained in Theorem 1.3 is invariant under translations of h.

We have made no effort to optimize the dependence on the test functions f and h in the theorems; rather, we have simply imposed as much smoothness and decay of these as needed to comfortably reach the best decay rate with respect to y that our method can give.

The proofs of Theorems 1.2 and 1.3 are given in Sections 4–8; the basic approach is to use Fourier decomposition with respect to the torus fiber variable, just as in [**38**]; however, there are several new difficulties that have to be tackled. In particular, the $\overline{\Gamma}'$-orbits in $\mathbb{Z}^{2k}$, which are used to partition the Fourier decomposition, are more complicated for $k \geqslant 2$ than for $k = 1$: There are two types of orbits, which we call 'A-orbits' and 'B-orbits', where B-orbits only appear for $k \geqslant 2$ (cf. Section 4). Establishing cancelation in the contribution from the B-orbits requires a novel treatment, which we give in Section 8. The treatment of the A-orbits (cf. Section 7) becomes more delicate for $k \geqslant 2$ than for $k = 1$, and this is where we need to require that the test function f decays sufficiently rapidly in the cusp (cf. the parameter 'a' in Theorems 1.2 and 1.3); this is not needed for $k = 1$. Other differences versus [**38**] are that we consider congruence subgroups and not just $\overline{\Gamma} = \mathrm{SL}(2, \mathbb{Z}) \ltimes (\mathbb{Z}^2)^{\oplus k}$ itself, and the fact that the Diophantine conditions are more complicated in the present paper, as they concern vectors in $\mathbb{R}^k$.

As will be seen, in the present paper we make crucial use of the assumptions in Theorems 1.2 and 1.3 that either $\boldsymbol{\xi}_1 = \mathbf{0}$ or $\boldsymbol{\xi}_2 = \mathbf{0}$. It is an interesting problem to seek a more general treatment so as to obtain an effective version of Theorem 1.1 for general $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$. We have some preliminary results on this problem and hope to return to it in a later paper.

We next turn to an application of Theorem 1.2: Following an approach introduced by Marklof in [**24**] using theta series, we will prove an effective quantitative Oppenheim-type result for the inhomogeneous quadratic form

$$Q(x_1, x_2, x_3, x_4) = (x_1 - \alpha)^2 + (x_2 - \beta)^2 - (x_3 - \alpha)^2 - (x_4 - \beta)^2 \tag{9}$$

for a fixed vector $(\alpha, \beta) \in \mathbb{R}^2$ subject to Diophantine conditions. Recall that the original Oppenheim conjecture states that for any indefinite nondegenerate homogeneous quadratic form $\widetilde{Q}$ in $n \geqslant 3$ variables, not proportional to a rational form, $\widetilde{Q}(\mathbb{Z}^n)$ is dense in $\mathbb{R}$. This was proved in celebrated work by Margulis [**21**]. An effective version of this result has more recently been obtained by Lindenstrauss and Margulis, [**20**]. A quantitative (but non-effective) version of the Oppenheim conjecture for forms of signature (p, q) with $p \geqslant 3$ and $q \geqslant 1$ was proved by Eskin, Margulis and Mozes, [**7**], and extended to forms of signature (2,2) subject to a Diophantine condition in [**8**]. Similar quantitative results were later proved also for inhomogeneous quadratic forms by Margulis and Mohammadi [**22**]; in particular the result proved by Marklof [**24**] for the form Q in (9) is a special case of the results in [**22**]; however, the method of proof in [**22**] is different and does not involve theta series.

Effective quantitative results for indefinite forms in $n \geqslant 5$ variables have been proved by Götze and Margulis [**11**]. However, we are not aware of any previous effective quantitative results for forms in 3 or 4 variables.

Returning to the form Q in (9), for $f \in \mathrm{C}_c(\mathbb{R}^4)$, $g \in \mathrm{C}(\mathbb{R}) \cap \mathrm{L}^1(\mathbb{R})$ and $T > 0$, set

$$N_{\alpha,\beta}(f, g, T) := \frac{1}{T^2} \sum_{\boldsymbol{m} \in \mathbb{Z}^4 \setminus \Delta} f(T^{-1}\boldsymbol{m}) g(Q(\boldsymbol{m})), \tag{10}$$

where $\Delta := \{(\boldsymbol{m}_1, \boldsymbol{m}_1) \,:\, \boldsymbol{m}_1 \in \mathbb{Z}^2\}$. We also set

$$\lambda_f := \frac{1}{2} \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} f(r\cos\zeta_1, r\sin\zeta_1, r\cos\zeta_2, r\sin\zeta_2)\, d\zeta_1\, d\zeta_2\, r\, dr. \tag{11}$$

One verifies easily that

$$\lim_{T\to\infty}\frac{1}{T^2}\int_{\mathbb{R}^4} f\big(T^{-1}\boldsymbol{x}\big)g(Q(\boldsymbol{x}))\,d\boldsymbol{x} = \lambda_f\int_{\mathbb{R}} g(r)\,dr.$$

We say that $\boldsymbol{\xi}\in\mathbb{R}^k$ is κ-Diophantine if there exists a constant $c>0$ such that $\|q\boldsymbol{\xi}-\boldsymbol{m}\|\geqslant cq^{-\kappa}$ for all $q\in\mathbb{Z}^+$ and $\boldsymbol{m}\in\mathbb{Z}^k$ (cf. [**23**, Section 1.5][†]). We also say that $\boldsymbol{\xi}$ is $[\kappa;c]$-Diophantine in this case. The smallest possible value for κ is $\kappa=k^{-1}$, and on the other hand Lebesgue-almost every $\boldsymbol{\xi}\in\mathbb{R}^k$ is $(k^{-1}+\varepsilon)$-Diophantine for every $\varepsilon>0$. In Section 3, we will also discuss a different (also standard) Diophantine condition, which is more directly connected to the decay properties of $\delta_{\beta,\boldsymbol{\xi}}(T)$.

In Section 9, we prove the following effective quantitative Oppenheim result for the form Q.

THEOREM 1.4. *There exists an absolute constant $B>0$ such that for any $[\kappa;c]$-Diophantine vector $(\alpha,\beta)\in\mathbb{R}^2$ with $|\alpha|,|\beta|\leqslant 1$, any $f\in \mathrm{C}_c^1(\mathbb{R}^4)$ with support contained in the unit ball centered at the origin, any $g\in \mathrm{C}^3(\mathbb{R})$ with $S_{1,2,3}(g)<\infty$, and any $T\geqslant 1$,*

$$\Big|N_{\alpha,\beta}(f,g,T)-\lambda_f\int_{\mathbb{R}} g(s)\,ds\Big| \ll \sum_{j=1}^{4}\Big\|\frac{\partial}{\partial x_j}f\Big\|_{\mathrm{L}^\infty} S_{1,2,3}(g)\,\kappa\,c^{-\frac{1}{\kappa}}\,\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)}, \tag{12}$$

where the implied constant is absolute.

The assumption in Theorem 1.4 that $\mathrm{supp}(f)$ is contained in the unit ball simplifies the statement of the theorem, but can easily be weakened by an *a posteriori* scaling argument; furthermore one can remove the assumption that $(\alpha,\beta)\in[-1,1]^2$, as long as T is large compared to $\|(\alpha,\beta)\|$ (cf. Corollary 9.12).

As we will show in Section 9.5, by a standard approximation argument, Theorem 1.4 implies the following effective counting result. For real numbers $a<b$ and $T>0$, set

$$N_{\alpha,\beta}(a,b,T) := \frac{1}{T^2}\#\big\{\boldsymbol{x}\in\mathbb{Z}^4\setminus\Delta\ :\ \|\boldsymbol{x}\|<T,\ a<Q(\boldsymbol{x})<b\big\}. \tag{13}$$

(One could also replace the ball $\{\|\boldsymbol{x}\|<T\}$ in (13) by a more general expanding region in $\mathbb{R}^4$; however, in order to keep the presentation simple we will not elaborate on this.)

COROLLARY 1.5. *There exists an absolute constant $B'>0$ such that for any $[\kappa;c]$-Diophantine vector $(\alpha,\beta)\in[-1,1]^2$ and any real numbers $a<b$ and $T\geqslant 1$,*

$$\Big|N_{\alpha,\beta}(a,b,T)-\tfrac{\pi^2}{2}(b-a)\Big| \ll (1+|a|+|b|)^3\kappa c^{-\frac{1}{\kappa}}\,\delta_{6,(\alpha,\beta)}(T)^{1/(B'\kappa)}, \tag{14}$$

where the implied constant is absolute.

Note that the right-hand sides of (12) and (14) tend to zero as $T\to\infty$ (keeping all other data fixed) whenever $1,\alpha,\beta$ are linearly independent over $\mathbb{Q}$ and the vector (α,β) is κ-Diophantine for some κ. If (α,β) furthermore satisfies a Diophantine condition of the type discussed in Section 3, then we even have a power rate decay with respect to T in (12) and (14). In particular, by a result of Schmidt [**34**] (or [**33**]), we have a power rate decay with respect to T whenever α,β are algebraic numbers such that $1,\alpha,\beta$ are linearly independent over $\mathbb{Q}$ (cf. Remark 4).

[†]Note that our κ corresponds to '$\kappa-1$' in [**23**, Section 1.5]. Both of these conventions are common in the literature, and we made our choice so as to make the statement of Theorem 1.4 and later results as simple as possible.

REMARK 1. The actual powers for the decay with respect to T which we obtain in Theorem 1.4 and Corollary 1.5 are quite small and depend strongly on the a and m appearing in the C_a^m-norm in Theorem 1.2 (which, as we remarked above, we have not attempted to optimize) (cf. Lemma 9.8 and Remark 12). It is an interesting problem to seek the *maximal* power η such that the difference in (12) decays like $T^{-\eta}$, for any fixed (α, β) subject to an appropriate Diophantine condition and any sufficiently nice test functions f and g.

REMARK 2. The relation $\lim_{T\to\infty} N_{\alpha,\beta}(a,b,T) = \frac{\pi^2}{2}(b-a)$ also holds for (α,β) κ-Diophantine with $1, \alpha, \beta$ linearly *de*pendent over $\mathbb{Q}$, except that for certain such pairs α, β, the definition of $N_{\alpha,\beta}(a,b,T)$ in (13) has to be modified by removing one more exceptional subspace besides Δ. This follows as a special case of the (ineffective) result of Margulis and Mohammadi [**22**, Theorem 1.9][†]. The reason why Theorem 1.4 and Corollary 1.5 fail to give the desired limiting result in the case when $1, \alpha, \beta$ are linearly dependent over $\mathbb{Q}$ is that as a crucial step in the proof, Theorem 1.2 is applied with $\boldsymbol{\xi}_2 = \binom{\alpha}{\beta}$, and as we discussed in connection with Theorem 1.1 (cf. (2)), the asymptotic equidistribution therein fails when $1, \alpha, \beta$ are $\mathbb{Q}$-linearly dependent. This situation is discussed in [**24**, Appendix A], and as indicated there, and carried out in some special cases, it is possible to extend the proof method of [**24**] to the case of $\mathbb{Q}$-linear dependence, by utilizing equidistribution in the appropriate homogeneous submanifold of $\Gamma\backslash G$. It would be interesting to make this approach effective, that is, to seek a satisfactory effective version of the statement that $\lim_{T\to\infty} N_{\alpha,\beta}(a,b,T) = \frac{\pi^2}{2}(b-a)$ for *all* κ-Diophantine vectors $(\alpha,\beta) \in \mathbb{R}^2$.

It should be noted that *some* Diophantine condition on (α,β) is certainly necessary in order for $\lim_{T\to\infty} N_{\alpha,\beta}(a,b,T) = \frac{\pi^2}{2}(b-a)$ to hold (cf. [**24**, Theorem 1.13, Section 9]. By contrast, the *non*-quantitative result that $Q(\mathbb{Z}^4)$ is dense in $\mathbb{R}$, and in fact even $\liminf_{T\to\infty} N_{\alpha,\beta}(a,b,T) \geqslant \frac{\pi^2}{2}(b-a)$ for all $a < b$, is known to hold for *all* irrational vectors (α,β), that is, for all $(\alpha,\beta) \in \mathbb{R}^2 \setminus \mathbb{Q}^2$. This is a special case of [**22**, Theorem 1.4].

Finally, let us note that Theorem 1.4 implies an effective version of the main theorem of [**24**], which says that under explicit Diophantine conditions on $(\alpha,\beta) \in \mathbb{R}^2$, the local two-point correlations of the sequence given by the values of $Q_1(m,n) = (m-\alpha)^2 + (n-\beta)^2$, with $(m,n) \in \mathbb{Z}^2$, are those of a Poisson process — a result which partly confirms a conjecture of Berry and Tabor [**1**] on quantized integrable systems. For fixed $(\alpha,\beta) \in \mathbb{R}^2$, denote by $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \to \infty$ the sequence of values of $Q_1(m,n)$ for $(m,n) \in \mathbb{Z}^2$, counted with multiplicity. One easily verifies that the asymptotic density of this sequence is π:

$$\#\{j \,:\, \lambda_j \leqslant \Lambda\} = \#\{(m,n) \in \mathbb{Z}^2 \,:\, (m-\alpha)^2 + (n-\beta)^2 < \Lambda\} \sim \pi\Lambda \quad \text{as} \quad \Lambda \to \infty.$$

For a given interval $[a,b] \subset \mathbb{R}$, the *pair correlation function* is then defined as

$$R_2[a,b](\Lambda) = \frac{1}{\pi\Lambda}\#\big\{(j,k) \in (\mathbb{Z}^+)^2 \,:\, j \neq k,\ \lambda_j, \lambda_k < \Lambda;\ \lambda_j - \lambda_k \in (a,b)\big\}. \tag{15}$$

In Section 9.5, we will prove the following.

COROLLARY 1.6. *There exists an absolute constant $B'' > 0$ such that for any $[\kappa; c]$-Diophantine vector $(\alpha,\beta) \in \mathbb{R}^2$, and any real numbers $a < b$ and $\Lambda \geqslant 1$,*

$$|R_2[a,b](\Lambda) - \pi(b-a)| \ll (1+|a|+|b|)^3 \kappa c^{-\frac{1}{\kappa}} \delta_{6,(\alpha,\beta)}(T)^{1/(B''\kappa)}, \tag{16}$$

where the implied constant is absolute.

[†]The notion of $\boldsymbol{\xi} \in \mathbb{R}^k$ being 'κ-Diophantine' in [**22**] is different from the one which we have defined; however, it is easy to verify that if $\boldsymbol{\xi}$ is κ-Diophantine in the sense of [**22**, Definition 1.7], then $\boldsymbol{\xi}$ is $(\kappa-1)$-Diophantine in our sense, and if $\boldsymbol{\xi}$ is κ-Diophantine in our sense, then $\boldsymbol{\xi}$ is $k(\kappa+1)$-Diophantine in the sense of [**22**, Definition 1.7]. One also verifies by a direct computation that the form Q in (9) with $(\alpha,\beta) \notin \mathbb{Q}^2$ admits at most one more exceptional subspace in the sense of [**22**, p. 124(bottom)] besides $\Delta = \{(\boldsymbol{m}_1, \boldsymbol{m}_1) \,:\, \boldsymbol{m}_1 \in \mathbb{Z}^2\}$, and such an exceptional subspace can only occur when $1, \alpha, \beta$ are linearly dependent over $\mathbb{Q}$.

This corollary indeed gives an effective version of Marklof [**24**, Theorem 1.8], as well as of [**23**, Theorem 1.6] in the case $k = 2$, since the right-hand side of (16) tends to zero as $T \to \infty$ for any fixed κ-Diophantine vector (α, β) (any κ) such that $1, \alpha, \beta$ are linearly independent over $\mathbb{Q}$.

The main result in Marklof [**23**, Theorem 1.6] generalizes [**24**, Theorem 1.8] to the case of the local pair correlation density of the sequence $\|\boldsymbol{m} - \boldsymbol{\alpha}\|^k$ ($\boldsymbol{m} \in \mathbb{Z}^k$) for any $k \geqslant 2$ (and also for $k = 2$ it is a stronger result, since the Diophantine condition imposed on the vector $(\alpha, \beta) \in \mathbb{R}^2$ is weaker). Unfortunately, it seems that Theorem 1.2 above cannot be used to prove an effective version of this more general result when $k \geqslant 3$. The reason is that the key equidistribution result required, [**23**, Theorem 5.1], concerns the integral

$$y^\sigma \int_{\mathbb{R}} f\Big(\Gamma\Big(1_2, \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\Big) u(x) a(y)\Big) h(y^\sigma x)\, dx \tag{17}$$

with $\sigma = \frac{k}{2} - 1$, that is, the integral which appears in Theorem 1.2 but with the function h replaced by $x \mapsto y^\sigma h(y^\sigma x)$. With this choice, the $S_{\infty,2+\varepsilon,2}$-norm in the right-hand side of (7) grows rapidly as $y \to 0$, making the bound useless. This failure may at first seem surprising, since the factor y^σ means, when $\sigma > 0$, that we are considering a unipotent orbit expanding at a *faster* rate than for $\sigma = 0$, so the result can be expected to be easier (or at least not more difficult) to prove. However, there is a genuine difference between x near zero and x far from zero in the integrand in (17); for example, for any $u(n) \in \Gamma$, using $u(n)\big(1_2, \binom{\mathbf{0}}{\boldsymbol{\xi}_2}\big) = \big(1_2, \binom{n\boldsymbol{\xi}_2}{\boldsymbol{\xi}_2}\big) u(n)$, we have

$$f\Big(\Gamma\Big(1_2, \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\Big) u(x) a(y)\Big) = f\Big(\Gamma\Big(1_2, \begin{pmatrix}n\boldsymbol{\xi}_2\\ \boldsymbol{\xi}_2\end{pmatrix}\Big) u(x+n) a(y)\Big).$$

It is clear from this that if one would solve the aforementioned problem of proving an effective version of Theorem 1.1 in the general case with both $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ allowed to be non-zero, this can be expected to also lead to an effective version of [**23**, Theorem 5.1], and so, with further work, should also lead to an effective version of [**23**, Theorem 1.6] for general $k \geqslant 2$.

2. *Some notation*

We use the standard notation $A = O(B)$ or $A \ll B$ meaning $|A| \leqslant CB$ for some constant $C > 0$. We shall also use $A \asymp B$ as a substitute for $A \ll B \ll A$. The implicit constant C will always be allowed to depend on k and N without any explicit mention. If we wish to indicate that C also depends on some other quantities f, g, h, we will use the notation $A \ll_{f,g,h} B$ or $A = O_{f,g,h}(B)$.

Recall from Section 1 that $G' = \mathrm{SL}(2, \mathbb{R})$ and $G = G' \ltimes \mathbb{R}^{2k}$. Let $\mathfrak{g}$ be the Lie algebra of G; it may be naturally identified with the space $\mathfrak{sl}(2, \mathbb{R}) \oplus \mathbb{R}^{2k}$, with Lie bracket $[(X, \boldsymbol{v}), (Y, \boldsymbol{w})] = (XY - YX, X\boldsymbol{w} - Y\boldsymbol{v})$ (see, for example, [**19**, Proposition 1.124]). Using this notation, we fix the following basis of $\mathfrak{g}$:

$$X_1 = \Big(\begin{pmatrix}0&1\\0&0\end{pmatrix}, \mathbf{0}\Big); \quad X_2 = \Big(\begin{pmatrix}0&0\\1&0\end{pmatrix}, \mathbf{0}\Big); \quad X_3 = \Big(\begin{pmatrix}1&0\\0&-1\end{pmatrix}, \mathbf{0}\Big); \tag{18}$$

$$X_{3+\ell} = \Big(0, \begin{pmatrix}\boldsymbol{e}_\ell\\ \mathbf{0}\end{pmatrix}\Big); \quad X_{3+k+\ell} = \Big(0, \begin{pmatrix}\mathbf{0}\\ \boldsymbol{e}_\ell\end{pmatrix}\Big) \quad (\ell = 1, \dots, k).$$

Here, $\boldsymbol{e}_1 = (1, 0, \dots, 0)$, $\boldsymbol{e}_2 = (0, 1, 0, \dots, 0)$, …, $\boldsymbol{e}_k = (0, \dots, 0, 1)$ are the standard basis vectors of $\mathbb{R}^k$.

We set

$$\overline{\Gamma}' = \mathrm{SL}(2,\mathbb{Z}) \quad \text{and} \quad \Gamma' = \Gamma(N), \quad \text{so that} \quad \overline{\Gamma} = \overline{\Gamma}' \ltimes \mathbb{Z}^{2k} \quad \text{and} \quad \Gamma = \Gamma' \ltimes \mathbb{Z}^{2k}$$

(cf. Section 1). Given a function f on $X = \Gamma\backslash G$, we will often view f as a function on G through $f(g) = f(\Gamma g)$, and we will write $f(M, \boldsymbol{v})$ in place of $f((M, \boldsymbol{v}))$, for $(M, \boldsymbol{v}) \in G$. Furthermore, given any $R \in \overline{\Gamma}'$, we set

$$f_R(M, \boldsymbol{v}) := f(R^{-1}(M, \boldsymbol{v})) = f(R^{-1}M, R^{-1}\boldsymbol{v}). \tag{19}$$

Since Γ' is normal in $\overline{\Gamma}'$, f_R is also left Γ-invariant, that is, f_R can be viewed as a function on X. Note also that $\|f_R\|_{C_a^m} = \|f\|_{C_a^m}$ for all $m \geqslant 0$, $a \in \mathbb{R}$.

3. *Linear form Diophantine conditions*

Given real numbers $\kappa \geqslant k$ and $\alpha \geqslant 1$, we say that a vector $\boldsymbol{\xi} \in \mathbb{R}^k$ is κ-LFD (short for κ-linear form Diophantine) if there is a constant $c > 0$ such that

$$\langle \boldsymbol{r}\boldsymbol{\xi} \rangle \geqslant c\|\boldsymbol{r}\|^{-\kappa} \quad \text{for all } \boldsymbol{r} \in \mathbb{Z}^k \setminus \{\boldsymbol{0}\}, \tag{20}$$

and we say that $\boldsymbol{\xi}$ is (κ, α)-LFD if there is a constant $c > 0$ such that

$$\langle j\boldsymbol{r}\boldsymbol{\xi} \rangle \geqslant cj^{-\alpha}\|\boldsymbol{r}\|^{-\kappa} \quad \text{for all } j \in \mathbb{Z}^+, \boldsymbol{r} \in \mathbb{Z}^k \setminus \{\boldsymbol{0}\}. \tag{21}$$

Recall here that for $x \in \mathbb{R}$, $\langle x \rangle$ denotes the distance to the nearest integer, and $\boldsymbol{r}\boldsymbol{\xi}$ is the scalar product, $\boldsymbol{r}\boldsymbol{\xi} = r_1\xi_1 + \cdots + r_k\xi_k$. The condition in (20) is very standard in the Diophantine approximation literature; however, we are not aware of any discussion of the more general condition in (21). When (20) holds, we will say that $\boldsymbol{\xi}$ is $[\kappa; c]$-LFD, and similarly when (21) holds, we will say that $\boldsymbol{\xi}$ is $[(\kappa, \alpha); c]$-LFD. Note that being $[\kappa; c]$-LFD is equivalent to being $[(\kappa, \alpha); c]$-LFD for any $\alpha \geqslant \kappa$. Hence, the notion of being $[(\kappa, \alpha); c]$-LFD is mainly relevant when $1 \leqslant \alpha < \kappa$, and in this case the condition (21) is equivalent to the same condition with $\boldsymbol{r}$ restricted to being a *primitive* vector in $\mathbb{Z}^k$ (namely, a vector with $\gcd(r_1, \ldots, r_k) = 1$).

Note that if $\boldsymbol{\xi}$ is (κ, α)-LFD, then $\boldsymbol{\xi}$ is also κ-LFD and furthermore each co-ordinate ξ_ℓ of $\boldsymbol{\xi}$ is an α-LFD ($\Leftrightarrow$ α-Diophantine) real number (apply (21) with $\boldsymbol{r} = \boldsymbol{e}_\ell$). Hence, if either $\kappa = k$ or $\alpha = 1$, then the set of (κ, α)-LFD $\boldsymbol{\xi} \in \mathbb{R}^k$ has Lebesgue measure zero [**16, 17, 29**]. On the other hand, if both $\kappa > k$ and $\alpha > 1$, then the *complement* of that set has Lebesgue measure zero, and moreover, it has Hausdorff dimension strictly less than k.

LEMMA 3.1. *If $\kappa > k$ and $\alpha > 1$, then the Hausdorff dimension of the set of all $\boldsymbol{\xi} \in \mathbb{R}^k$ which are not (κ, α)-LFD equals $k - 1 + \max\left(\frac{k+1}{\kappa+1}, \frac{2}{\alpha+1}\right)$.*

Proof. The set in the statement of the Lemma contains the set of all $\boldsymbol{\xi} \in \mathbb{R}^k$ which are not κ-LFD, and the latter set has (Hausdorff) dimension $k - 1 + \frac{k+1}{\kappa+1}$ (cf. Bovey and Dodson [**2**]). Furthermore, taking $\boldsymbol{r} = \boldsymbol{e}_1$ in (21) we see that the set in the statement of the lemma contains the set of all $\boldsymbol{\xi} \in \mathbb{R}^k$ for which ξ_1 is not α-LFD, and this set has dimension $k - 1 + \frac{2}{\alpha+1}$.

Hence, it remains to prove that the dimension in the statement of the lemma is bounded *above* by $k - 1 + \max\left(\frac{k+1}{\kappa+1}, \frac{2}{\alpha+1}\right)$. It suffices to consider $\boldsymbol{\xi} \in [0,1)^k$. Set

$$\Delta_{j,\boldsymbol{r},m} = \left\{\boldsymbol{\xi} \in [0,1)^k \ : \ |j\boldsymbol{r}\boldsymbol{\xi} - m| < j^{-\alpha}\|\boldsymbol{r}\|^{-\kappa}\right\}.$$

Then every non-(κ, α)-LFD $\boldsymbol{\xi}$ in $[0,1)^k$ belongs to $\Delta_{j,\boldsymbol{r},m}$ for infinitely many $(j, \boldsymbol{r}, m) \in \mathbb{Z}^+ \times (\mathbb{Z}^k \setminus \{\boldsymbol{0}\}) \times \mathbb{Z}$. Note also that $\Delta_{j,\boldsymbol{r},m} = \emptyset$ unless $|m| \ll j\|\boldsymbol{r}\|$, and for any $(j, \boldsymbol{r}, m) \in \mathbb{Z}^+ \times (\mathbb{Z}^k \setminus \{\boldsymbol{0}\}) \times \mathbb{Z}$, if we set $\ell = \ell_{j,\boldsymbol{r}} = j^{-\alpha-1}\|\boldsymbol{r}\|^{-\kappa-1}$, then the set $\Delta_{j,\boldsymbol{r},m}$ can be covered by $\ll \ell^{1-k}$ open hypercubes each having sides of length $\ll \ell$, with the normal to each face being parallel to a co-ordinate axis. If $s > k - 1 + \max\left(\frac{k+1}{\kappa+1}, \frac{2}{\alpha+1}\right)$, then the total s-volume of the family of

hypercubes obtained as $(j, \boldsymbol{r}, m)$ runs through $\mathbb{Z}^+ \times (\mathbb{Z}^k \setminus \{\mathbf{0}\}) \times \mathbb{Z}$ (subject to $\Delta_{j,\boldsymbol{r},m} \neq \emptyset$) is

$$\ll \sum_{j=1}^{\infty} \sum_{\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}} j\|\boldsymbol{r}\| \cdot \big(j^{-\alpha-1}\|\boldsymbol{r}\|^{-\kappa-1}\big)^{1-k+s} < \infty.$$

Note also that for any $\delta > 0$ there are only a finite number of non-empty sets $\Delta_{j,\boldsymbol{r},m}$ satisfying $\ell_{j,\boldsymbol{r}} \geqslant \delta$; hence every non-$(\kappa,\alpha)$-LFD $\boldsymbol{\xi} \in [0,1)^k$ is contained in the union of hypercubes in the above family restricted by $\ell_{j,\boldsymbol{r}} < \delta$. It follows that for every $s > k - 1 + \max\big(\frac{k+1}{\kappa+1}, \frac{2}{\alpha+1}\big)$, the s-dimensional outer Hausdorff measure of the set of all non-(κ,α)-LFD $\boldsymbol{\xi}$ in $[0,1)^k$ equals zero. This completes the proof. □

We will need the following auxiliary result.

LEMMA 3.2. *Let $\eta \in \mathbb{R}$, $c > 0$, $\kappa \geqslant 1$, and assume that $\langle j\eta\rangle \geqslant cj^{-\kappa}$ for all $j \in \mathbb{Z}^+$. Then*

$$\sum_{j=1}^{\infty} \frac{1}{j^2 + Tj\langle j\eta\rangle} \ll (cT)^{-\frac{2}{1+\kappa}} \log^2(2+T) \quad \text{for all } T > 0. \tag{22}$$

(The bound is essentially optimal. Indeed, if $\langle j\eta\rangle \leqslant cj^{-\kappa}$ holds for some j, then for $T = j^{1+\kappa}/c$, already the term $\frac{1}{j^2+Tj\langle j\eta\rangle}$ is bounded below by $\frac{1}{2}(cT)^{-\frac{2}{1+\kappa}}$.)

Proof. We assume $cT > 1$ since otherwise the bound is trivial. Note that the assumptions of the lemma imply that η is irrational, and $0 < c \leqslant \langle\eta\rangle \leqslant \frac{1}{2}$. Thus, $T > 2$.

Let p_k/q_k be the kth convergent of the (simple) continued fraction expansion of η (see, for example, [14, Chapter X]; in particular $1 = q_0 \leqslant q_1 < q_2 < \cdots$). For any $\ell \geqslant 1$ we have

$$\sum_{1 \leqslant j \leqslant q_\ell/2} \frac{1}{j\langle j\eta\rangle} = \sum_{k=1}^{\ell} \sum_{q_{k-1}/2 < j \leqslant q_k/2} \frac{1}{j\langle j\eta\rangle} \ll \sum_{k=1}^{\ell} q_{k-1}^{-1} \sum_{1 \leqslant j \leqslant q_k/2} \frac{1}{\langle j\eta\rangle} \ll \sum_{k=1}^{\ell} \frac{q_k \log q_k}{q_{k-1}},$$

where the last bound follows from [28, Lemma 4.8], since $|\eta - \frac{p_k}{q_k}| < \frac{1}{q_k q_{k+1}}$ [14, Theorem 171]. But for every $k \geqslant 1$ we have $cq_{k-1}^{-\kappa} \leqslant \langle q_{k-1}\eta\rangle < q_k^{-1}$, that is, $q_k < c^{-1}q_{k-1}^{\kappa}$; hence we get

$$\sum_{1 \leqslant j \leqslant q_\ell/2} \frac{1}{j\langle j\eta\rangle} \ll c^{-1}(\log q_\ell) \sum_{k=1}^{\ell} q_{k-1}^{\kappa-1} \ll c^{-1}(\log q_\ell)^2 q_{\ell-1}^{\kappa-1}, \tag{23}$$

where we used the fact that q_ℓ is bounded below by the ℓth Fibonacci number.

Next note that for any $\ell \geqslant 1$ and $h \geqslant 1$, by [28, Lemma 4.9],

$$\sum_{hq_\ell+1 \leqslant j \leqslant (h+1)q_\ell} \frac{1}{j^2 + Tj\langle j\eta\rangle} \leqslant \frac{1}{Thq_\ell} \sum_{r=1}^{q_\ell} \min\left(\frac{T}{hq_\ell}, \frac{1}{\langle (hq_\ell + r)\eta\rangle}\right) \ll \frac{1}{(hq_\ell)^2} + \frac{\log q_\ell}{Th}. \tag{24}$$

Similarly,

$$\sum_{q_\ell/2 < j \leqslant q_\ell} \frac{1}{j^2 + Tj\langle j\eta\rangle} \ll \frac{1}{Tq_\ell} \sum_{r=1}^{q_\ell} \min\left(\frac{T}{q_\ell}, \frac{1}{\langle r\eta\rangle}\right) \ll \frac{1}{q_\ell^2} + \frac{\log q_\ell}{T}. \tag{25}$$

Adding (25) and (24) for all $h \leqslant T/q_\ell$, we obtain

$$\sum_{q_\ell/2 < j \leqslant T} \frac{1}{j^2 + Tj\langle j\eta\rangle} \ll \frac{1}{q_\ell^2} + \frac{\log q_\ell \log(1 + \frac{T}{q_\ell})}{T}. \tag{26}$$

Now choose $\ell \geqslant 1$, so that $q_{\ell-1} \leqslant (cT)^{\frac{1}{1+\kappa}} < q_\ell$. Then $q_\ell < c^{-1} q_{\ell-1}^{\kappa} \leqslant (cT)^{-\frac{1}{\kappa+1}} T < T$. Now (22) follows from (23), (26) and the bound $\sum_{j>T} j^{-2} \ll T^{-1}$. □

We now give a result on the rate of decay of the majorant function $\delta_{\beta,\boldsymbol{\xi}}(T)$ (cf. (5)), assuming that $\boldsymbol{\xi}$ is of an appropriate LFD type. In fact, we consider the following slightly simpler majorant:

$$\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T) = \sum_{\boldsymbol{r}\in\mathbb{Z}^k\setminus\{\mathbf{0}\}} \|\boldsymbol{r}\|^{-\beta} \sum_{j=1}^{\infty} \frac{1}{j^2 + Tj\langle j\boldsymbol{r}\boldsymbol{\xi}\rangle}. \tag{27}$$

Note that $\delta_{\beta,\boldsymbol{\xi}}(T)$ and $\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T)$ decay with very similar rates, since

$$\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T) \leqslant \delta_{\beta,\boldsymbol{\xi}}(T) \leqslant (2\log T)\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T), \quad \forall T \geqslant e. \tag{28}$$

LEMMA 3.3. *For any $\kappa \geqslant k$, $\alpha \geqslant 1$ and $\beta > k + \frac{2\kappa}{1+\alpha}$, if $\boldsymbol{\xi} \in \mathbb{R}^k$ is $[(\kappa,\alpha);c]$-LFD, then*

$$\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T) \ll_{\beta,\kappa,\alpha} (cT)^{-\frac{2}{1+\alpha}} \log^2(2+T) \quad \textit{for all } T > 0.$$

Proof. Using Lemma 3.2 and the assumption that $\boldsymbol{\xi}$ is $[(\kappa,\alpha);c]$-LFD, we have

$$\sum_{j=1}^{\infty} \frac{1}{j^2 + Tj\langle j\boldsymbol{r}\boldsymbol{\xi}\rangle} \ll (c\|\boldsymbol{r}\|^{-\kappa}T)^{-\frac{2}{1+\alpha}} \log^2(2+T), \quad \text{for each } \boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}.$$

Multiplying by $\|\boldsymbol{r}\|^{-\beta}$ and adding over all $\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, we obtain the stated bound. □

REMARK 3. A standard argument also shows that given any $\beta > k$ and $\varepsilon > 0$, the bound $\delta_{\beta,\boldsymbol{\xi}}(T) \ll T^{\varepsilon-1}$ as $T \to \infty$ holds for *Lebesgue almost all* $\boldsymbol{\xi} \in \mathbb{R}^k$. We here give an outline of the proof: One verifies that for T large, $\int_{[0,1]^k} \widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T)\, d\boldsymbol{\xi} \ll T^{\frac{1}{3}\varepsilon-1}$, and hence the set of $\boldsymbol{\xi} \in [0,1]^k$ satisfying $\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T) \geqslant T^{\frac{2}{3}\varepsilon-1}$ has Lebesgue measure $\ll T^{-\frac{1}{3}\varepsilon}$. The sum of these measures over $T = 2^1, 2^2, 2^3, \ldots$ is finite, and so, by Borel–Cantelli, for almost every $\boldsymbol{\xi} \in [0,1]^k$ there is some $M \in \mathbb{Z}^+$ such that $\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(2^m) < (2^m)^{\frac{2}{3}\varepsilon-1}$ for all integers $m \geqslant M$, and thus $\widetilde{\delta}_{\beta,\boldsymbol{\xi}}(T) < 2T^{\frac{2}{3}\varepsilon-1}$ for all (real) $T \geqslant 2^M$. The desired claim then follows using (28) and the fact that $\delta_{\beta,\boldsymbol{\xi}}(T)$ is invariant under $\boldsymbol{\xi} \mapsto \boldsymbol{\xi} + \boldsymbol{m}$, $\boldsymbol{m} \in \mathbb{Z}^k$.

REMARK 4. By Schmidt, [34], if $\xi_1, \ldots, \xi_k$ are (real) *algebraic* numbers such that $1, \xi_1, \ldots, \xi_k$ are linearly independent over $\mathbb{Q}$, then $\boldsymbol{\xi}$ is κ-LFD (and thus $[\kappa,\kappa]$-LFD) for every $\kappa > k$. Hence, for such a $\boldsymbol{\xi}$, Lemma 3.3 implies that for any $\beta > k\big(1 + \frac{2}{1+k}\big)$ and $\varepsilon > 0$ we have $\delta_{\beta,\boldsymbol{\xi}}(T) \ll_{\boldsymbol{\xi},\beta,\varepsilon} T^{\varepsilon - \frac{2}{k+1}}$ for all $T > 0$. In connection with Theorem 1.4, it should be noted that any such $\boldsymbol{\xi}$ is also κ-Diophantine for every $\kappa > k^{-1}$; again (cf. [34]).

4. *Fourier decomposition with respect to the torus variable*

We now start with the proof of Theorems 1.2 and 1.3. In this section, which generalizes [38, Section 4], we consider the Fourier decomposition of a given test function on X with respect to the torus variable, and prove bounds on the resulting Fourier coefficients. Some parts of our discussion is a close mimic of [38, Section 4], but there are also some new aspects that have to be considered; see, in particular, all of Section 4.2.

To start with, we consider an arbitrary function $f \in \mathrm{C}(\mathbb{Z}^{2k}\backslash G)$, where $\mathbb{Z}^{2k}$ is viewed as a subgroup of G through $\boldsymbol{n} \mapsto (1_2, \boldsymbol{n})$. We view f as a function on G by composing with the

projection $G \mapsto \mathbb{Z}^{2k}\backslash G$. Then $f(M,\boldsymbol{\xi}) = f((1_2,\boldsymbol{n})(M,\boldsymbol{\xi})) = f(M,\boldsymbol{\xi}+\boldsymbol{n})$ for all $\boldsymbol{n} \in \mathbb{Z}^{2k}$, which means that for any fixed $M \in G'$, $\boldsymbol{\xi} \mapsto f(M,\boldsymbol{\xi})$ is a function on the torus $\mathbb{T}^{2k} = \mathbb{Z}^{2k}\backslash\mathbb{R}^{2k}$. We write $\widehat{f}(M,\boldsymbol{m})$ for the Fourier coefficients in the torus variable:

$$\widehat{f}(M,\boldsymbol{m}) = \int_{\mathbb{Z}^{2k}\backslash\mathbb{R}^{2k}} f(M,\boldsymbol{\xi}) e(-\boldsymbol{m}\boldsymbol{\xi})\, d\boldsymbol{\xi}, \quad M \in G',\ \boldsymbol{m} \in \mathbb{Z}^{2k}. \tag{29}$$

Here, $d\boldsymbol{\xi}$ denotes Lebesgue measure on $\mathbb{R}^{2k}$. Thus, for $f \in \mathrm{C}^{k+1}(\mathbb{Z}^{2k}\backslash G)$ we have [**12**, Theorem 3.2.16]

$$f(M,\boldsymbol{\xi}) = \sum_{\boldsymbol{m}\in\mathbb{Z}^{2k}} \widehat{f}(M,\boldsymbol{m}) e(\boldsymbol{m}\boldsymbol{\xi}), \tag{30}$$

with a uniform absolute convergence[†] over $(M,\boldsymbol{\xi})$ in any compact subset of G. (Indeed, the function $\boldsymbol{\xi} \mapsto f(M,\boldsymbol{\xi})$ is in $\mathrm{C}^{k+1}(\mathbb{T}^{2k})$, with $\|f(M,\cdot)\|_{\mathrm{C}^{k+1}(\mathbb{T}^{2k})}$ depending continuously on $M \in G'$.)

If f is also invariant under some $T \in \overline{\Gamma}' = \mathrm{SL}(2,\mathbb{Z})$, this leads to a corresponding invariance relation for $\widehat{f}(M,\boldsymbol{m})$:

LEMMA 4.1. *For any $T \in \overline{\Gamma}'$, if $f \in \mathrm{C}(\mathbb{Z}^{2k}\backslash G)$ is left T-invariant, then*

$$\widehat{f}(TM,\boldsymbol{m}) = \widehat{f}(M,\,{}^{\mathrm{t}}T\boldsymbol{m}), \quad \forall M \in G',\ \boldsymbol{m} \in \mathbb{Z}^{2k}, \tag{31}$$

where ${}^{\mathrm{t}}T$ is the transpose of T.

Proof. We have

$$\begin{aligned}\widehat{f}(TM,\boldsymbol{m}) &= \int_{\mathbb{T}^{2k}} f(TM,\boldsymbol{\xi}) e(-\boldsymbol{m}\boldsymbol{\xi})\, d\boldsymbol{\xi} = \int_{\mathbb{T}^{2k}} f(TM,T\boldsymbol{\xi}) e(-\boldsymbol{m}(T\boldsymbol{\xi}))\, d\boldsymbol{\xi} \\ &= \int_{\mathbb{T}^{2k}} f(T(M,\boldsymbol{\xi})) e(-\boldsymbol{m}(T\boldsymbol{\xi}))\, d\boldsymbol{\xi} = \int_{\mathbb{T}^{2k}} f(M,\boldsymbol{\xi}) e(-\boldsymbol{m}(T\boldsymbol{\xi}))\, d\boldsymbol{\xi},\end{aligned}$$

where in the second equality we used the fact that $\boldsymbol{\xi} \mapsto T\boldsymbol{\xi}$ is a diffeomorphism of $\mathbb{T}^{2k}$ preserving $d\boldsymbol{\xi}$, and in the last equality we used the fact that f is left T-invariant. Using $\boldsymbol{m}(T\boldsymbol{\xi}) = ({}^{\mathrm{t}}T\boldsymbol{m})\boldsymbol{\xi}$ we obtain (31). □

Because of Lemma 4.1, if $f \in \mathrm{C}^{k+1}(\overline{\Gamma}\backslash G)$, then it is convenient to group the terms in (30) together according to the orbits for the action of $\overline{\Gamma}'$ on $\mathbb{Z}^{2k}$. We call an orbit for this action an *A-orbit* if it contains some element of the form $\binom{\boldsymbol{0}}{\boldsymbol{r}}$, where $\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\boldsymbol{0}\}$. Every other non-zero orbit is called a *B-orbit*.

LEMMA 4.2. *Every B-orbit contains an element $\boldsymbol{\eta} = \binom{\boldsymbol{q}}{\boldsymbol{r}}$ ($\boldsymbol{q} = {}^{\mathrm{t}}(q_1,\dots,q_k)$, $\boldsymbol{r} = {}^{\mathrm{t}}(r_1,\dots,r_k)$) with the property that there are some $1 \leqslant \ell_1 < \ell_2 \leqslant k$ such that $r_j = 0$ for all $j < \ell_1$, $q_j = 0$ for all $j < \ell_2$, and $r_{\ell_1} > 0$, $0 \leqslant r_{\ell_2} < |q_{\ell_2}|$.*

Proof. Let $\boldsymbol{\eta} = \binom{\boldsymbol{q}}{\boldsymbol{r}}$ be an element in a B-orbit. Then $\boldsymbol{\eta} \neq \boldsymbol{0}$, and we may take ℓ_1 to be the smallest index for which $\binom{q_{\ell_1}}{r_{\ell_1}} \neq \binom{0}{0}$. After replacing $\boldsymbol{\eta}$ by $T\boldsymbol{\eta}$ for an appropriate $T \in \overline{\Gamma}'$ we can ensure that $q_{\ell_1} = 0$ and $r_{\ell_1} > 0$, while clearly still $q_j = r_j = 0$ for all $j < \ell_1$. Now since $\boldsymbol{\eta}$ is not in an A-orbit, we cannot have $q_j = 0$ for all j, and we take $\ell_2 > \ell_1$ to be the smallest index for which $q_{\ell_2} \neq 0$. Finally, by replacing $\boldsymbol{\eta}$ by $\left(\begin{smallmatrix}1 & 0\\ x & 1\end{smallmatrix}\right)\boldsymbol{\eta}$ for an appropriate $x \in \mathbb{Z}$ we can make $0 \leqslant r_{\ell_2} < |q_{\ell_2}|$ hold, while q_j and r_j for $j < \ell_2$ remain unchanged. □

[†]For any fixed ordering of $\mathbb{Z}^{2k}$.

Let us fix, once and for all, a set of representatives $A_k, B_k \subset \mathbb{Z}^{2k}$ such that A_k contains exactly one element from each A-orbit and B_k contains exactly one element from each B-orbit, and furthermore each $\boldsymbol{\eta} \in A_k$ is of the form $\boldsymbol{\eta} = \binom{0}{r}$ and each $\boldsymbol{\eta} \in B_k$ has the property described in Lemma 4.2.

LEMMA 4.3. *The stabilizer in $\overline{\Gamma}'$ of any $\boldsymbol{\eta} \in A_k$ equals $\left\{\left(\begin{smallmatrix}1 & 0\\ n & 1\end{smallmatrix}\right) : n \in \mathbb{Z}\right\}$. The stabilizer in $\overline{\Gamma}'$ of any $\boldsymbol{\eta} \in B_k$ is trivial.*

Proof. Immediate verification. □

The lemma implies that we can decompose $\mathbb{Z}^{2k}$ as a disjoint union of singleton sets as follows:

$$\mathbb{Z}^{2k} = \{\mathbf{0}\} \bigsqcup \left(\bigsqcup_{\boldsymbol{\eta}\in A_k} \bigsqcup_{T\in \overline{\Gamma}'_\infty\backslash\overline{\Gamma}'} \{{}^tT\boldsymbol{\eta}\} \right) \bigsqcup \left(\bigsqcup_{\boldsymbol{\eta}\in B_k} \bigsqcup_{T\in\overline{\Gamma}'} \{{}^tT\boldsymbol{\eta}\} \right), \tag{32}$$

where $\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'$ denotes any set of representatives for the right cosets inside $\overline{\Gamma}'$ of the subgroup

$$\overline{\Gamma}'_\infty := \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}. \tag{33}$$

Grouping together the terms in (30) according to (32), and then applying Lemma 4.1, we get, for any $f \in \mathrm{C}^{k+1}(\overline{\Gamma}\backslash G)$:

$$f(M,\boldsymbol{\xi}) = \widehat{f}(M,\mathbf{0}) + \sum_{\boldsymbol{\eta}\in A_k} \sum_{T\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'} \widehat{f}(TM,\boldsymbol{\eta}) e((\,{}^tT\boldsymbol{\eta})\boldsymbol{\xi}) + \sum_{\boldsymbol{\eta}\in B_k} \sum_{T\in\overline{\Gamma}'} \widehat{f}(TM,\boldsymbol{\eta}) e((\,{}^tT\boldsymbol{\eta})\boldsymbol{\xi}). \tag{34}$$

If $k = 1$, then $B_k = \emptyset$ and (34) can be seen to agree with [**38**, Lemma 4.1]. However, B_k is easily seen to be nonempty for every $k \geqslant 2$.

We now wish to give a similar decomposition of a general function $f \in \mathrm{C}^{k+1}(X)$. Recall that $X = \Gamma\backslash G$ and $\Gamma = \Gamma' \ltimes \mathbb{Z}^{2k}$ with $\Gamma' = \Gamma(N)$, a normal subgroup of $\overline{\Gamma}' = \mathrm{SL}(2,\mathbb{Z})$. For any subgroup H of G' and any subset $A \subset G'$ satisfying $HA = A$, we denote by $H\backslash A$ a set of representatives for the distinct cosets Ha $(a \in A)$. We also write $\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'$ for a set of representatives for the double cosets of the form $\overline{\Gamma}'_\infty R\Gamma'$ with $R \in \overline{\Gamma}'$. Let

$$\Gamma'_\infty := \Gamma' \cap \overline{\Gamma}'_\infty = \left\{ \begin{pmatrix} 1 & Nn \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}.$$

One then verifies that $\bigsqcup_{R\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'} \bigsqcup_{T\in\Gamma'_\infty\backslash\Gamma' R}\{T\}$ is a set of representatives for $\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'$. Hence, from (32) we get

$$\mathbb{Z}^{2k} = \{\mathbf{0}\} \bigsqcup \left(\bigsqcup_{\boldsymbol{\eta}\in A_k} \bigsqcup_{R\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'} \bigsqcup_{T\in\Gamma'_\infty\backslash\Gamma' R} \{{}^tT\boldsymbol{\eta}\} \right) \bigsqcup \left(\bigsqcup_{\boldsymbol{\eta}\in B_k} \bigsqcup_{R\in\overline{\Gamma}'/\Gamma'} \bigsqcup_{T\in\Gamma' R} \{{}^tT\boldsymbol{\eta}\} \right). \tag{35}$$

Using $\Gamma' R = R\Gamma'$ and ${}^t(R\gamma)\boldsymbol{\eta} = {}^t\gamma({}^tR\boldsymbol{\eta})$ for $\gamma \in \Gamma'$, this formula is seen to provide a decomposition of $\mathbb{Z}^{2k}$ into orbits for the action of ${}^t\Gamma' = \Gamma'$. In order to get a convenient corresponding partition of the sum in (30), recall (19), and note that for any $R \in \overline{\Gamma}'$, $M \in G'$, $\boldsymbol{m} \in \mathbb{Z}^{2k}$ we have

$$\widehat{f_R}(M,\boldsymbol{m}) = \widehat{f}(R^{-1}M,\, {}^tR\boldsymbol{m}).$$

This is proved by a computation similar to the proof of Lemma 4.1. Using Lemma 4.1, we get $\widehat{f}(M, {}^t\gamma\,{}^tR\boldsymbol{\eta}) = \widehat{f_R}(R\gamma M, \boldsymbol{\eta})$ for all $\gamma \in \Gamma'$, or in other words:

$$\widehat{f}(M, {}^tT\boldsymbol{\eta}) = \widehat{f_R}(TM, \boldsymbol{\eta}), \quad \forall R \in \overline{\Gamma}', \ T \in \Gamma' R, \ M \in G', \ \boldsymbol{\eta} \in \mathbb{Z}^{2k}. \tag{36}$$

Now from (30), (35) and (36) we get

$$\begin{aligned} f(M, \boldsymbol{\xi}) = \widehat{f}(M, \mathbf{0}) &+ \sum_{\boldsymbol{\eta} \in A_k} \sum_{R \in \overline{\Gamma}'_\infty \backslash \overline{\Gamma}' / \Gamma'} \sum_{T \in \Gamma'_\infty \backslash \Gamma' R} \widehat{f_R}(TM, \boldsymbol{\eta}) e(({}^tT\boldsymbol{\eta})\boldsymbol{\xi}) \\ &+ \sum_{\boldsymbol{\eta} \in B_k} \sum_{R \in \overline{\Gamma}' / \Gamma'} \sum_{T \in \Gamma' R} \widehat{f_R}(TM, \boldsymbol{\eta}) e(({}^tT\boldsymbol{\eta})\boldsymbol{\xi}). \end{aligned} \tag{37}$$

Note here that for any $\boldsymbol{\eta} \in A_k$ and $R \in \overline{\Gamma}'$, the function $M \mapsto \widehat{f_R}(M, \boldsymbol{\eta})$ is left Γ'_∞-invariant, by (36) and Lemma 4.3. However, for $\boldsymbol{\eta} \in B_k$ there is no such invariance present.

4.1. *Bounds when $\boldsymbol{\eta} \in A_k$*

We now give bounds on $\widehat{f}(T, \boldsymbol{\eta})$ for $\boldsymbol{\eta} \in A_k$.

LEMMA 4.4. *For any $m \geqslant 0$, $\alpha \in \mathbb{R}_{\geqslant 0}$, $\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ and $f \in \mathrm{C}^m_\alpha(X)$, we have*

$$\left| \widehat{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix} \right) \right| \ll_{m,\alpha} \|f\|_{\mathrm{C}^m_\alpha} \|\boldsymbol{r}\|^{-m} (c^2 + d^2)^{-\frac{m}{2}} \min\left(1, (c^2 + d^2)^{\alpha}\right). \tag{38}$$

Proof. The left invariant differential operator corresponding to $Y \in \mathfrak{g}$ is given by $Yf(g) = \lim_{t \to 0}(f(g \exp(tY)) - f(g))/t$. In particular, if we parameterize G as $\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), \left(\begin{smallmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{smallmatrix}\right)\right)$, where $\boldsymbol{y} = {}^t(y_1, \dots, y_k)$ and $\boldsymbol{z} = {}^t(z_1, \dots, z_k)$, then (cf. (18))

$$X_{3+\ell} = a\frac{\partial}{\partial y_\ell} + c\frac{\partial}{\partial z_\ell} \quad \text{and} \quad X_{3+k+\ell} = b\frac{\partial}{\partial y_\ell} + d\frac{\partial}{\partial z_\ell}, \quad \ell \in \{1, \dots, k\}. \tag{39}$$

Now

$$\widehat{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix} \right) = \int_{\mathbb{T}^k} \int_{\mathbb{T}^k} f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} \right) e(-\boldsymbol{r}\boldsymbol{z}) \, d\boldsymbol{y} \, d\boldsymbol{z},$$

and hence by repeated integration by parts we have

$$(2\pi i r_\ell c)^m \cdot \widehat{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix} \right) = \int_{\mathbb{T}^k} \int_{\mathbb{T}^k} [X^m_{3+\ell} f]\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} \right) e(-\boldsymbol{r}\boldsymbol{z}) \, d\boldsymbol{y} \, d\boldsymbol{z}$$

and

$$(2\pi i r_\ell d)^m \cdot \widehat{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix} \right) = \int_{\mathbb{T}^k} \int_{\mathbb{T}^k} [X^m_{3+k+\ell} f]\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} \right) e(-\boldsymbol{r}\boldsymbol{z}) \, d\boldsymbol{y} \, d\boldsymbol{z}.$$

Hence,

$$|r_\ell|^m \max(|c|^m, |d|^m) \cdot \left| \widehat{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix} \right) \right| \leqslant (2\pi)^{-m} \|f\|_{\mathrm{C}^m_\alpha} \mathcal{Y}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right)^{-\alpha},$$

for each $\ell \in \{1, \dots, k\}$. Using $\mathcal{Y}(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)) \geqslant \max(\sqrt{3}/2, (c^2 + d^2)^{-1})$, we get (38). □

Using Lemma 4.4 we immediately obtain bounds on *derivatives* of $\widehat{f}(\cdot,\cdot)$ with respect to the first variable. We express these in terms of Iwasawa co-ordinates, that is we write (by a slight abuse of notation)

$$\widehat{f}(u,v,\theta;\boldsymbol{\eta}) := \widehat{f}\left(\begin{pmatrix}1 & u\\ 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{v} & 0\\ 0 & 1/\sqrt{v}\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\boldsymbol{\eta}\right), \tag{40}$$

for $u\in\mathbb{R}$, $v>0$, $\theta\in\mathbb{R}/2\pi\mathbb{Z}$, $\boldsymbol{\eta}\in\mathbb{Z}^{2k}$.

LEMMA 4.5. *For any $\alpha\in\mathbb{R}_{\geqslant 0}$, $\boldsymbol{r}\in\mathbb{Z}^k\setminus\{\mathbf{0}\}$, integers $m,\ell_1,\ell_2,\ell_3\geqslant 0$ and $f\in \mathrm{C}_\alpha^{m+\ell}(X)$, where $\ell=\ell_1+\ell_2+\ell_3$, we have*

$$\left|\left(\frac{\partial}{\partial u}\right)^{\ell_1}\left(\frac{\partial}{\partial v}\right)^{\ell_2}\left(\frac{\partial}{\partial\theta}\right)^{\ell_3}\widehat{f}\left(u,v,\theta;\begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)\right| \ll_{m,\ell,\alpha} \|f\|_{\mathrm{C}_\alpha^{m+\ell}}\|\boldsymbol{r}\|^{-m}v^{\frac{m}{2}-\ell_1-\ell_2}\min(1,v^{-\alpha}). \tag{41}$$

Proof. This is just as in [38, Lemmas 4.3 and 4.4]. □

4.2. *Bounds when $\boldsymbol{\eta}\in B_k$*

We now give bounds on $\widehat{f}(T,\boldsymbol{\eta})$ when $\boldsymbol{\eta}\in B_k$. We will use the Frobenius matrix norm,

$$\|T\| := \operatorname{tr}\left(T\,{}^tT\right) = \sqrt{a^2+b^2+c^2+d^2},\quad T=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in G'.$$

In Iwasawa co-ordinates, we have, for any $u\in\mathbb{R}$, $v>0$, $\theta\in\mathbb{R}/2\pi\mathbb{Z}$,

$$\left\|\begin{pmatrix}1 & u\\ 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{v} & 0\\ 0 & 1/\sqrt{v}\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\right\| = \sqrt{\frac{u^2+v^2+1}{v}}. \tag{42}$$

This is verified by a quick computation, where the first step is to use the fact that the matrix $U=\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ lies in SO(2), that is, $U\,{}^tU=1_2$.

We have the following analogue of Lemma 4.4.

LEMMA 4.6. *For any $\boldsymbol{\eta}\in B_k$, $m\geqslant 0$, $f\in \mathrm{C}_0^m(X)$ and $T\in G'$,*

$$\left|\widehat{f}(T,\boldsymbol{\eta})\right| \ll_m \frac{\|f\|_{\mathrm{C}_0^m}}{(\|T\|+\|\boldsymbol{\eta}\|/\|T\|)^m}. \tag{43}$$

REMARK 5. As a consequence, for any $0<\beta<\frac{1}{2}$ we have

$$\left|\widehat{f}(T,\boldsymbol{\eta})\right| \ll_m \frac{\|f\|_{\mathrm{C}_0^m}}{\|T\|^{m(1-2\beta)}\|\boldsymbol{\eta}\|^{m\beta}}.$$

Proof. We write $\boldsymbol{\eta}=\left(\begin{smallmatrix}\boldsymbol{q}\\ \boldsymbol{r}\end{smallmatrix}\right)$ and $T=\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$. Repeated integration by parts gives (cf. the proof of Lemma 4.4)

$$(2\pi i(q_\ell a+r_\ell c))^m\cdot\widehat{f}(T,\boldsymbol{\eta}) = \int_{\mathbb{T}^k}\int_{\mathbb{T}^k}[X_{3+\ell}^m f]\left(T,\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{pmatrix}\right)e\left(-\boldsymbol{\eta}\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{pmatrix}\right)d\boldsymbol{y}\,d\boldsymbol{z}$$

and

$$(2\pi i(q_\ell b+r_\ell d))^m\cdot\widehat{f}(T,\boldsymbol{\eta}) = \int_{\mathbb{T}^k}\int_{\mathbb{T}^k}[X_{3+k+\ell}^m f]\left(T,\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{pmatrix}\right)e\left(-\boldsymbol{\eta}\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{z}\end{pmatrix}\right)d\boldsymbol{y}\,d\boldsymbol{z}.$$

Hence, if we write $\boldsymbol{\eta}^{(\ell)} := \binom{q_\ell}{r_\ell} \in \mathbb{R}^2$, then we conclude that for each $\ell \in \{1, \ldots, k\}$ and for each column vector $\boldsymbol{v}$ of T, we have

$$\left|\widehat{f}(T, \boldsymbol{\eta})\right| \leqslant \frac{\|f\|_{\mathrm{C}_0^m}}{(2\pi)^m |\boldsymbol{\eta}^{(\ell)}\boldsymbol{v}|^m}. \tag{44}$$

Now fix a column vector $\boldsymbol{v}$ of T with the largest norm. Then $\|T\| \leqslant \sqrt{2}\|\boldsymbol{v}\|$. By our definition of B_k, $\boldsymbol{\eta}$ has the property described in Lemma 4.2, that is, there are $1 \leqslant \ell_1 < \ell_2 \leqslant k$ such that $r_j = 0$ for all $j < \ell_1$, $q_j = 0$ for all $j < \ell_2$, and $r_{\ell_1} > 0$, $0 \leqslant r_{\ell_2} < |q_{\ell_2}|$. In particular, the vectors $\boldsymbol{\eta}^{(\ell_1)}$ and $\boldsymbol{\eta}^{(\ell_2)}$ are non-zero, hence both have length $\geqslant 1$, and the angle between the lines $\mathbb{R}\boldsymbol{\eta}^{(\ell_1)}$ and $\mathbb{R}\boldsymbol{\eta}^{(\ell_2)}$ in $\mathbb{R}^2$ is $> \frac{\pi}{4}$. Hence, the normal line to $\boldsymbol{v}$ in $\mathbb{R}^2$ has an angle $\geqslant \frac{\pi}{8}$ to at least one of the lines $\mathbb{R}\boldsymbol{\eta}^{(\ell_1)}$ and $\mathbb{R}\boldsymbol{\eta}^{(\ell_2)}$, and it follows that at least one of the scalar products $\boldsymbol{\eta}^{(\ell_1)}\boldsymbol{v}$ and $\boldsymbol{\eta}^{(\ell_2)}\boldsymbol{v}$ has an absolute value $\geqslant \sin(\frac{\pi}{8})\|\boldsymbol{v}\|$. Hence, using (44) we get

$$\left|\widehat{f}(T, \boldsymbol{\eta})\right| \ll_m \frac{\|f\|_{\mathrm{C}_0^m}}{\|T\|^m}. \tag{45}$$

Next let $\boldsymbol{v}'$ be the *other* column vector of T, and let $\alpha \in (0, \frac{\pi}{2}]$ be the angle between the lines $\mathbb{R}\boldsymbol{v}$ and $\mathbb{R}\boldsymbol{v}'$; then $\|\boldsymbol{v}\|\|\boldsymbol{v}'\| \sin\alpha = 1$, since $\det T = 1$. Let $\ell \in \{1, \ldots, k\}$ be the index for which $\|\boldsymbol{\eta}^{(\ell)}\|$ is maximal; then $\|\boldsymbol{\eta}\| \leqslant \sqrt{k}\|\boldsymbol{\eta}^{(\ell)}\|$. Now the normal line to $\boldsymbol{\eta}^{(\ell)}$ in $\mathbb{R}^2$ must have an angle $\geqslant \frac{\alpha}{2}$ to at least one of the lines $\mathbb{R}\boldsymbol{v}$ and $\mathbb{R}\boldsymbol{v}'$. Hence, either

$$|\boldsymbol{\eta}^{(\ell)}\boldsymbol{v}| \geqslant \|\boldsymbol{\eta}^{(\ell)}\|\|\boldsymbol{v}\| \sin(\tfrac{\alpha}{2}) > \tfrac{1}{2}\|\boldsymbol{\eta}^{(\ell)}\|\|\boldsymbol{v}\| \sin\alpha = \frac{\|\boldsymbol{\eta}^{(\ell)}\|}{2\|\boldsymbol{v}'\|} \geqslant \frac{\|\boldsymbol{\eta}\|}{2\sqrt{k}\|T\|}$$

or else

$$|\boldsymbol{\eta}^{(\ell)}\boldsymbol{v}'| \geqslant \|\boldsymbol{\eta}^{(\ell)}\|\|\boldsymbol{v}'\| \sin(\tfrac{\alpha}{2}) > \tfrac{1}{2}\|\boldsymbol{\eta}^{(\ell)}\|\|\boldsymbol{v}'\| \sin\alpha = \frac{\|\boldsymbol{\eta}^{(\ell)}\|}{2\|\boldsymbol{v}\|} \geqslant \frac{\|\boldsymbol{\eta}\|}{2\sqrt{k}\|T\|}.$$

Applying (44) for the appropriate column vector of T we get

$$\left|\widehat{f}(T, \boldsymbol{\eta})\right| \ll_m \frac{\|f\|_{\mathrm{C}_0^m}}{(\|\boldsymbol{\eta}\|/\|T\|)^m}. \tag{46}$$

Together, (45) and (46) imply (43). □

Using Iwasawa co-ordinates, the bound in Remark 5 can be expressed as follows, for any $0 < \beta < \frac{1}{2}$ (cf. (42)):

$$\left|\widehat{f}(u, v, \theta; \boldsymbol{\eta})\right| \ll_m \|f\|_{\mathrm{C}_0^m} \left(\frac{v}{u^2 + v^2 + 1}\right)^{m(\frac{1}{2}-\beta)} \|\boldsymbol{\eta}\|^{-m\beta}. \tag{47}$$

Arguing again as in [**38**, Lemmas 4.3 and 4.4], we now obtain the following bound on derivatives.

LEMMA 4.7. *Fix $0 < \beta < \frac{1}{2}$ and integers $m, \ell_1, \ell_2, \ell_3 \geqslant 0$. For any $\boldsymbol{\eta} \in B_k$ and $f \in \mathrm{C}_0^{m+\ell}(X)$, where $\ell = \ell_1 + \ell_2 + \ell_3$, we have*

$$\left|\left(\frac{\partial}{\partial u}\right)^{\ell_1}\left(\frac{\partial}{\partial v}\right)^{\ell_2}\left(\frac{\partial}{\partial \theta}\right)^{\ell_3}\widehat{f}(u, v, \theta; \boldsymbol{\eta})\right| \ll_{m,\ell} \|f\|_{\mathrm{C}_0^{m+\ell}} v^{-\ell_1-\ell_2}\left(\frac{v}{u^2 + v^2 + 1}\right)^{m(\frac{1}{2}-\beta)} \|\boldsymbol{\eta}\|^{-m\beta}.$$

5. *Obtaining the leading term*

Our task is to study the integral

$$\int_{\mathbb{R}} f(\Gamma\,(1_2,\boldsymbol{\xi})u(x)a(y))h(x)\,dx = \int_{\mathbb{R}} f\left(\Gamma\left(\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\boldsymbol{\xi}\right)\right)h(x)\,dx. \tag{48}$$

We may assume $0 < y \leqslant 1$ from the start, since (7) and (8) are otherwise trivial (indeed, the left-hand sides of (7) and (8) are always $\ll \|f\|_{\mathrm{C}_0^0} S_{\infty,0,2}(h)$). Decomposing f as in (37), we get that (48) is

$$\begin{aligned}
&= \int_{\mathbb{R}} \widehat{f}\left(\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\mathbf{0}\right)h(x)\,dx\\
&\quad + \sum_{\boldsymbol{\eta}\in A_k}\sum_{R\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'}\sum_{\substack{T\in\Gamma'_\infty\backslash\overline{\Gamma}'\\ T\equiv R \bmod N}} e((\,{}^t T\boldsymbol{\eta})\boldsymbol{\xi})\int_{\mathbb{R}}\widehat{f}_R\left(T\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\boldsymbol{\eta}\right)h(x)\,dx\\
&\quad + \sum_{\boldsymbol{\eta}\in B_k}\sum_{R\in\overline{\Gamma}'/\Gamma'}\sum_{\substack{T\in\overline{\Gamma}'\\ T\equiv R \bmod N}} e((\,{}^t T\boldsymbol{\eta})\boldsymbol{\xi})\int_{\mathbb{R}}\widehat{f}_R\left(T\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\boldsymbol{\eta}\right)h(x)\,dx.
\end{aligned} \tag{49}$$

Here, the change of order of summation and integration will be justified by an absolute convergence which holds for any f and h as in Theorems 1.2 or 1.3 (see Lemmata 7.3 and 8.2 as well as (102) and (103)).

Recall that $M \mapsto \widehat{f}(M,\mathbf{0})$ is invariant under $\Gamma' = \Gamma(N)$; hence the first integral in (49) is simply a weighted average along a closed horocycle in $\Gamma'\backslash G'$, a case which has been thoroughly studied in the literature (for arbitrary lattices in $G' = \mathrm{SL}(2,\mathbb{R})$) (see, in particular, [**4, 9, 37**]). By the bound by Kim and Sarnak [**18**] toward the Ramanujan conjecture, the smallest non-zero eigenvalue of the Laplace operator on the hyperbolic surface $\Gamma(N)\backslash\mathbb{H}$ satisfies $\lambda_1 \geqslant \frac{1}{4} - (\frac{7}{64})^2$. Using this in [**37**, Theorem 1, Rem. 3.4], we obtain

$$\int_{\mathbb{R}} \widehat{f}\left(\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\mathbf{0}\right)h(x)\,dx = \int_{\Gamma\backslash G} f\,d\mu\int_{\mathbb{R}} h\,dx + O\Big(\|f\|_{\mathrm{C}_0^4}\,S_{1,0,1}(h)\,y^{\frac{1}{2}-\frac{7}{64}}\Big). \tag{50}$$

Remark 6. Note that in the more general setting of Theorem 1.1, we could have, for example, $\Gamma = \Lambda \ltimes (\mathbb{Z}^2)^{\oplus k}$ with Λ being a *non-congruence* subgroup of $\mathrm{SL}(2,\mathbb{Z})$. If we would seek to extend the present methods to that case, when carrying out this first step of using equidistribution on $\Lambda\backslash\mathrm{SL}(2,\mathbb{R})$, we would obtain an analogue of (50) with an error term decaying as $O(y^{c(\Lambda)})$ for some $0 < c(\Lambda) \leqslant \frac{1}{2}$. However, in this case it is known that for certain choices of Λ the spectral gap for $\Lambda\backslash\mathrm{SL}(2,\mathbb{R})$ can be made arbitrarily small [**35**], meaning that there is no uniform lower bound on the exponent $c(\Lambda)$.

6. *Cancellation in an exponential sum*

In this section, we derive bounds on certain exponential sums which give nontrivial cancellations in various sums that arise frequently in our arguments in the rest of the paper. Recall that $\overline{\Gamma}' = \mathrm{SL}(2,\mathbb{Z})$ and $\Gamma' = \Gamma(N)$. Let $R = \left(\begin{smallmatrix}a_0 & b_0\\ c_0 & d_0\end{smallmatrix}\right) \in \overline{\Gamma}'$ be given. We set $[R] := \Gamma' R = R\Gamma'$; this is the set of all matrices in $\overline{\Gamma}'$ which are congruent to R modulo N. We let $\Gamma'_\infty\backslash[R]$ be a set of representatives for the right cosets of Γ'_∞ contained in $[R]$, and let $\Gamma'_\infty\backslash[R]/\Gamma'_\infty$ be a set of representatives for the double cosets of the form $\Gamma'_\infty T\Gamma'_\infty$ with $T \in [R]$. For any given integer

$c \equiv c_0 \bmod N$, we consider the following subsets:

$$[\Gamma'_\infty\backslash[R]\,;\,c] := \left\{ \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \Gamma'_\infty\backslash[R] \;:\; c_1 = c \right\}$$

and

$$[\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c] := \left\{ \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \Gamma'_\infty\backslash[R]/\Gamma'_\infty \;:\; c_1 = c \right\}.$$

Note that $[\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c]$ is a finite set. We introduce the symbol $\sum^{(1)}$ to denote summation over all matrices in $[\Gamma'_\infty\backslash[R]\,;\,c]$, and $\sum^{(2)}$ to denote summation over all matrices in $[\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c]$. Note that the summation range in both $\sum^{(1)}$ and $\sum^{(2)}$ depend implicitly on c, N and R.

REMARK 7. In the rest of this section, we will assume $c \neq 0$. Note that we have an obvious bijection, $T \mapsto -T$, between the two sets $[\Gamma'_\infty\backslash[R]\,;\,c]$ and $[\Gamma'_\infty\backslash[-R]\,;\,-c]$. Hence, without loss of generality we may assume $c > 0$.

For any N, R, c as above with $c > 0$, and $m, n \in \mathbb{Z}$, we introduce the following generalized Kloosterman sum:

$$S(m,n;c;R,N) = \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(2)} e\left(m\frac{d}{cN} + n\frac{a}{cN}\right). \tag{51}$$

This sum is well defined, since, for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in [\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c]$, both $d \bmod cN$ and $a \bmod cN$ are independent of the choice of coset representative. We begin by deriving bounds for the sums $S(m,n;c;R,N)$.

LEMMA 6.1. *Let c and N be positive integers, let $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'$ with $c_0 \equiv c \bmod N$, and let M_1, M_2 be coprime positive integers such that $cN = M_1M_2$. Then*

$$S(m,n;c;R,N) = S(m,\overline{M}_2^2 n;K_3;R_1,K_1)S(m,\overline{M}_1^2 n;K_4;R_2,K_2), \tag{52}$$

where $K_1 = (N,M_1), K_2 = (N,M_2), K_3 = (c,M_1), K_4 = (c,M_2)$, and $\overline{M}_1 \in \mathbb{Z}$ is a multiplicative inverse of M_1 mod M_2, $\overline{M}_2 \in \mathbb{Z}$ is a multiplicative inverse of M_2 mod M_1, and

$$R_1 \equiv \begin{pmatrix} M_2a_0 & K_4b_0 \\ K_3 & \overline{M}_2d_0 \end{pmatrix} \bmod K_1, \quad R_2 \equiv \begin{pmatrix} M_1a_0 & K_3b_0 \\ K_4 & \overline{M}_1d_0 \end{pmatrix} \bmod K_2. \tag{53}$$

Note that the existence of matrices $R_1, R_2 \in \overline{\Gamma}'$ satisfying (53) is guaranteed (see, for example, [**26**, Theorem 4.2.1]).

Proof. By a straightforward analysis one verifies that the map $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \langle a \bmod cN, d \bmod cN\rangle$ gives a bijection from $[\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c]$ onto the set $U[c,N;a_0,b_0,d_0]$ consisting of all pairs $\langle a,d\rangle$ in $(\mathbb{Z}/cN\mathbb{Z})^2$ satisfying $a \equiv a_0 \bmod N$, $d \equiv d_0 \bmod N$ and $ad \equiv 1 + b_0c \bmod cN$. Hence,

$$S(m,n;c;R,N) = \sum_{\langle a,d\rangle \in U[c,N;a_0,b_0,d_0]} e\left(m\frac{d}{cN} + n\frac{a}{cN}\right). \tag{54}$$

The formula (52) now follows since the map

$$\langle\langle a,d\rangle,\langle a',d'\rangle\rangle \mapsto \langle M_2\overline{M}_2^2 a + M_1\overline{M}_1^2 a', M_2d + M_1d'\rangle$$

is a bijection from $U\left[K_3, K_1; M_2 a_0, K_4 b_0, \overline{M_2} d_0\right] \times U\left[K_4, K_2; M_1 a_0, K_3 b_0, \overline{M_1} d_0\right]$ onto $U[c, N; a_0, b_0, d_0]$, with inverse $\langle A, D\rangle \mapsto \langle\langle M_2 A, \overline{M}_2 D\rangle, \langle M_1 A, \overline{M}_1 D\rangle\rangle$. □

For n a positive integer, we write $\sigma(n)$ for the number of (positive) divisors of n, and $\sigma_1(n)$ for their sum: $\sigma(n) = \sum_{d|n} 1$ and $\sigma_1(n) = \sum_{d|n} d$.

We now use the multiplicativity relation to prove that the generalized Kloosterman sums satisfy a Weil-type bound (cf. (55)), and to give an explicit formula in the case $n = 0$.

LEMMA 6.2. *For any $m, n \in \mathbb{Z}$, $c, N \in \mathbb{Z}^+$ and $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'$ with $c_0 \equiv c \bmod N$,*

$$|S(m, n; c; R, N)| \ll_N \sigma(c)(m, n, c)^{1/2} c^{1/2}, \tag{55}$$

where $\sigma(c)$ is the number of (positive) divisors of c. Moreover, in the case $n = 0$, if we write $c = c_1 c_2$ with $c_1, c_2 \in \mathbb{Z}^+$, $c_1 \mid N^{\infty}$ [†] and $(c_2, N) = 1$, then

$$\mathrm{S}(m, 0; c; R, N) = I(c_1 \mid m)\, \mu\left(\frac{c}{(c, m)}\right) \frac{\phi(c_2) c_1}{\phi(c/(c, m))} e\left(m \frac{\overline{c_2} d_0}{c_1 N}\right), \tag{56}$$

where $I(\cdot)$ is the indicator function and $\overline{c_2}$ is a multiplicative inverse of $c_2 \bmod N$. In particular,

$$|\mathrm{S}(m, 0; c; R, N)| \leqslant (c, m). \tag{57}$$

Proof. Let c_1, c_2 be as in the statement of the lemma, and set $N' = (c_1^{\infty}, N)$ and $N'' = N/N'$. Applying Lemma 6.1 twice, first with $M_1 = c_1 N$, $M_2 = c_2$ and next with $M_1 = c_1 N'$, $M_2 = N''$, we get

$$S(m, n; c; R, N) = S(m, n'; c_1; R', N') S(m, n''; 1; R'', N'') S(m, n_2; c_2; R_2, 1), \tag{58}$$

for some $n', n'', n_2 \in \mathbb{Z}$ and $R', R'', R_2 \in \overline{\Gamma}'$. Here, $|S(m, n''; 1; R'', N'')| = 1$, and the third factor is a standard Kloosterman sum; $S(m, n_2; c_2; R_2, 1) = S(m, n_2; c_2)$. Regarding the first factor, elementary arguments give, with $R' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$:

$$\begin{aligned} S(m, n'; c_1; R', N') &= e\left(n' \frac{\overline{d'} b'}{N'}\right) \sum_{\substack{d \in \mathbb{Z}/c_1 N' \mathbb{Z} \\ d \equiv d' \bmod N'}} e\left(m \frac{d}{c_1 N'} + n' \frac{\overline{d}}{c_1 N'}\right), \\ &= \frac{e(n' \overline{d'} b'/N')}{N'} \sum_{j \in \mathbb{Z}/N'\mathbb{Z}} e\left(-\frac{j d'}{N'}\right) S(m + j c_1, n'; c_1 N'). \end{aligned}$$

Now (55) follows using Weil's bound on the standard Kloosterman sum [**39**], [**15**, Chapter 11.7]. Also (56) and (57) follow, using basic facts about Ramanujan sums (see, for example, [**15**, Chapter 3.2]). □

We are now set to state and prove the main lemma in this section.

LEMMA 6.3. *Let $N, c \in \mathbb{Z}^+$ and $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'$, with $c_0 \equiv c \bmod N$. Write $c = c_1 c_2$, where $c_1 \mid N^{\infty}$ and $(c_2, N) = 1$. Let $F(x_1, x_2)$ be a function in $\mathrm{C}^4(\mathbb{R} \times (\mathbb{R}/N\mathbb{Z}))$ such that F and its derivatives $\partial_{x_1}^j \partial_{x_2}^k F$ for $j, k \leqslant 2$ are in $\mathrm{L}^1(\mathbb{R} \times (\mathbb{R}/N\mathbb{Z}))$. Then for any subset $K \subset \mathbb{Z}$ and any $\alpha \in \mathbb{R}$,*

[†] That is, $c_1 \mid N^{\ell}$ for all sufficiently large ℓ; equivalently, each prime divisor of c_1 divides N.

$$\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(1)} e(d\alpha) F\left(\frac{d}{c}, \frac{a}{c}\right)$$

$$= \sum_{\substack{m \in K \\ c_1 | m}} \left(\int_{\mathbb{R} \times (\mathbb{R}/\mathbb{Z})} F(Nx_1, Nx_2) e((cN\alpha - m)x_1)\, dx_1\, dx_2 \right) \mu\left(\frac{c}{(c,m)}\right) \frac{\phi(c_2)c_1}{\phi(c/(c,m))} e\left(\frac{m\overline{c_2}d_0}{c_1 N}\right)$$

$$+ O\big(\|F\|_{\mathrm{L}^1} + \|\partial_{x_1}^2 F\|_{\mathrm{L}^1}\big) \sum_{m \in \mathbb{Z} \setminus K} \frac{(c,m)}{1 + |m - cN\alpha|^2} + O\big(\|\partial_{x_2}^2 F\|_{\mathrm{L}^1} + \|\partial_{x_1}^2 \partial_{x_2}^2 F\|_{\mathrm{L}^1}\big) \sigma(c) \sqrt{c}, \tag{59}$$

where $\overline{c_2}$ is a multiplicative inverse of c_2 mod N.

We remark that the sum in the left-hand side of (59) is well defined, since, for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in [\Gamma'_\infty \backslash [R]\,;\, c]$, both d and the congruence class of a modulo cN are independent of the choice of a coset representative.

Proof. Set

$$H(x_1, x_2) = \sum_{\ell \in \mathbb{Z}} f(x_1 + \ell, x_2), \quad \text{where } f(x_1, x_2) := F(Nx_1, Nx_2) e(cN\alpha x_1). \tag{60}$$

Note that since $f \in \mathrm{C}^4 \cap \mathrm{L}^1(\mathbb{R} \times (\mathbb{R}/\mathbb{Z}))$; the sum defining $H(x_1, x_2)$ is absolutely convergent for almost all $(x_1, x_2) \in \mathbb{R} \times (\mathbb{R}/\mathbb{Z})$, and any $H \in \mathrm{L}^1(\mathbb{R}^2/\mathbb{Z}^2)$. We will use the notation $F_{j,k} = \partial_{x_1}^j \partial_{x_2}^k F$ and $f_{j,k} = \partial_{x_1}^j \partial_{x_2}^k f$. In order to get a stronger convergence statement, we note that

$$|f(x_1, x_2)| \leqslant \int_{x_1 - \frac{1}{2}}^{x_1 + \frac{1}{2}} (|f(r, x_2)| + |f_{1,0}(r, x_2)|)\, dr. \tag{61}$$

This follows by integrating the inequality $|f(x_1, x_2)| \leqslant |f(r, x_2)| + \int_r^{x_1} |f_{1,0}(t, x_2)|\, |dt|$ over $r \in (x_1 - \frac{1}{2}, x_1 + \frac{1}{2})$. Similarly, we have $|f_{j,0}(r, x_2)| \leqslant \int_{\mathbb{R}/\mathbb{Z}} (|f_{j,0}(r, s)| + |f_{j,1}(r, s)|)\, ds$, and using this in (61), we obtain the following elementary Sobolev embedding-type inequality:

$$|f(x_1, x_2)| \leqslant \int_{x_1 - \frac{1}{2}}^{x_1 + \frac{1}{2}} \int_{\mathbb{R}/\mathbb{Z}} (|f(r, s)| + |f_{1,0}(r, s)| + |f_{0,1}(r, s)| + |f_{1,1}(r, s)|)\, ds\, dr. \tag{62}$$

Using (62) and the fact that $f_{j,k} \in \mathrm{L}^1(\mathbb{R} \times (\mathbb{R}/\mathbb{Z}))$ for $j, k \leqslant 1$, we conclude that the sum in (60) is absolutely convergent for *all* (x_1, x_2), uniformly over (x_1, x_2) in any compact set. In particular, the function $H(x_1, x_2)$ is defined everywhere on $\mathbb{R}^2/\mathbb{Z}^2$, and is continuous.

Consider the Fourier coefficients of H,

$$a_{m,n} = \int_{\mathbb{R}^2/\mathbb{Z}^2} H(x_1, x_2) e(-mx_1 - nx_2)\, dx_1\, dx_2$$

$$= \int_{\mathbb{R} \times (\mathbb{R}/\mathbb{Z})} F(Nx_1, Nx_2) e((cN\alpha - m)x_1 - nx_2)\, dx_1\, dx_2. \tag{63}$$

Note that for any $j \leqslant 1$ and $k \leqslant 2$,

$$\int_{\mathbb{R}/\mathbb{Z}} |F_{j,k}(x_1, Nx_2)|\, dx_2 \to 0 \quad \text{as } x_1 \to \pm\infty. \tag{64}$$

This follows by applying (61) to $F_{j,k}(x_1, Nx_2)$ and using $F_{j,k}, F_{j+1,k} \in \mathrm{L}^1(\mathbb{R} \times (\mathbb{R}/N\mathbb{Z}))$. We may now integrate by parts repeatedly in (63), using (64) to justify convergence, to obtain

$$a_{m,n} = \frac{N^{j+k}}{(2\pi i)^{j+k}(m - cN\alpha)^j n^k} \int_{\mathbb{R}\times\mathbb{R}/\mathbb{Z}} F_{j,k}(Nx_1, Nx_2) e((cN\alpha - m)x_1 - nx_2)\, dx_1\, dx_2, \quad (65)$$

for any $0 \leqslant j, k \leqslant 2$ and any integers m, n subject to $m \neq cN\alpha$ if $j > 0$ and $n \neq 0$ if $k > 0$. Using this formula for $j \in \{0, 2\}$ and $k = 2$ gives

$$|a_{m,n}| \ll_N (\|F_{0,2}\|_{\mathrm{L}^1} + \|F_{2,2}\|_{\mathrm{L}^1}) \min(1, |m - cN\alpha|^{-2}) n^{-2}, \quad \forall m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\}. \quad (66)$$

Similarly, using (65) for $j \in \{0, 2\}$ and $k = 0$,

$$|a_{m,0}| \ll_N (\|F\|_{\mathrm{L}^1} + \|F_{2,0}\|_{\mathrm{L}^1}) \min(1, |m - cN\alpha|^{-2}), \quad \forall m \in \mathbb{Z}. \quad (67)$$

These bounds imply that the Fourier series of H is absolutely convergent; and since H is continuous, H is in fact equal to its Fourier series at every point (see, for example, [**12**, Proposition 3.1.14]):

$$H(x_1, x_2) = \sum_{m,n\in\mathbb{Z}} a_{m,n} e(mx_1 + nx_2). \quad (68)$$

Now we consider the sum in the left-hand side of (59). We have

$$\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(1)} e(d\alpha) F\left(\frac{d}{c}, \frac{a}{c}\right) = \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(2)} \sum_{\ell\in\mathbb{Z}} F\left(\frac{d + \ell cN}{c}, \frac{a}{c}\right) e(\alpha(d + \ell cN)) = \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(2)} H\left(\frac{d}{cN}, \frac{a}{cN}\right). \quad (69)$$

Here, all sums are absolutely convergent, since the sum in (60) is absolutely convergent and $\sum^{(2)}$ runs over a finite set. Substituting (68) in the last sum, and using (51), we obtain

$$\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(1)} e(d\alpha) F\left(\frac{d}{c}, \frac{a}{c}\right) = \sum_{m,n\in\mathbb{Z}} a_{m,n} S(m, n; c; R, N). \quad (70)$$

Now we bound the contribution from all terms with $n \neq 0$ in (70) using (66), (55) and $\sum_{n\neq 0}(m, n, c)^{1/2} n^{-2} \leqslant \sum_{n\neq 0} |n|^{-3/2} \ll 1$, while the terms with $n = 0$ are handled using (63) and (56) when $m \in K$, and using (67) and (57) when $m \notin K$. In this way, we obtain (59). □

Remark 8. If $c < 0$, $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'$ and $c_0 \equiv c \bmod N$, then we see from Remark 7 that the sum $\sum^{(1)} e(d\alpha) F\left(\frac{d}{c}, \frac{a}{c}\right)$ remains the same if we replace $\langle c, R, \alpha\rangle$ by $\langle -c, -R, -\alpha\rangle$; after this replacement, Lemma 6.3 applies to the sum.

Lemma 6.3 will suffice for most parts of our discussion. However, at one step in the treatment of the sum over B_k in (49), we will need a more delicate estimate. The point here is to obtain a bound which only involves derivatives $\partial_{x_1}^{\ell_1} \partial_{x_2}^{\ell_2} F$ with ℓ_2 as small as possible. Lemma 6.3 requires using $\ell_2 = 2$ but the following lemma will effectively allow us to take $\ell_2 = \frac{1}{2} + \varepsilon$ (see also Remark 11). We define a mixed $\mathrm{L}^1, \mathrm{L}^2$ norm for functions F on $\mathbb{R} \times \mathbb{R}/N\mathbb{Z}$ as follows:

$$\|F\|_{\mathrm{L}^{1,2}} = \left(\int_{\mathbb{R}/N\mathbb{Z}} \left(\int_{\mathbb{R}} |F(x_1, x_2)|\, dx_1 \right)^2 dx_2 \right)^{1/2}.$$

LEMMA 6.4. *Let $0<\varepsilon<1$ and let N, c, R be as before. Let $F(x_1,x_2)$ be a function in $\mathrm{C}^3(\mathbb{R}\times(\mathbb{R}/N\mathbb{Z}))$ such that $\|\partial_{x_1}^j\partial_{x_2}^k F\|_{\mathrm{L}^{1,2}}<\infty$ for $j\leqslant 2$, $k\leqslant 1$. Then for any $\alpha\in\mathbb{R}$,*

$$\left|\sum_{\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)}^{(1)} e(d\alpha)F\left(\frac{d}{c},\frac{a}{c}\right)\right| \ll_\varepsilon (\|F\|_{\mathrm{L}^1}+\|\partial_{x_1}^2F\|_{\mathrm{L}^1})\sum_{\ell\in\mathbb{Z}}\frac{(c,\lfloor cN\alpha+\ell\rfloor)}{1+\ell^2}$$

$$+\left(\|F\|_{\mathrm{L}^{1,2}}+\|\partial_{x_1}^2F\|_{\mathrm{L}^{1,2}}\right)^{\frac{1-\varepsilon}{2}}\left(\|\partial_{x_2}F\|_{\mathrm{L}^{1,2}}+\|\partial_{x_1}^2\partial_{x_2}F\|_{\mathrm{L}^{1,2}}\right)^{\frac{1+\varepsilon}{2}}\sigma(c)^{3/2}\sqrt{c}. \tag{71}$$

Proof. Note that $\|F_{j,k}\|_{\mathrm{L}^1}\leqslant\sqrt{N}\|F_{j,k}\|_{\mathrm{L}^{1,2}}$ by Cauchy–Schwarz. Hence, as in the proof of Lemma 6.3, $H(x_1,x_2)$ in (60) is a well-defined continuous function on $\mathbb{R}^2/\mathbb{Z}^2$, and its Fourier coefficients $a_{m,n}$ satisfy (65) for any $j\leqslant 2$, $k\leqslant 1$; that is,

$$a_{m,n}=\frac{N^{j+k}}{(2\pi i)^{j+k}(m-cN\alpha)^j n^k}\int_{\mathbb{R}/\mathbb{Z}}F_{m,j,k}(x_2)e(-nx_2)\,dx_2,$$

where $F_{m,j,k}(x_2)=\int_{\mathbb{R}}F_{j,k}(Nx_1,Nx_2)e((cN\alpha-m)x_1)\,dx_1$ is a function on $\mathbb{R}/\mathbb{Z}$. This gives a relation between $a_{m,n}$ and the nth Fourier coefficient of $F_{m,j,k}$. Using this relation for $j\in\{0,2\}$ and applying Parseval's identity, for any $k\geqslant 0$ and $m\in\mathbb{Z}$, we get

$$\sum_{n\in\mathbb{Z}\setminus\{0\}}n^{2k}|a_{m,n}|^2\ll_{k,N}(\|F_{m,0,k}\|_{L^2}^2+\|F_{m,2,k}\|_{L^2}^2)\min(1,|m-cN\alpha|^{-4})$$

$$\ll_N(\|F_{0,k}\|_{L^{1,2}}^2+\|F_{2,k}\|_{L^{1,2}}^2)\min(1,|m-cN\alpha|^{-4}). \tag{72}$$

Using this bound, $\sum_{n\neq0}|a_{m,n}|\leqslant(\sum_{n\neq0}|n|^{-2})^{\frac12}(\sum_{n\neq0}n^2|a_{m,n}|^2)^{\frac12}$, and (67), we conclude that the Fourier series of H is absolutely convergent, and hence as in the proof of Lemma 6.3, we again have

$$\sum_{\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)}^{(1)} e(d\alpha)F\left(\frac{d}{c},\frac{a}{c}\right)=\sum_{m,n\in\mathbb{Z}}a_{m,n}S(m,n;c;R,N). \tag{73}$$

Using (67) and (57) for $n=0$, and the generalized Weil bound (55) for $n\neq0$, we see that (73) is

$$\ll(\|F\|_{\mathrm{L}^1}+\|F_{2,0}\|_{\mathrm{L}^1})\sum_{\ell\in\mathbb{Z}}\frac{(c,\lfloor cN\alpha+\ell\rfloor)}{1+\ell^2}+\sigma(c)\sqrt{c}\sum_{m\in\mathbb{Z}}\sum_{n\neq0}|a_{m,n}|\sqrt{(n,c)}. \tag{74}$$

Note that for any integer m,

$$\sum_{n\neq0}|a_{m,n}|\sqrt{(n,c)}\leqslant\sqrt{\sum_{n\neq0}\frac{(n,c)}{|n|^{1+\varepsilon}}}\sqrt{\sum_{n\neq0}|a_{m,n}|^2|n|^{1+\varepsilon}}.$$

Now, since $0<\varepsilon<1$, we may apply Hölder's inequality with $p=\frac{2}{1-\varepsilon}$ and $q=\frac{2}{1+\varepsilon}$, to get

$$\sum_{n\neq0}|a_{m,n}|^2|n|^{1+\varepsilon}=\sum_{n\neq0}|a_{m,n}|^{\frac2p}\cdot\left(|a_{m,n}|^{\frac2q}|n|^{1+\varepsilon}\right)\leqslant\left(\sum_{n\neq0}|a_{m,n}|^2\right)^{\frac1p}\left(\sum_{n\neq0}|a_{m,n}|^2|n|^{(1+\varepsilon)q}\right)^{\frac1q}$$

$$\ll(\|F\|_{\mathrm{L}^{1,2}}+\|F_{2,0}\|_{\mathrm{L}^{1,2}})^{1-\varepsilon}(\|F_{0,1}\|_{\mathrm{L}^{1,2}}+\|F_{2,1}\|_{\mathrm{L}^{1,2}})^{1+\varepsilon}(1+|cN\alpha-m|)^{-4}.$$

Here, in the last step we use the Parseval bound, (72), for $k=0$ and $k=1$. Furthermore,

$$\sum_{n\neq 0}\frac{(n,c)}{|n|^{1+\varepsilon}} = 2\sum_{d|c}\sum_{\substack{m\geqslant 1\\(m,c)=d}}\frac{d}{m^{1+\varepsilon}} \leqslant 2\sum_{d|c} d\sum_{k=1}^{\infty}\frac{1}{(kd)^{1+\varepsilon}} \ll_\varepsilon \sum_{d|c} d^{-\varepsilon} \leqslant \sigma(c).$$

Hence, for any m,

$$\sum_{n\neq 0}|a_{m,n}|\sqrt{(n,c)}$$

$$\ll_\varepsilon (\|F\|_{\mathrm{L}^{1,2}} + \|F_{2,0}\|_{\mathrm{L}^{1,2}})^{\frac{1-\varepsilon}{2}}(\|F_{0,1}\|_{\mathrm{L}^{1,2}} + \|F_{2,1}\|_{\mathrm{L}^{1,2}})^{\frac{1+\varepsilon}{2}}(1+|cN\alpha - m|)^{-2}\sqrt{\sigma(c)}.$$

Using this bound in (74), we obtain (71). □

7. *The contribution from A_k-orbits*

7.1. *The case of Diophantine $\boldsymbol{\xi}_2$*

We next study the sum in the second line of (49). This sum will be bounded by a generalization of the method in [**38**]. We first prove a bound which is adequate for *any* $\boldsymbol{\xi} = \binom{\boldsymbol{\xi}_1}{\boldsymbol{\xi}_2} \in \mathbb{R}^{2k}$ for which $\boldsymbol{\xi}_2$ has good Diophantine properties. This bound will be used in the special case $\boldsymbol{\xi}_1 = \mathbf{0}$ in the proof of Theorem 1.2. We note that we allow the special case $k=1$ in the present section, to allow comparison with [**38**, Proposition 8.3] (cf. Remark 10).

PROPOSITION 7.1. *Fix an integer $m \geqslant \max(8, k+3)$ and real numbers $a \in (\frac{k}{2}-\frac{1}{2}, \frac{m}{2}-1)$ and $\varepsilon > 0$. Then for any $f \in \mathrm{C}_a^m(X)$, $h \in \mathrm{C}^2(\mathbb{R})$ with $S_{1,0,2}(h) < \infty$, $\boldsymbol{\xi} = \binom{\boldsymbol{\xi}_1}{\boldsymbol{\xi}_2} \in \mathbb{R}^{2k}$ and $0 < y \leqslant 1$, we have*

$$\sum_{\boldsymbol{\eta}\in A_k}\sum_{R\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'}\sum_{T\in\Gamma'_\infty\backslash[R]} e((\,{}^{\mathrm{t}}T\boldsymbol{\eta})\boldsymbol{\xi})\int_{\mathbb{R}}\widehat{f}_R\left(T\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\boldsymbol{\eta}\right)h(x)\,dx$$

$$\ll_{m,a,\varepsilon} \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h)\left(\widetilde{\delta}_{2a+1,\boldsymbol{\xi}_2}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon}\right). \tag{75}$$

(Recall that the majorant function $\widetilde{\delta}_{\beta,\boldsymbol{\xi}_2}(T)$ was introduced in (27).)

To start with the proof of Proposition 7.1, let us fix some $\boldsymbol{\eta} = \binom{0}{\boldsymbol{r}} \in A_k$ and $R = \left(\begin{smallmatrix}a_0 & b_0\\ c_0 & d_0\end{smallmatrix}\right) \in \overline{\Gamma}'$. Using the notation introduced in Section 6, the corresponding inner sum in (75) can be written as

$$\sum_{c\in c_0+N\mathbb{Z}}\sum_{\binom{a\ b}{c\ d}}^{(1)}\int_{\mathbb{R}}\widehat{f}_R\left(\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)e\left(\begin{pmatrix}c\boldsymbol{r}\\ d\boldsymbol{r}\end{pmatrix}\boldsymbol{\xi}\right)h(x)\,dx. \tag{76}$$

The contribution from the terms with $c=0$ can be bounded easily. Indeed, there are at most two such terms in (76), and by Lemma 4.4 and the remarks below (19), for any $b\in\mathbb{Z}$ we have

$$\int_{\mathbb{R}}\left|\widehat{f}_R\left(\pm\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix},\begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)h(x)\right|dx \ll \|h\|_{\mathrm{L}^1}\|f\|_{\mathrm{C}_0^m}y^{m/2}\|\boldsymbol{r}\|^{-m}. \tag{77}$$

Using this with $m=k+1$ and adding over all $\boldsymbol{\eta}\in A_k$, we see that the contribution from all the terms with $T = \left(\begin{smallmatrix}* & *\\ 0 & *\end{smallmatrix}\right)$ in the second line of (49) is $O(\|h\|_{\mathrm{L}^1}\|f\|_{\mathrm{C}_0^{k+1}}y^{(k+1)/2})$, which is clearly subsumed by the bound in (75).

Hence, from now on we focus on the terms with $c \neq 0$. The following lemma expresses the integral in (76) in the Iwasawa notation (cf. (40)).

LEMMA 7.2. *For any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G'$ with $c > 0$, any $y > 0$ and any $f \in \mathrm{C}(G')$,*

$$\int_{\mathbb{R}} f\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}\right) h(x)\,dx = \int_0^{\pi} f\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta\right) h\left(-\frac{d}{c} + y\cot\theta\right) \frac{y\,d\theta}{\sin^2\theta},$$

in the sense that if either of the integrals is absolutely convergent then so is the other, and the equality holds.

REMARK 9. In the case $c < 0$ one obtains exactly the same formula, except that $\int_0^{\pi}$ is replaced by $\int_{-\pi}^{0}$ in the right-hand side.

Proof. See [**38**, Lemma 6.1]. □

We now prove that we have an absolute convergence in the left-hand side of (75); this fact is important in order to justify the manipulations which we will carry out later.

LEMMA 7.3. *Set $m = \max(3, k+1)$. Then for any $f \in \mathrm{C}_0^m(X)$ and any $h \in \mathrm{C}^1(\mathbb{R})$ with $S_{1,0,1}(h) < \infty$, the expression*

$$\sum_{\boldsymbol{\eta} \in A_k} \sum_{R \in \overline{\Gamma}'_\infty \backslash \overline{\Gamma}'/\Gamma'} \sum_{T \in \Gamma'_\infty \backslash [R]} \int_{\mathbb{R}} \left| \widehat{f}_R\left(T\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}, \boldsymbol{\eta}\right) h(x) \right| dx \tag{78}$$

is finite for all $y > 0$. If, furthermore, $f \in \mathrm{C}_a^m(X)$ for some a and m subject to $a \geqslant 0$, $a > \frac{k}{2} - 1$ and $m > 2a + 2$, then the expression in (78) stays bounded as $y \to 0$.

(Note that the lemma in particular applies to any f and h as in Proposition 7.1.)

Proof. As previously, we write $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. The contribution from terms with $c = 0$ in (78) is treated by (77). Thus, we only consider the terms with $c > 0$; the terms with $c < 0$ can be dealt with similarly. By Lemma 7.2, and since $\overline{\Gamma}'_\infty \backslash \overline{\Gamma}'/\Gamma'$ is finite, it suffices to prove that for each fixed $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'_\infty \backslash \overline{\Gamma}'/\Gamma'$,

$$\sum_{\boldsymbol{r} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}} \sum_{\substack{c \equiv c_0 \bmod N \\ c > 0}} \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(1)} \int_0^{\pi} \left| \widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix}\right) h\left(-\frac{d}{c} + y\cot\theta\right) \right| \frac{y\,d\theta}{\sin^2\theta} < \infty. \tag{79}$$

By Lemma 4.5 (and the observations below (19)), for any $m \geqslant 0$ and $a \in \mathbb{R}_{\geqslant 0}$ we have

$$\left| \widehat{f}_R\left(u, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix}\right) \right| \ll \|f\|_{\mathrm{C}_a^m} \|\boldsymbol{r}\|^{-m} \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^m \min\left(1, \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{-2a}\right),$$

uniformly over $u \in \mathbb{R}$. Using this bound for both $m = 0$ and a general $m \geqslant 0$, we conclude

$$\begin{aligned} &\left| \widehat{f}_R\left(u, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{r} \end{pmatrix}\right) \right| \\ &\ll \|f\|_{\mathrm{C}_a^m} \min\left(\|\boldsymbol{r}\|^{-m} \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^m, \|\boldsymbol{r}\|^{-m} \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{m-2a}, \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{-2a} \right). \end{aligned} \tag{80}$$

We decompose the innermost sum in (79) in the same way as in (69), and then use the fact that

$$\sum_{n\in\mathbb{Z}} |h(\delta+n)| \leqslant S_{1,0,1}(h), \quad \forall \delta\in\mathbb{R}, \tag{81}$$

which holds since $|h(\alpha)| \leqslant \int_{\alpha-1/2}^{\alpha+1/2} (|h(x)|+|h'(x)|)\,dx$ for all $\alpha\in\mathbb{R}$. From the proof of Lemma 6.1, we also have

$$\#[\Gamma'_\infty\backslash[R]/\Gamma'_\infty\,;\,c] = \#U[c,N;a_0,b_0,d_0] \leqslant c.$$

Hence, we conclude that if $f\in \mathrm{C}_a^m(X)$ and $S_{1,0,1}(h)<\infty$, then the left-hand side of (79) is

$$\ll \sum_{\boldsymbol{r}\in\mathbb{Z}^k\backslash\{\mathbf{0}\}} \sum_{c=1}^{\infty} c\int_0^\pi \min\left(\|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^m, \|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{m-2a}, \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{-2a}\right)\frac{y\,d\theta}{\sin^2\theta}.$$

Assuming $m>2a+2$, we get (cf. Lemma 7.4)

$$\begin{aligned} &\ll \sum_{\boldsymbol{r}\in\mathbb{Z}^k\backslash\{\mathbf{0}\}} \sum_{c=1}^{\infty} cy \begin{Bmatrix} \|\boldsymbol{r}\|^{-m}(c\sqrt{y})^{-m} & \text{if } 1\leqslant c\sqrt{y} \\ \|\boldsymbol{r}\|^{-m}(c\sqrt{y})^{2a-m} & \text{if } \|\boldsymbol{r}\|^{-1}\leqslant c\sqrt{y}\leqslant 1 \\ \|\boldsymbol{r}\|^{-2a-1}(c\sqrt{y})^{-1} & \text{if } c\sqrt{y}\leqslant \|\boldsymbol{r}\|^{-1} \end{Bmatrix} \\ &\ll \sum_{\boldsymbol{r}\in\mathbb{Z}^k\backslash\{\mathbf{0}\}} \min\left(\|\boldsymbol{r}\|^{-2-2a}, \|\boldsymbol{r}\|^{-m}y^{1+a-\frac{m}{2}}\right). \end{aligned}$$

(Here, $m>2a+1$ suffices for the first step, while $m>2a+2$ is needed to get the last bound.) The last sum converges provided that either $m>k$ or $2a+2>k$; and if $2a+2>k$, then it also stays bounded as $y\to 0$. □

In the proof above, we used the following bound, which we will need again later.

LEMMA 7.4. *Fix $a\geqslant 0$ and $m>2a+1$. Then for any $u>0$ and $r\geqslant 1$, we have*

$$\begin{aligned} &\int_0^\pi \min\left(r^{-m}\left(\frac{|\sin\theta|}{u}\right)^m, r^{-m}\left(\frac{|\sin\theta|}{u}\right)^{m-2a}, \left(\frac{|\sin\theta|}{u}\right)^{-2a}\right)\frac{d\theta}{\sin^2\theta} \\ &\ll \begin{cases} r^{-m}u^{-m} & \text{if } 1\leqslant u \\ r^{-m}u^{2a-m} & \text{if } r^{-1}\leqslant u\leqslant 1 \\ r^{-2a-1}u^{-1} & \text{if } u\leqslant r^{-1}. \end{cases} \end{aligned}$$

Proof. This is a straightforward case-by-case analysis. □

We continue with the proof of Proposition 7.1. Using Lemma 7.2 and Remark 9, the sum in (76), excluding all terms with $c=0$, can be expressed as

$$\begin{aligned} &\sum_{\substack{c\equiv c_0 \bmod N\\ c>0}} \int_0^\pi \sum_{\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)}^{(1)} h\left(-\frac{d}{c}+y\cot\theta\right)\widehat{f}_R\left(\frac{a}{c}-\frac{\sin 2\theta}{2c^2y}, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right) e\left(\begin{pmatrix} c\boldsymbol{r}\\ d\boldsymbol{r}\end{pmatrix}\boldsymbol{\xi}\right)\frac{y\,d\theta}{\sin^2\theta} \\ &+ \sum_{\substack{c\equiv c_0 \bmod N\\ c<0}} \int_{-\pi}^0 \sum_{\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)}^{(1)} h\left(-\frac{d}{c}+y\cot\theta\right)\widehat{f}_R\left(\frac{a}{c}-\frac{\sin 2\theta}{2c^2y}, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right) e\left(\begin{pmatrix} c\boldsymbol{r}\\ d\boldsymbol{r}\end{pmatrix}\boldsymbol{\xi}\right)\frac{y\,d\theta}{\sin^2\theta}. \end{aligned} \tag{82}$$

Here, the change of order of summation and integration is justified by Lemma 7.3. We will only deal with the first sum in (82); the second sum can be dealt with similarly (cf. Remark 8). By Lemma 6.3 (applied with $K = \emptyset$ and $\alpha = \boldsymbol{r}\boldsymbol{\xi}_2$), for any positive integer $c \equiv c_0 \bmod N$ and any $\theta \in (0, \pi)$, we have

$$\begin{aligned}
&\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(1)} h\Big(-\frac{d}{c} + y\cot\theta\Big)\, \widehat{f}_R\Big(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\Big) e(c\boldsymbol{r}\boldsymbol{\xi}_1 + d\boldsymbol{r}\boldsymbol{\xi}_2) \\
&\quad \ll S_{1,0,2}(h)\Big(\int_{\mathbb{R}/N\mathbb{Z}} \Big|\widehat{f}_R\Big(u, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\Big)\Big|\, du\Big) \sum_{\ell\in\mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{r}\boldsymbol{\xi}_2 + \ell\rfloor)}{1+\ell^2} \qquad (83)\\
&\quad + S_{1,0,2}(h)\Big(\int_{\mathbb{R}/N\mathbb{Z}} \Big|\frac{\partial^2}{\partial u^2}\widehat{f}_R\Big(u, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\Big)\Big|\, du\Big)\sigma(c)\sqrt{c}.
\end{aligned}$$

Here, we will use the bound (80). By a similar application of Lemma 4.5 as in (80), we have for any $m' \in \mathbb{Z}_{\geqslant 6}$ and $a' \in \mathbb{R}_{\geqslant 0}$, uniformly over $u \in \mathbb{R}$:

$$\begin{aligned}
&\Big|\frac{\partial^2}{\partial u^2}\widehat{f}_R\Big(u, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\Big)\Big| \\
&\ll \|f\|_{\mathrm{C}^{m'}_{a'}} \|\boldsymbol{r}\|^{-4} \min\left(\|\boldsymbol{r}\|^{6-m'}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m'-6}, \|\boldsymbol{r}\|^{6-m'}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m'-6-2a'}, \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{-2a'}\right).
\end{aligned}$$

Using these bounds, we conclude that the first sum in (82) is

$$\begin{aligned}
&\ll S_{1,0,2}(h)\Bigg\{\|f\|_{\mathrm{C}^m_a} \sum_{c=1}^{\infty}\int_0^{\pi} \min\left(\|\boldsymbol{r}\|^{-m}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m}, \|\boldsymbol{r}\|^{-m}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m-2a}, \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{-2a}\right) \\
&\quad \times \frac{y\,d\theta}{\sin^2\theta} \sum_{\ell\in\mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{r}\boldsymbol{\xi}_2 + \ell\rfloor)}{1+\ell^2} \\
&\quad + \|f\|_{\mathrm{C}^{m'}_{a'}} \|\boldsymbol{r}\|^{-4} \sum_{c=1}^{\infty}\int_0^{\pi} \min\left(\Big(\frac{|\sin\theta|}{\|\boldsymbol{r}\|c\sqrt{y}}\Big)^{m'-6}, \|\boldsymbol{r}\|^{6-m'}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m'-6-2a'}, \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{-2a'}\right) \\
&\quad \times \frac{y\,d\theta}{\sin^2\theta}\sigma(c)\sqrt{c}\Bigg\}. \qquad (84)
\end{aligned}$$

By Lemma 7.4, assuming $m > 2a+1$ and $m' > 2a'+7$, (84) is

$$\begin{aligned}
&\ll \|f\|_{\mathrm{C}^m_a} S_{1,0,2}(h) \|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}} \sum_{c=1}^{\infty} c^{-1}\Big((\|\boldsymbol{r}\|\sqrt{y})^{-1} + c\Big)^{1+2a-m} \sum_{\ell\in\mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{r}\boldsymbol{\xi}_2 + \ell\rfloor)}{1+\ell^2} \\
&\quad + \|f\|_{\mathrm{C}^{m'}_{a'}} S_{1,0,2}(h) \|\boldsymbol{r}\|^{2-m'} y^{4+a'-\frac{m'}{2}} \sum_{c=1}^{\infty}\Big((\|\boldsymbol{r}\|\sqrt{y})^{-1} + c\Big)^{7+2a'-m'} \frac{\sigma(c)}{\sqrt{c}}. \qquad (85)
\end{aligned}$$

LEMMA 7.5. *Fix $\beta > 1$. Then for any $\alpha \in \mathbb{R}$ and $X > 0$,*

$$\sum_{c=1}^{\infty} c^{-1}(X+c)^{-\beta} \sum_{k\in\mathbb{Z}} \frac{(c,k)}{1+|k-c\alpha|^2} \ll_\beta \begin{cases} X^{1-\beta} \sum_{j=1}^{\infty} \big(j^2 + Xj\langle j\alpha\rangle\big)^{-1} & \text{if } X \geqslant 1 \\ 1 & \text{if } X < 1. \end{cases} \tag{86}$$

Proof. If $\beta \in \mathbb{Z}$, then this is [**38**, Lemma 8.2] (with $\eta = 1$ and $m = \beta + 1$). The proof extends without changes to the case of an arbitrary real $\beta > 1$. □

LEMMA 7.6. *For any $X > 0$ and $\beta > \frac{1}{2}$,*

$$\sum_{c=1}^{\infty} (X+c)^{-\beta} \frac{\sigma(c)}{\sqrt{c}} \ll_\beta \begin{cases} X^{\frac{1}{2}-\beta} \log(1+X) & \text{if } X \geqslant 1 \\ 1 & \text{if } X < 1. \end{cases}$$

Proof. See [**38**, Lemma 8.1]. This follows by using $\sum_{1\leqslant c\leqslant x} \sigma(c) \ll x \log(1+x)$, $\forall x \geqslant 1$ (see, for example, [**15**, (1.75)]), and integration by parts. □

Using Lemmas 7.5 and 7.6, and assuming from now on that $m > 2a+2$ and $m' > 2a' + \frac{15}{2}$, we find that (85) (and thus (84)) is

$$\begin{aligned} &\ll \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h) \|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}} \left\{ \begin{matrix} (\|\boldsymbol{r}\|\sqrt{y})^{m-2a-2} \sum_{j=1}^{\infty} \min(j^{-2}, \frac{\|\boldsymbol{r}\|\sqrt{y}}{j\langle j\boldsymbol{r}\boldsymbol{\xi}_2\rangle}) & \text{if } \|\boldsymbol{r}\|\sqrt{y} \leqslant 1 \\ 1 & \text{if } \|\boldsymbol{r}\|\sqrt{y} > 1 \end{matrix} \right\} \\ &\quad + \|f\|_{\mathrm{C}_{a'}^{m'}} S_{1,0,2}(h) \|\boldsymbol{r}\|^{2-m'} y^{4+a'-\frac{m'}{2}} \\ &\quad \times \left\{ \begin{matrix} (\|\boldsymbol{r}\|\sqrt{y})^{m'-2a'-\frac{15}{2}} \log(1+(\|\boldsymbol{r}\|\sqrt{y})^{-1}) & \text{if } \|\boldsymbol{r}\|\sqrt{y} \leqslant 1 \\ 1 & \text{if } \|\boldsymbol{r}\|\sqrt{y} > 1 \end{matrix} \right\} \\ &\ll \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h) \|\boldsymbol{r}\|^{-2a-1} \left\{ \begin{matrix} \sum_{j=1}^{\infty} \min(j^{-2}, \frac{\sqrt{y}}{j\langle j\boldsymbol{r}\boldsymbol{\xi}_2\rangle}) & \text{if } \|\boldsymbol{r}\|\sqrt{y} \leqslant 1 \\ (\|\boldsymbol{r}\|\sqrt{y})^{2a+1-m}\sqrt{y} & \text{if } \|\boldsymbol{r}\|\sqrt{y} > 1 \end{matrix} \right\} \\ &\quad + \|f\|_{\mathrm{C}_{a'}^{m'}} S_{1,0,2}(h) \|\boldsymbol{r}\|^{2-m'} y^{4+a'-\frac{m'}{2}} \left\{ \begin{matrix} (\|\boldsymbol{r}\|\sqrt{y})^{m'-2a'-\frac{15}{2}-\varepsilon} & \text{if } \|\boldsymbol{r}\|\sqrt{y} \leqslant 1 \\ 1 & \text{if } \|\boldsymbol{r}\|\sqrt{y} > 1 \end{matrix} \right\}. \end{aligned} \tag{87}$$

In order to obtain a bound on the left-hand side of (75), we have to sum over R running through the finite set $\overline{\Gamma}'_\infty \backslash \overline{\Gamma}'/\Gamma'$, and add over all $\boldsymbol{\eta} \in A_k$, which means that $\boldsymbol{r}$ runs through a subset of $\mathbb{Z}^k \setminus \{\mathbf{0}\}$. For this to give a satisfactory result, we have to assume $2a+1 > k$, while in the second bound we choose $a' = \max(\frac{k}{2} - \frac{11}{4}, 0)$; with this choice, $m' = \max(8, k+3)$ satisfies the condition $m' > 2a' + \frac{15}{2}$. Adding now over R and $\boldsymbol{\eta}$, we conclude that the left-hand side of (75) is $\ll \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h) \widetilde{\delta}_{2a+1,\boldsymbol{\xi}_2}(y^{-\frac{1}{2}}) + \|f\|_{\mathrm{C}_{a'}^{m'}} S_{1,0,2}(h) y^{\frac{1}{4}-\varepsilon}$. Finally, we note that $a > a'$, and so if we also assume $m \geqslant m'$, then $\|f\|_{\mathrm{C}_{a'}^{m'}} \ll \|f\|_{\mathrm{C}_a^m}$, and we obtain the bound stated in Proposition 7.1.

REMARK 10. In (87), the somewhat crude inequality $\min\big(j^{-2}, \frac{\|\boldsymbol{r}\|\sqrt{y}}{j\langle j\boldsymbol{r}\boldsymbol{\xi}_2\rangle}\big) \leqslant \|\boldsymbol{r}\| \min\big(j^{-2}, \frac{\sqrt{y}}{j\langle j\boldsymbol{r}\boldsymbol{\xi}_2\rangle}\big)$ was used. In the special case $k = 1$, by avoiding using this bound one can keep $a = 0$ in the treatment, that is, no cuspidal decay of f has to be required (cf. [**38**, Proposition 8.3]). Note also that [**38**, Proposition 8.3] has a better dependence on the test function h (called 'ν' in [**38**]) than Proposition 7.1, namely, essentially, '$S_{1,0,1+\varepsilon}(h)$' in place of $S_{1,0,2}(h)$. We have avoided this in the present paper for simplicity of presentation.

7.2. *The case* $\boldsymbol{\xi}_2 = \mathbf{0}$

In this case, we prove the following bound.

PROPOSITION 7.7. *Fix an integer* $m \geqslant \max(8, k+3)$ *and real numbers* $a \in (\frac{k}{2} - \frac{1}{2}, \frac{m}{2} - 1)$ *and* $\varepsilon > 0$. *Then for any* $f \in \mathrm{C}_a^m(X)$, $h \in \mathrm{C}^2(\mathbb{R})$ *with* $S_{1,0,2}(h) < \infty$, $\boldsymbol{\xi} = \left(\begin{smallmatrix}\boldsymbol{\xi}_1\\ \mathbf{0}\end{smallmatrix}\right) \in \mathbb{R}^{2k}$ *and* $0 < y \leqslant 1$, *we have*

$$\sum_{\boldsymbol{\eta}\in A_k} \sum_{R\in\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'} \sum_{T\in\Gamma'_\infty\backslash[R]} e((\,{}^t T\boldsymbol{\eta})\boldsymbol{\xi}) \int_{\mathbb{R}} \widehat{f}_R\left(T\begin{pmatrix}\sqrt{y} & x/\sqrt{y}\\ 0 & 1/\sqrt{y}\end{pmatrix}, \boldsymbol{\eta}\right) h(x)\,dx$$

$$\ll_{m,a,\varepsilon} \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h)\left(\delta_{2a+1,\boldsymbol{\xi}_1}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon}\right). \tag{88}$$

Proof. Arguing as in the proof of Proposition 7.1, we arrive again at the expression in (82), where we now have $\left(\begin{smallmatrix}cr\\ dr\end{smallmatrix}\right)\boldsymbol{\xi} = c\boldsymbol{r}\boldsymbol{\xi}_1$. Applying Lemma 6.3 with $\alpha = 0$ and $K = \{0\}$ gives, for any positive integer $c \equiv c_0 \bmod N$, decomposed as $c = c_1c_2$ where $c_1 \mid N^\infty$ and $(c_2, N) = 1$, and any $\theta \in (0, \pi)$:

$$\sum_{\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)}^{(1)} h\left(-\frac{d}{c} + y\cot\theta\right) \widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2y}, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)$$

$$= c_1\phi(c_2)\int_{\mathbb{R}} h(-Nx_1 + y\cot\theta)dx_1 \int_{\mathbb{R}/\mathbb{Z}} \widehat{f}_R\left(Nx_2, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)dx_2$$

$$+ O\left(S_{1,0,2}(h)\int_{\mathbb{R}/N\mathbb{Z}} \left|\widehat{f}_R\left(u, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)\right| du\right) \sum_{\ell\in\mathbb{Z}\backslash\{0\}} \frac{(c,\ell)}{1+\ell^2}$$

$$+ O\left(S_{1,0,2}(h)\int_{\mathbb{R}/N\mathbb{Z}} \left|\frac{\partial^2}{\partial u^2}\widehat{f}_R\left(u, \frac{\sin^2\theta}{c^2y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)\right| du\right)\sigma(c)\sqrt{c}. \tag{89}$$

Note here that the error term in the last line is the same as in the last line in (83); hence it can be bounded as before (cf. (84) and (87)). In the remaining error term in (89), we have

$$\sum_{\ell\in\mathbb{Z}\backslash\{0\}} \frac{(c,\ell)}{1+\ell^2} \leqslant 2\sum_{m|c}\sum_{n=1}^{\infty} \frac{m}{(mn)^2} \ll \sum_{m|c}\frac{1}{m} \leqslant \sigma(c).$$

We can now argue as in the proof of Proposition 7.1, but instead of Lemma 7.5 using the simple bound

$$\|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}} \sum_{c=1}^{\infty} c^{-1}\left((\|\boldsymbol{r}\|\sqrt{y})^{-1} + c\right)^{1+2a-m}\sigma(c) \ll \begin{cases}\|\boldsymbol{r}\|^{-1-2a-\varepsilon} y^{\frac{1-\varepsilon}{2}} & \text{if } \|\boldsymbol{r}\|\sqrt{y} \leqslant 1\\ \|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}} & \text{if } \|\boldsymbol{r}\|\sqrt{y} \geqslant 1,\end{cases}$$

which is valid under the assumption that $m > 2a+1$, and for any fixed $\varepsilon > 0$. This leads to the conclusion that the contribution from the error terms in the last two lines of (89) to the left-hand side of (88) is

$$\ll \|f\|_{\mathrm{C}_a^m} S_{1,0,2}(h) y^{\frac{1}{2}-\varepsilon} + \|f\|_{\mathrm{C}_{a'}^{m'}} S_{1,0,2}(h) y^{\frac{1}{4}-\varepsilon},$$

with $a = \frac{k}{2} - \frac{1}{2}$, $m = k+1$, $a' = \max(\frac{k}{2} - \frac{11}{4}, 0)$ and $m' = \max(8, k+3)$. This is clearly subsumed by the right-hand side of (88).

Now, it only remains to consider the first line in the right-hand side of (89). The contribution from this line to the expression in the first line of (82) can be written as follows, after expressing the indicator function of $c \equiv c_0 \bmod N$ as $N^{-1}\sum_{b \bmod N} e(b(c-c_0)/N)$:

$$\frac{1}{N^2}\int_{\mathbb{R}} h(x)\,dx \sum_{b \bmod N} e\left(-\frac{bc_0}{N}\right)\int_0^\pi \int_{\mathbb{R}/\mathbb{Z}} \sum_{c>0} e(c\alpha)c_1\phi(c_2)\widehat{f}_R\left(Nx_2, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right) dx_2\,\frac{y\,d\theta}{\sin^2\theta}, \tag{90}$$

where $\alpha := \boldsymbol{r}\boldsymbol{\xi}_1 + b/N$. We will use integration by parts to handle the sum over c. Thus, we let

$$B_\alpha(X) = \sum_{1\leqslant c\leqslant X} e(c\alpha)c_1\phi(c_2) = \sum_{\substack{1\leqslant c_2\leqslant X\\ (c_2,N)=1}} \sum_{\substack{1\leqslant c_1\leqslant X/c_2\\ c_1|N^\infty}} e(c_1c_2\alpha)c_1\phi(c_2).$$

We have the following bound, analogous to [**38**, Lemma 9.2].

LEMMA 7.8. *For any $\alpha \in \mathbb{R}$, and $X \geqslant 1$,*

$$B_\alpha(X) \ll X^2 \sum_{1\leqslant j\leqslant X} \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle j\alpha\rangle}\right).$$

Proof. For any $c_2 > 0$, we have $\phi(c_2) = \sum_{d|c_2}\mu(c_2/d)d$. Using this formula and substituting $c_2 = jd$, we get

$$B_\alpha(X) = \sum_{\substack{1\leqslant j\leqslant X\\ (j,N)=1}}\mu(j)\sum_{\substack{1\leqslant d\leqslant X/j\\ (d,N)=1}}\sum_{\substack{1\leqslant c_1\leqslant X/(jd)\\ c_1|N^\infty}} dc_1 e(jdc_1\alpha) = \sum_{\substack{1\leqslant j\leqslant X\\ (j,N)=1}}\mu(j)\sum_{1\leqslant k\leqslant X/j} ke(jk\alpha). \tag{91}$$

However, for any $j, n \in \mathbb{Z}^+$,

$$\sum_{1\leqslant k\leqslant n} ke(jk\alpha) \ll \min\left(n^2, \frac{n}{\langle j\alpha\rangle}\right).$$

(See the proof of [**38**, Lemma 9.2].) Applying this bound to (91), we get the lemma. □

For any $m \geqslant 0$ and $a \in \mathbb{R}_{\geqslant 0}$, by Lemma 4.5 we have (in a similar way as in (80))

$$\begin{aligned}&\frac{\partial}{\partial X}\widehat{f}_R\left(Nx_2, \frac{\sin^2\theta}{X^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)\\ &\quad\ll \|f\|_{\mathrm{C}_a^{m+1}} X^{-1}\min\left(\|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^m, \|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^{m-2a}, \left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^{-2a}\right).\end{aligned} \tag{92}$$

Using summation by parts in (90) (justified using (92) and $B_\alpha(X) \ll X^2$), we have

$$\sum_{c>0} e(c\alpha)c_1\phi(c_2)\widehat{f}_R\left(Nx_2, \frac{\sin^2\theta}{c^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right) = -\int_1^\infty \left(\frac{\partial}{\partial X}\widehat{f}_R\left(Nx_2, \frac{\sin^2\theta}{X^2 y}, \theta; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{r}\end{pmatrix}\right)\right) B_\alpha(X)\,dX.$$

Furthermore, Lemma 7.4 implies that for $m > 2a+1$,

$$\begin{aligned}&\int_0^\pi \min\left(\|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^m, \|\boldsymbol{r}\|^{-m}\left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^{m-2a}, \left(\frac{|\sin\theta|}{X\sqrt{y}}\right)^{-2a}\right)\frac{y\,d\theta}{\sin^2\theta}\\ &\quad\ll X^{-1}\|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}}\left((\|\boldsymbol{r}\|\sqrt{y})^{-1} + X\right)^{1+2a-m}.\end{aligned}$$

Hence, also using Lemma 7.8 and $\langle j\alpha\rangle = \langle j(\boldsymbol{r}\boldsymbol{\xi}_1 + b/N)\rangle \geqslant N^{-1}\langle jN\boldsymbol{r}\boldsymbol{\xi}_1\rangle$, we find that the expression in (90) is

$$\ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}} \frac{y^{1+a-\frac{m}{2}}}{\|\boldsymbol{r}\|^m} \int_1^\infty \Big((\|\boldsymbol{r}\|\sqrt{y})^{-1} + X\Big)^{1+2a-m} \sum_{1\leqslant j\leqslant X} \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle jN\boldsymbol{r}\boldsymbol{\xi}_1\rangle}\right) dX. \tag{93}$$

Lemma 7.9. *Assume $m > 2a+2$. Then for any $\beta \in \mathbb{R}$ and $U > 0$ we have*

$$\int_1^\infty (U+X)^{1+2a-m} \sum_{1\leqslant j\leqslant X} \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle j\beta\rangle}\right) dX$$

$$\ll_{m,a} (U+1)^{2+2a-m} \sum_{j=1}^\infty \min\left(\frac{1}{j^2}, \frac{1}{Uj\langle j\beta\rangle}\right)\left(1 + \log^+\left(\frac{U\langle j\beta\rangle}{j}\right)\right). \tag{94}$$

Proof. Changing the order of summation and integration, the left-hand side of (94) becomes

$$\sum_{j=1}^\infty \int_j^\infty (U+X)^{1+2a-m} \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle j\beta\rangle}\right) dX.$$

Here, for each $j \geqslant U$ we use $(U+X)^{1+2a-m} \leqslant X^{1+2a-m}$ and $\min(\frac{1}{j^2}, \frac{1}{Xj\langle j\beta\rangle}) \leqslant j^{-2}$, to see that $\int_j^\infty \cdots dX \leqslant j^{2a-m}$. On the other hand, for $j < U$ we have

$$\int_j^\infty \cdots dX \leqslant U^{1+2a-m} \int_j^U \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle j\beta\rangle}\right) dX + \min\left(\frac{1}{j^2}, \frac{1}{Uj\langle j\beta\rangle}\right) \int_j^\infty X^{1+2a-m}\, dX$$

$$\ll U^{2+2a-m} \min\left(\frac{1}{j^2}, \frac{1}{Uj\langle j\beta\rangle}\right)\left(1 + \log^+\left(\frac{U\langle j\beta\rangle}{j}\right)\right),$$

where the last bound is proved by splitting into the two cases $U \leqslant \frac{j}{\langle j\beta\rangle}$ and $U > \frac{j}{\langle j\beta\rangle}$ and evaluating the integrals. The proof of the lemma is completed by adding up our bounds over all positive integers j, and noticing that $\sum_{j\geqslant U} j^{2a-m} \ll (U+1)^{1+2a-m}$, which is bounded above by the contribution from $j = 1$ in the right-hand side of (94). □

Assuming now $m > 2a+2$, using the lemma we get, via (93), that the expression in (90) is

$$\ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}} \frac{(1+\|\boldsymbol{r}\|\sqrt{y})^{2+2a-m}}{\|\boldsymbol{r}\|^{2+2a}} \sum_{j=1}^\infty \min\left(\frac{1}{j^2}, \frac{\|\boldsymbol{r}\|\sqrt{y}}{j\langle jN\boldsymbol{r}\boldsymbol{\xi}_1\rangle}\right)\left(1 + \log^+\left(\frac{\langle jN\boldsymbol{r}\boldsymbol{\xi}_1\rangle}{\|\boldsymbol{r}\|\sqrt{y}j}\right)\right)$$

$$\ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}} \frac{(1+\|\boldsymbol{r}\|\sqrt{y})^{2+2a-m}}{\|\boldsymbol{r}\|^{2+2a}} \sum_{j=1}^\infty \min\left(\frac{1}{j^2}, \frac{\|\boldsymbol{r}\|\sqrt{y}}{j\langle j\boldsymbol{r}\boldsymbol{\xi}_1\rangle}\right)\left(1 + \log^+\left(\frac{\langle j\boldsymbol{r}\boldsymbol{\xi}_1\rangle}{\|\boldsymbol{r}\|\sqrt{y}j}\right)\right). \tag{95}$$

(Indeed, the last bound holds even if the last sum over j is restricted to $j = N, 2N, 3N, \ldots$.) Finally, we have to add this bound over all R in the finite set $\overline{\Gamma}'_\infty\backslash\overline{\Gamma}'/\Gamma'$, and over all $\boldsymbol{\eta} \in A_k$, which means that $\boldsymbol{r}$ runs through a subset of $\mathbb{Z}^k \setminus \{\mathbf{0}\}$. Comparing with the definition (5), assuming now $a > \frac{k-1}{2}$ (that is, $2a+1 > k$), we immediately find that the sum of the bound in (95) over all $\boldsymbol{r} \in \mathbb{Z}^k$ with $0 < \|\boldsymbol{r}\| < y^{-1/2}$ is

$$\ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}}\, \delta_{2a+1,\boldsymbol{\xi}_1}(y^{-\frac{1}{2}}). \tag{96}$$

On the other hand, for $\boldsymbol{r}$ with $\|\boldsymbol{r}\| \geqslant y^{-1/2}$, the sum over j in (95) equals $\sum_{j=1}^{\infty} j^{-2} = \pi^2/6$, and hence the sum of the bound in (95) over all such $\boldsymbol{r}$ is, assuming $m > k$

$$\ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}} \sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^k \\ (\|\boldsymbol{r}\|\geqslant y^{-1/2})}} \|\boldsymbol{r}\|^{-m} y^{1+a-\frac{m}{2}} \ll S_{1,0,0}(h)\|f\|_{\mathrm{C}_a^{m+1}}\, y^{1+a-\frac{k}{2}}.$$

However, this is subsumed by the bound (96), since $a > \frac{k-1}{2}$ and $\delta_{\mu,\boldsymbol{\xi}}(T) \geqslant (T+1)^{-1}$ for all $T > 0$ (as is clear by taking $\boldsymbol{r} = \boldsymbol{e}_1$, $j = 1$ in (5)). Hence, for any fixed $a > \frac{k-1}{2}$ and $m > 2a+2$, $m \in \mathbb{Z}$, we have proved that the contribution from the first line in the right-hand side of (89) to the left-hand side of (88) is bounded by (96). This completes the proof of Proposition 7.7. □

8. *The contribution from B_k-orbits*

8.1. *The case $\boldsymbol{\xi}_1 = \mathbf{0}$*

In this section, we will bound the sum in the third line of (49). We will assume $k \geqslant 2$ throughout this section, since B_k is empty for $k = 1$. We will prove the following.

PROPOSITION 8.1. *Let $k \geqslant 2$. Fix a real number $\varepsilon > 0$ and an integer $m \geqslant \max(8, 2k+1)$. For any $f \in \mathrm{C}_0^{3m+3}(X)$, $h \in \mathrm{C}^1(\mathbb{R})$ with $S_{\infty,2+\varepsilon,1}(h) < \infty$, $\boldsymbol{\xi}_2 \in \mathbb{R}^k$ and $0 < y \leqslant 1$, we have*

$$\sum_{\boldsymbol{\eta}\in B_k} \sum_{R\in\overline{\Gamma}'/\Gamma'} \sum_{T\in[R]} e\left(({}^{\mathrm{t}}T\boldsymbol{\eta}) \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix} \right) \int_{\mathbb{R}} \widehat{f}_R\left(T \begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}, \boldsymbol{\eta} \right) h(x)\, dx$$

$$\ll_{m,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+3}} S_{\infty,2+\varepsilon,1}(h) \left(\delta_{m-k,\boldsymbol{\xi}_2}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon} \right). \tag{97}$$

Note that Theorem 1.2 follows from Proposition 8.1 together with Proposition 7.1 and the relations (49) and (50).

To start the proof of Proposition 8.1, note that taking $\beta = \frac{1}{3}$ in Lemma 4.7, replacing m by $3m$ and using the remarks below (19), we get

$$\left| \partial_u^{\ell_1} \partial_v^{\ell_2} \partial_\theta^{\ell_3} \widehat{f}_R(u, v, \theta; \boldsymbol{\eta}) \right| \ll_{m,\ell} \|f\|_{\mathrm{C}_0^{3m+\ell}} \|\boldsymbol{\eta}\|^{-m}\, v^{-\ell_1-\ell_2} \left(\frac{v}{u^2+v^2+1} \right)^{m/2}, \tag{98}$$

for all $R \in \overline{\Gamma}'$, $u \in \mathbb{R}$, $v > 0$, $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, $\boldsymbol{\eta} \in B_k$ and $\ell_1, \ell_2, \ell_3 \geqslant 0$, with $\ell = \ell_1 + \ell_2 + \ell_3$.

Any $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \overline{\Gamma}'$ with $c = 0$ can be expressed as $T = \varepsilon \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$, where $\varepsilon \in \{-1, 1\}$ and $n \in \mathbb{Z}$, and the contribution from these T to the left-hand side of (97) is

$$\ll \sum_{\boldsymbol{\eta}\in B_k} \sum_{\varepsilon\in\{-1,1\}} \sum_{n\in\mathbb{Z}} \int_{\mathbb{R}} \left| \widehat{f}_R\left(\varepsilon \begin{pmatrix} \sqrt{y} & (x+n)/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}, \boldsymbol{\eta} \right) h(x) \right| dx, \tag{99}$$

wherein R denotes the unique element in our chosen system of representatives $\overline{\Gamma}'/\Gamma'$ satisfying $R \equiv \varepsilon \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right) \bmod N$. Using (98) and $m \geqslant 2k+1$, we get that the sum in consideration is

$$\ll \|f\|_{\mathrm{C}_0^{3m}} \sum_{\boldsymbol{\eta}\in B_k} \|\boldsymbol{\eta}\|^{-m} \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} \frac{y^{m/2}}{(1+|x+n|)^m} |h(x)|\, dx \ll \|f\|_{\mathrm{C}_0^{3m}} y^{\frac{m}{2}} \sum_{\boldsymbol{\eta}\in B_k} \|\boldsymbol{\eta}\|^{-m} \int_{\mathbb{R}} |h(x)|\, dx$$

$$\ll \|f\|_{\mathrm{C}_0^{3m}} \|h\|_{\mathrm{L}^1} y^{\frac{m}{2}}. \tag{100}$$

This is clearly subsumed by the bound in (97).

Hence, from now on we focus on the terms for $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $c \neq 0$ in the left-hand side of (97). We will restrict to the case $c > 0$; the case $c < 0$ can be handled completely analogously. We fix some $\boldsymbol{\eta} = \left(\begin{smallmatrix} q \\ r \end{smallmatrix}\right) \in B_k$ and $R = \left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right) \in \overline{\Gamma}'$. Using Lemma 7.2, the inner sum can be expressed as:

$$\sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in [R] \\ c>0}} e((b\boldsymbol{q} + d\boldsymbol{r})\boldsymbol{\xi}_2) \int_0^\pi \widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) h\left(-\frac{d}{c} + y\cos\theta\right) \frac{y\, d\theta}{\sin^2\theta}. \tag{101}$$

Let us first record a trivial upper bound on (101), variants of which will be used repeatedly below.

LEMMA 8.2. *For any $\boldsymbol{\eta} \in B_k$ and $R \in \overline{\Gamma}'/\Gamma'$,*

$$\sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in [R] \\ c>0}} \int_0^\pi \left|\widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) h\left(-\frac{d}{c} + y\cos\theta\right)\right| \frac{y\, d\theta}{\sin^2\theta}$$

$$\ll_m \|f\|_{\mathrm{C}_0^{3m}} S_{\infty,2,0}(h) \|\boldsymbol{\eta}\|^{-m}. \tag{102}$$

Proof. We overestimate the sum by letting $\langle a, c, d\rangle$ run through *all* integer triples with $c > 0$ and $ad \equiv 1 \bmod c$. Using (98) we then get that the left-hand side of (102) is

$$\ll \|f\|_{\mathrm{C}_0^{3m}} \|\boldsymbol{\eta}\|^{-m} \sum_{c=1}^\infty \sum_{\substack{d \in \mathbb{Z} \\ (d,c)=1}} \int_0^\pi \sum_{n\in\mathbb{Z}} \left(\frac{v}{u_n^2 + v^2 + 1}\right)^{m/2} \left|h\left(-\frac{d}{c} + y\cos\theta\right)\right| \frac{y\, d\theta}{\sin^2\theta}, \tag{103}$$

where $v = v(y, c, \theta) = \frac{\sin^2\theta}{c^2 y}$ and $u_n = u_n(y, c, d, \theta) = n + \frac{\alpha}{c} - \frac{\sin 2\theta}{2c^2 y}$, with $\alpha = \alpha(c, d)$ being the unique integer between 1 and c satisfying $\alpha d \equiv 1 \bmod c$. But here

$$\sum_{n\in\mathbb{Z}} \left(\frac{v}{u_n^2 + v^2 + 1}\right)^{m/2} \ll \sum_{\substack{n\in\mathbb{Z} \\ |u_n| \leqslant 1+v}} \left(\frac{v}{v^2+1}\right)^{m/2} + \sum_{\substack{n\in\mathbb{Z} \\ |u_n| > 1+v}} \left(\frac{v}{u_n^2}\right)^{m/2} \ll \min\left(v^{\frac{m}{2}}, v^{1-\frac{m}{2}}\right), \tag{104}$$

where we used the fact that $m > 2k \geqslant 2$. Furthermore, if $S_{\infty,2,0}(h) < \infty$, then we have

$$\sum_{d\in\mathbb{Z}} \left|h\left(-\frac{d}{c} + y\cos\theta\right)\right| \leqslant S_{\infty,2,0}(h) \sum_{d\in\mathbb{Z}} \left(1 + \left|-\frac{d}{c} + y\cos\theta\right|\right)^{-2} \ll S_{\infty,2,0}(h) c. \tag{105}$$

Hence, we obtain that (103) is

$$\ll \|f\|_{\mathrm{C}_0^{3m}} S_{\infty,2,0}(h) \|\boldsymbol{\eta}\|^{-m} y \sum_{c=1}^\infty c \int_0^\pi \min\left(v^{\frac{m}{2}}, v^{1-\frac{m}{2}}\right) \frac{d\theta}{\sin^2\theta}. \tag{106}$$

However,

$$\int_0^\pi \min\left(v^{\frac{m}{2}}, v^{1-\frac{m}{2}}\right) \frac{d\theta}{\sin^2\theta} \ll \min\left((c\sqrt{y})^{-1}, (c\sqrt{y})^{-m}\right), \tag{107}$$

as one verifies by treating the two cases $c^2 y \geqslant 1$ and $c^2 y < 1$ separately, and in the latter case, splitting the interval for θ into the parts $\{\theta : |\sin\theta| < c\sqrt{y}\}$ and $\{\theta : |\sin\theta| \geqslant c\sqrt{y}\}$. Now the lemma follows by using (107) in (106). □

Adding the bound in Lemma 8.2 over all $R \in \overline{\Gamma}'/\Gamma'$ and $\boldsymbol{\eta} \in B_k$ (again using $m > 2k$), we immediately see that the sum in the left-hand side of (97) stays *bounded* as $y \to 0$. In order to show that the sum actually *decays* as $y \to 0$, we have to establish cancellation in (101).

It will be convenient later (cf. Lemma 8.4) to note that we may restrict the integral in (101) to those $\theta \in (0, \pi)$ which satisfy $y|\cot\theta| \leqslant 1$. Indeed, if $y|\cot\theta| > 1$, then $|\sin\theta| < y$, and we note that for any $c \geqslant 1$ we have, with $v = \frac{\sin^2\theta}{c^2 y}$ as in the proof of Lemma 8.2,

$$\int_{\substack{0<\theta<\pi \\ (|\sin\theta|<y)}} \min\left(v^{\frac{m}{2}}, v^{1-\frac{m}{2}}\right) \frac{d\theta}{\sin^2\theta} = \int_{\substack{0<\theta<\pi \\ (|\sin\theta|<y)}} v^{\frac{m}{2}} \frac{d\theta}{\sin^2\theta} \ll_m c^{-m} y^{\frac{m}{2}-1}.$$

Using this bound in place of (107) in the proof of Lemma 8.2, we conclude that the contribution from θ with $y|\cot\theta| > 1$ in (101) is $\ll \|f\|_{\mathrm{C}_0^{3m}} S_{\infty,2,0}(h) \|\boldsymbol{\eta}\|^{-m} y^{m/2}$. Adding this over R and $\boldsymbol{\eta}$ as in the left-hand side of (97), we again obtain a bound which is (by far) subsumed by the bound in (97).

Let us also note that if $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in (101) has $d = 0$, then necessarily $c = 1$, and inspecting the proof of Lemma 8.2 we see that the contribution from all such T in (101) is $\ll \|f\|_{\mathrm{C}_0^{3m}} S_{\infty,2,0}(h) \|\boldsymbol{\eta}\|^{-m} \sqrt{y}$. This gives a contribution $\ll \|f\|_{\mathrm{C}_0^{3m}} S_{\infty,2,0}(h) \sqrt{y}$ in the left-hand side of (97), which is ok. Hence, from now on we may consider the sum in (101) restricted by $d \neq 0$.

Next we will make use of the approximation $\frac{a}{c} = \frac{1+bc}{dc} \approx \frac{b}{d}$. The error in doing so is controlled by the following lemma.

LEMMA 8.3. *Assuming that $m \geqslant 4$, we have*

$$\sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in [R] \\ c>0,\ d\neq 0}} \int_0^\pi \left| \widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) - \widehat{f}_R\left(\frac{b}{d} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) \right| \times \left| h\left(-\frac{d}{c} + y\cot\theta\right) \right| \frac{y\, d\theta}{\sin^2\theta} \ll_m \|f\|_{\mathrm{C}_0^{3m+1}} S_{\infty,2,0}(h) \frac{\sqrt{y}\log(2+y^{-1})}{\|\boldsymbol{\eta}\|^m}. \tag{108}$$

Proof. For any $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \overline{\Gamma}'$ with $c, d \neq 0$ we have, letting J be the interval with endpoints $\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}$ and $\frac{b}{d} - \frac{\sin 2\theta}{2c^2 y}$, and using $\frac{a}{c} - \frac{b}{d} = \frac{1}{dc}$ and (98),

$$\left| \widehat{f}_R\left(\frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) - \widehat{f}_R\left(\frac{b}{d} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) \right| \ll \frac{1}{|dc|} \sup_{x\in J} \left| \partial_x \widehat{f}_R\left(x, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\right) \right| \ll \|f\|_{\mathrm{C}_0^{3m+1}} \|\boldsymbol{\eta}\|^{-m} |c|^{-1} v^{-1} \left(\frac{v}{u^2+v^2+1}\right)^{m/2},$$

with $v = \frac{\sin^2\theta}{c^2 y}$ and $u = \frac{a}{c} - \frac{\sin 2\theta}{2c^2 y}$. (We used the crude bound $|d|^{-1} \leqslant 1$, and the fact that $(u+\xi)^2 + 1 \asymp u^2 + 1$ for all $u \in \mathbb{R}$, $|\xi| \leqslant 1$.) Hence, arguing as in the proof of Lemma 8.2, and using the same notation 'u_n' as there, we find that the left-hand side of (108) is

$$\ll \|f\|_{\mathrm{C}_0^{3m+1}} \|\boldsymbol{\eta}\|^{-m} \sum_{c=1}^\infty c^{-1} \sum_{\substack{d\in\mathbb{Z} \\ (d,c)=1}} \int_0^\pi \sum_{n\in\mathbb{Z}} v^{-1} \left(\frac{v}{u_n^2+v^2+1}\right)^{m/2} \left| h\left(-\frac{d}{c} + y\cot\theta\right) \right| \frac{y\, d\theta}{\sin^2\theta}.$$

The rest of the proof is very similar to Lemma 8.2, except that we now use

$$\int_0^\pi \min\left(v^{\frac{m}{2}-1}, v^{-\frac{m}{2}}\right) \frac{d\theta}{\sin^2\theta} \ll \min\left((c\sqrt{y})^{-1}, (c\sqrt{y})^{2-m}\right)$$

in place of (107). □

Adding the bound in Lemma 8.3 over all $R \in \overline{\Gamma}'/\Gamma'$ and $\boldsymbol{\eta} \in B_k$ gives a bound $\|f\|_{\mathrm{C}_0^{3m+1}} S_{\infty,2,0}(h)\sqrt{y}\log(2+y^{-1})$, and this is subsumed by the bound in (97). Hence, from now on we may replace $\frac{a}{c}$ by $\frac{b}{d}$ in (101). Restricting the summation to $d>0$ (the case $d<0$ being completely analogous), and writing $I_y := \{\theta \in (0,\pi) \,:\, y|\cot\theta| \leqslant 1\}$, the resulting sum is:

$$\sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)\in[R] \\ c>0,\ d>0}} e((b\boldsymbol{q}+d\boldsymbol{r})\boldsymbol{\xi}_2) \int_{I_y} \widehat{f}_R\Big(\frac{b}{d} - \frac{\sin 2\theta}{2c^2 y}, \frac{\sin^2\theta}{c^2 y}, \theta; \boldsymbol{\eta}\Big) h\Big(-\frac{d}{c} + y\cot\theta\Big) \frac{y\,d\theta}{\sin^2\theta}.$$

Replacing $\langle a,b,c,d\rangle$ by $\langle -b,d,-a,c\rangle$ in this sum gives, with $\widetilde{R} := \left(\begin{smallmatrix} -c_0 & -a_0 \\ d_0 & b_0\end{smallmatrix}\right)$:

$$= \sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)\in[\widetilde{R}] \\ a<0,\ c>0}} e((d\boldsymbol{q}+c\boldsymbol{r})\boldsymbol{\xi}_2) \int_{I_y} \widehat{f}_R\Big(\frac{d}{c} - \frac{\sin 2\theta}{2a^2 y}, \frac{\sin^2\theta}{a^2 y}, \theta; \boldsymbol{\eta}\Big) h\Big(\frac{c}{a} + y\cot\theta\Big) \frac{y\,d\theta}{\sin^2\theta}$$

$$= \sum_{\substack{c \equiv d_0 \bmod N \\ c>0}} e(c\boldsymbol{r}\boldsymbol{\xi}_2) \int_{I_y} \sum_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)}^{(\widetilde{1})} e(d\boldsymbol{q}\boldsymbol{\xi}_2) F_{c,\theta}\Big(\frac{d}{c}, \frac{a}{c}\Big) \frac{y\,d\theta}{\sin^2\theta}, \tag{109}$$

where $\sum^{(\widetilde{1})}$ is the same as $\sum^{(1)}$ (cf. p. 16) but using $\widetilde{R}$ in place of R, and for any $c \in \mathbb{Z}^+$ and $\theta \in (0,\pi)$, $F_{c,\theta}(x_1,x_2)$ is the function on $\mathbb{R}\times(\mathbb{R}/N\mathbb{Z})$ given by

$$F_{c,\theta}(x_1,x_2) := \sum_{\substack{s \in x_2 + N\mathbb{Z} \\ s<0}} \widehat{f}_R\Big(x_1 - \frac{\sin 2\theta}{2yc^2 s^2}, \frac{\sin^2\theta}{yc^2 s^2}, \theta; \boldsymbol{\eta}\Big) h\Big(\frac{1}{s} + y\cot\theta\Big). \tag{110}$$

(Note that $F_{c,\theta}$ also depends on $N, y, R, \boldsymbol{\eta}$.) Using $|\widehat{f}_R(u,v,\theta;\boldsymbol{\eta})| \ll \min(v, v^{-1})^{m/2}$ (cf. (98)), we see that the sum defining $F_{c,\theta}(x_1,x_2)$ is absolutely convergent, and that $F_{c,\theta}(x_1,x_2)$ is continuous on $\mathbb{R}\times(\mathbb{R}/N\mathbb{Z})$. If $F_{c,\theta}$ is sufficiently differentiable with the first few derivatives being in $\mathrm{L}^{1,2}$, then we may apply Lemma 6.4, to see that, for any $0<\varepsilon<\frac{1}{2}$,

$$\sum_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)}^{(\widetilde{1})} e(d\boldsymbol{q}\boldsymbol{\xi}_2) F_{c,\theta}\Big(\frac{d}{c}, \frac{a}{c}\Big) \ll_\varepsilon (\|F_{c,\theta}\|_{\mathrm{L}^1} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^1}) \sum_{\ell\in\mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{q}\boldsymbol{\xi}_2 + \ell\rfloor)}{1+\ell^2}$$

$$+ \big(\|F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^{1,2}}\big)^{\frac{1}{2}-\varepsilon} \big(\|\partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 \partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}}\big)^{\frac{1}{2}+\varepsilon} \sigma(c)^{3/2}\sqrt{c}. \tag{111}$$

Bounds on the $\mathrm{L}^{1,2}$-norms of derivatives of $F_{c,\theta}$ are provided by the following lemma.

LEMMA 8.4. *For any integer ℓ with $1 \leqslant \ell < \frac{1}{2}(m-1)$, we have $F_{c,\theta} \in \mathrm{C}^\ell(\mathbb{R}\times(\mathbb{R}/N\mathbb{Z}))$ provided that $f \in \mathrm{C}_0^{3m+\ell}(X)$ and $h \in \mathrm{C}^\ell(\mathbb{R})$ with $S_{\infty,0,\ell}(h) < \infty$. Furthermore, for any integers $\ell_1, \ell_2 \geqslant 0$, and $a \in \mathbb{R}_{\geqslant 0}$, $0<\varepsilon<1$, if $\ell = \ell_1+\ell_2$, $m > 2\ell+1$, $f \in \mathrm{C}_0^{3m+\ell}$, $h \in \mathrm{C}^{\ell_2}(\mathbb{R})$, $S_{\infty,a,\ell_2}(h) < \infty$ and $y|\cot\theta| \leqslant 1$, then we have*

$$\left(\int_{\mathbb{R}/N\mathbb{Z}} \left(\int_{\mathbb{R}} \left|\frac{\partial^{\ell_1+\ell_2}}{\partial x_1^{\ell_1}\partial x_2^{\ell_2}} F_{c,\theta}(x_1,x_2)\right| dx_1\right)^2 dx_2\right)^{1/2} \ll_{m,\ell,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+\ell}} S_{\infty,a,\ell_2}(h) \|\boldsymbol{\eta}\|^{-m}$$

$$\times \begin{cases} |\sin\theta|^{-\ell_2}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{1-\ell_2+\frac{\varepsilon}{2}} & \text{if } c\sqrt{y} \leqslant |\sin\theta| \\ |\sin\theta|^{-\ell_2}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{\frac12+a-\ell_2}\Big\{1+\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m-\frac12-a-2\ell_1+\ell_2}\Big\} & \text{if } |\sin\theta| \leqslant c\sqrt{y} \leqslant 1 \\ \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{\frac12+a-2\ell_2}\Big\{1+|\sin\theta|^{-\ell_2}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m-\frac12-a-2\ell_1+2\ell_2}\Big\} & \text{if } c\sqrt{y} \geqslant 1. \end{cases} \tag{112}$$

Proof. By repeated differentiation we obtain, for any $\ell_1 \geqslant 0$ and $\ell_2 \geqslant 1$,

$$\begin{aligned} &\frac{\partial^\ell}{\partial x_1^{\ell_1}\partial s^{\ell_2}}\Big(\widehat{f}_R\Big(x_1-\frac{\sin 2\theta}{2yc^2s^2},\frac{\sin^2\theta}{yc^2s^2},\theta;\boldsymbol{\eta}\Big)h\Big(\frac1s+y\cot\theta\Big)\Big) \\ &\quad=\sum_{1\leqslant\alpha+\beta+\gamma\leqslant\ell_2} K^{(\ell_2)}_{\alpha,\beta,\gamma}\frac{(\sin 2\theta)^\alpha(\sin\theta)^{2\beta}}{(yc^2s^2)^{\alpha+\beta}s^{\gamma+\ell_2}}\Big[\partial_1^{\ell_1+\alpha}\partial_2^\beta\widehat{f}_R\Big]\Big(x_1-\frac{\sin 2\theta}{2yc^2s^2},\frac{\sin^2\theta}{yc^2s^2},\theta;\boldsymbol{\eta}\Big) \\ &\qquad\times h^{(\gamma)}\Big(\frac1s+y\cot\theta\Big), \end{aligned} \tag{113}$$

where $\langle\alpha,\beta,\gamma\rangle$ runs through all triples of nonnegative integers satisfying $1\leqslant\alpha+\beta+\gamma\leqslant\ell_2$, each coefficient $K^{(\ell_2)}_{a,\beta,\gamma}$ is an integer, and ∂_1 and ∂_2 denote differentiation with respect to the first and second argument of $\widehat{f}_R$. Using (98), we find that the absolute value of (113) is

$$\begin{aligned} &\ll_{m,\ell}\|f\|_{\mathrm{C}_0^{3m+\ell}}\|\boldsymbol{\eta}\|^{-m}\bigg(\Big(x_1-\frac{\sin 2\theta}{2yc^2s^2}\Big)^2+\Big(\frac{\sin^2\theta}{yc^2s^2}\Big)^2+1\bigg)^{-\frac{m}{2}}\Big(\frac{\sin^2\theta}{yc^2s^2}\Big)^{\frac{m}{2}-\ell_1}|s|^{-\ell_2} \\ &\quad\times\sum_{1\leqslant\alpha+\beta+\gamma\leqslant\ell_2}|\sin\theta|^{-\alpha}|s|^{-\gamma}\Big|h^{(\gamma)}\Big(\frac1s+y\cot\theta\Big)\Big|. \end{aligned} \tag{114}$$

Here, the sum in the second line is $\asymp_{\ell_2}\sum_{\gamma=0}^{\ell_2}|\sin\theta|^{\gamma-\ell_2}|s|^{-\gamma}|h^{(\gamma)}(s^{-1}+y\cot\theta)|$. On the other hand for $\ell_2=0$, the left-hand side of (113) trivially equals $[\partial_1^{\ell_1}\widehat{f}_R](\dots)\,h(s^{-1}+y\cot\theta)$, and thus the bound in (114) is again valid, with the last sum replaced by '$|h(s^{-1}+y\cot\theta)|$'.

Now assume $S_{\infty,0,\ell_2}(h)<\infty$. Then the bound in (114) is $\ll|s|^{-m+2\ell_1-\ell_2}$ for $|s|\geqslant1$ and $\ll|s|^{m+2\ell_1-2\ell_2}$ for $0<|s|\leqslant1$, uniformly with respect to $x_1\in\mathbb{R}$ when keeping all other parameters fixed. Hence, if $m>\max(2\ell_1-\ell_2+1,-2\ell_1+2\ell_2)$, then the sum obtained by a term-wise application of $\partial^\ell/(\partial x_1^{\ell_1}\partial x_2^{\ell_2})$ in (110) is absolutely convergent, uniformly with respect to $(x_1,x_2)\in\mathbb{R}\times(\mathbb{R}/N\mathbb{Z})$, and defines a continuous function of (x_1,x_2). (The continuity along the line $x_2=0$ holds since the bound in (114) tends to 0 as $|s|\to0$.) In particular, if $m>2\ell+1$ and $S_{\infty,0,\ell}(h)<\infty$, then it follows that $F_{c,\theta}\in\mathrm{C}^\ell$ and that $(\partial^{\ell_1+\ell_2}/(\partial x_1^{\ell_1}\partial x_2^{\ell_2}))F_{c,\theta}$ may be computed by term-wise differentiation in the sum in (110), for any $\ell_1,\ell_2\geqslant0$ with $\ell_1+\ell_2\leqslant\ell$.

We now turn to the proof of the bound (112). Using $S_{\infty,a,\ell_2}(h)<\infty$, $y|\cot\theta|\leqslant1$ and (114), we see that (113) is $\ll_{m,\ell}B(x_1,s)$, where

$$\begin{aligned} B(x_1,s)&:=\|f\|_{\mathrm{C}_0^{3m+\ell}}S_{\infty,a,\ell_2}(h)\|\boldsymbol{\eta}\|^{-m}\bigg(\Big(x_1-\frac{\sin 2\theta}{2yc^2s^2}\Big)^2+\Big(\frac{\sin^2\theta}{yc^2s^2}\Big)^2+1\bigg)^{-\frac{m}{2}} \\ &\quad\times\Big(\frac{\sin^2\theta}{yc^2s^2}\Big)^{\frac{m}{2}-\ell_1}|s|^{-\ell_2}\sum_{\gamma=0}^{\ell_2}|\sin\theta|^{\gamma-\ell_2}|s|^{-\gamma+a}(1+|s|)^{-a}. \end{aligned} \tag{115}$$

This bound is also valid when $\ell_2 = 0$. Next, using the fact that $\int_{\mathbb{R}} (u^2 + A)^{-m/2}\, du \ll A^{(1-m)/2}$ for all $A \geqslant 1$, we have $\int_{\mathbb{R}} B(x_1, s)\, dx_1 \ll_{m,\ell} B_1(s)$, where

$$B_1(s) := \|f\|_{\mathrm{C}_0^{3m+\ell}} S_{\infty,a,\ell_2}(h) \|\boldsymbol{\eta}\|^{-m} \left(1 + \left(\frac{\sin^2\theta}{yc^2s^2}\right)^2\right)^{\frac{1-m}{2}} \left(\frac{\sin^2\theta}{yc^2s^2}\right)^{\frac{m}{2}-\ell_1} |s|^{-\ell_2}$$

$$\times \sum_{\gamma=0}^{\ell_2} |\sin\theta|^{\gamma-\ell_2} |s|^{-\gamma+a} (1+|s|)^{-a}. \tag{116}$$

It follows that the left-hand side of (112), after squaring, is

$$\ll_{m,\ell} \int_{\mathbb{R}/N\mathbb{Z}} \left(\int_{\mathbb{R}} \sum_{\substack{s \in x_2 + N\mathbb{Z} \\ s<0}} B(x_1, s)\, dx_1 \right)^2 dx_2 \ll_{m,\ell} \int_{\mathbb{R}/N\mathbb{Z}} \left(\sum_{\substack{s \in x_2 + N\mathbb{Z} \\ s<0}} B_1(s) \right)^2 dx_2$$

$$\ll_{\varepsilon} \int_{\mathbb{R}/N\mathbb{Z}} \sum_{\substack{s \in x_2 + N\mathbb{Z} \\ s<0}} B_1(s)^2 (1+|s|)^{1+\varepsilon}\, dx_2 = \int_{-\infty}^{0} B_1(s)^2 (1+|s|)^{1+\varepsilon}\, ds$$

$$= \|f\|^2_{\mathrm{C}_0^{3m+\ell}} S_{\infty,a,\ell_2}(h)^2 \|\boldsymbol{\eta}\|^{-2m} \sum_{\gamma=0}^{\ell_2} |\sin\theta|^{2(\gamma-\ell_2)}$$

$$\times \int_0^{\infty} \left(1 + \left(\frac{\sin^2\theta}{yc^2s^2}\right)^2\right)^{1-m} \left(\frac{\sin^2\theta}{yc^2s^2}\right)^{m-2\ell_1} s^{-2\ell_2-2\gamma+2a} (s+1)^{1+\varepsilon-2a}\, ds. \tag{117}$$

Using $m > 2\ell + \frac{3}{2} > 2\ell + 1 + \frac{\varepsilon}{2}$, we find that the integral in the last line of (117) is

$$\ll_{m,\ell,\varepsilon} \begin{cases} \left(\frac{|\sin\theta|}{\sqrt{y}c}\right)^{2-2\ell_2-2\gamma+\varepsilon} & \text{if } \frac{|\sin\theta|}{\sqrt{y}c} \geqslant 1 \\ \left(\frac{|\sin\theta|}{\sqrt{y}c}\right)^{-2\ell_2-2\gamma+2a+1} + \left(\frac{|\sin\theta|}{\sqrt{y}c}\right)^{2m-4\ell_1} & \text{if } \frac{|\sin\theta|}{\sqrt{y}c} \leqslant 1. \end{cases}$$

Carrying out the addition over γ, we obtain the bound in (112). □

Note that Lemma 8.4 also applies to give a bound on $\|\partial_{x_1}^{\ell_1} \partial_{x_2}^{\ell_2} F_{c,\theta}\|_{\mathrm{L}^1}$, since $\|F\|_{\mathrm{L}^1} \leqslant \sqrt{N}\|F\|_{\mathrm{L}^{1,2}}$ for any function F on $\mathbb{R} \times (\mathbb{R}/N\mathbb{Z})$, by Cauchy–Schwarz. However, in the case $c\sqrt{y} \leqslant |\sin\theta|$, we need to get rid of the ε-power in (112). Thus, we prove.

LEMMA 8.5. *For any integers $\ell_1 \geqslant 0$ and $m > 2\ell_1 + 1$, for any $f \in \mathrm{C}_0^{3m+\ell_1}(X)$ and $h \in \mathrm{C}^{\ell_1}(\mathbb{R})$ with $S_{\infty,0,0}(h) < \infty$, if $0 < c\sqrt{y} \leqslant |\sin\theta|$, then*

$$\int_{\mathbb{R}/N\mathbb{Z}} \int_{\mathbb{R}} \left| \frac{\partial^{\ell_1}}{\partial x_1^{\ell_1}} F_{c,\theta}(x_1, x_2) \right| dx_1\, dx_2 \ll_{m,\ell_1} \|f\|_{\mathrm{C}_0^{3m+\ell_1}} S_{\infty,0,0}(h) \|\boldsymbol{\eta}\|^{-m} \frac{|\sin\theta|}{c\sqrt{y}}. \tag{118}$$

Proof. Following the proof of Lemma 8.4, we see that the left-hand side of (118) is $\ll \int_{-\infty}^{0} B_1(s)\, ds$, where $B_1(s)$ is given by (116) (with $a = \ell_2 = 0$). This integral is bounded by a direct computation, and we obtain the bound in (118). □

We are now ready to complete the proof of Proposition 8.1. Take $m \geqslant \max(8, 2k+1)$, $2 < a < 3$ and $\varepsilon, \varepsilon' \in (0, \frac{1}{2})$; also take $f \in \mathrm{C}_0^{3m+3}(X)$ and $h \in \mathrm{C}^1(\mathbb{R})$ with $S_{\infty,a,1}(h) < \infty$. Let

$\boldsymbol{\xi}_2 \in \mathbb{R}^k$ and $0 < y \leqslant 1$ be given. By Lemmata 8.4 and 8.5, we have for every $c \in \mathbb{Z}^+$ and every θ with $y|\cot\theta| \leqslant 1$ (and using the fact that $m - a - \frac{9}{2} > 0$),

$$\|F_{c,\theta}\|_{\mathrm{L}^1} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^1} \ll_{m,\varepsilon'} \|f\|_{\mathrm{C}_0^{3m+2}} S_{\infty,a,0}(h)\|\boldsymbol{\eta}\|^{-m} \begin{cases} \frac{|\sin\theta|}{c\sqrt{y}} & \text{if } c\sqrt{y} \leqslant |\sin\theta| \\ \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{\frac{1}{2}+a} & \text{if } |\sin\theta| \leqslant c\sqrt{y}. \end{cases}$$

Therefore,

$$\int_{I_y} \left(\|F_{c,\theta}\|_{\mathrm{L}^1} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^1}\right) \frac{y\,d\theta}{\sin^2\theta}$$

$$\ll_{m,a} \|f\|_{\mathrm{C}_0^{3m+2}} S_{\infty,a,0}(h)\|\boldsymbol{\eta}\|^{-m} y \begin{cases} (c\sqrt{y})^{-1}(1 + \log((c\sqrt{y})^{-1})) & \text{if } c\sqrt{y} \leqslant 1 \\ (c\sqrt{y})^{-\frac{1}{2}-a} & \text{if } c\sqrt{y} \geqslant 1. \end{cases} \tag{119}$$

Lemma 8.4 also gives (again using $m - a - \frac{9}{2} > 0$, and also using the fact that if $c\sqrt{y} \geqslant 1$, then $|\sin\theta|^{-1}\left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{m-a-\frac{5}{2}} \leqslant |\sin\theta|^{-1}\frac{|\sin\theta|}{c\sqrt{y}} \leqslant 1$)

$$\left(\|F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^{1,2}}\right)^{\frac{1}{2}-\varepsilon} \left(\|\partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 \partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}}\right)^{\frac{1}{2}+\varepsilon}$$

$$\ll_{m,\varepsilon'} \|f\|_{\mathrm{C}_0^{3m+3}} S_{\infty,a,1}(h)\|\boldsymbol{\eta}\|^{-m} \begin{cases} |\sin\theta|^{-\frac{1}{2}-\varepsilon}\left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{\frac{1}{2}+\varepsilon'-\varepsilon} & \text{if } c\sqrt{y} \leqslant |\sin\theta| \\ |\sin\theta|^{-\frac{1}{2}-\varepsilon}\left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{a-\varepsilon} & \text{if } |\sin\theta| \leqslant c\sqrt{y} \leqslant 1 \\ \left(\frac{|\sin\theta|}{c\sqrt{y}}\right)^{a-\frac{1}{2}-2\varepsilon} & \text{if } c\sqrt{y} \geqslant 1. \end{cases} \tag{120}$$

This leads to (using $a > 2 > \frac{3}{2} + \varepsilon$)

$$\int_{I_y} \left(\|F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 F_{c,\theta}\|_{\mathrm{L}^{1,2}}\right)^{\frac{1}{2}-\varepsilon} \left(\|\partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}} + \|\partial_{x_1}^2 \partial_{x_2} F_{c,\theta}\|_{\mathrm{L}^{1,2}}\right)^{\frac{1}{2}+\varepsilon} \frac{y\,d\theta}{\sin^2\theta}$$

$$\ll_{m,a,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+3}} S_{\infty,a,1}(h)\|\boldsymbol{\eta}\|^{-m} y \begin{cases} (c\sqrt{y})^{-\frac{3}{2}-\varepsilon} & \text{if } c\sqrt{y} \leqslant 1 \\ (c\sqrt{y})^{-a+\frac{1}{2}+2\varepsilon} & \text{if } c\sqrt{y} \geqslant 1. \end{cases} \tag{121}$$

Let us now also assume $\varepsilon < \frac{a-2}{3}$. Then $-a + \frac{1}{2} + 2\varepsilon < -\frac{3}{2} - \varepsilon$, and using (111), (119) and (121), it follows that the expression in (109) is

$$\ll_{m,a,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+3}} S_{\infty,a,1}(h)\|\boldsymbol{\eta}\|^{-m} \sum_{c=1}^{\infty} \Bigg\{ y^{\frac{3}{4}-\frac{a}{2}} c^{-1} (y^{-\frac{1}{2}} + c)^{\frac{1}{2}-a} \left(1 + \log^+\left(\frac{1}{c\sqrt{y}}\right)\right)$$

$$\times \sum_{\ell\in\mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{q}\boldsymbol{\xi}_2 + \ell\rfloor)}{1+\ell^2} + y^{\frac{1}{4}-\frac{\varepsilon}{2}} c^{-1-\frac{\varepsilon}{2}} \Bigg\}. \tag{122}$$

We now need the following modification of Lemma 7.5.

LEMMA 8.6. *Fix $\beta > 1$. Then for any $\alpha \in \mathbb{R}$ and $X \geqslant 1$ we have*

$$\sum_{c=1}^{\infty} c^{-1}(X+c)^{-\beta} \left(1 + \log^+\left(\frac{X}{c}\right)\right) \sum_{k\in\mathbb{Z}} \frac{(c,k)}{1+|k-c\alpha|^2} \tag{123}$$

$$\ll_\beta X^{1-\beta} \sum_{j=1}^{\infty} \min\left(\frac{1}{j^2}, \frac{1}{Xj\langle j\alpha\rangle}\right)\left(1 + \log^+\left(\frac{X\langle j\alpha\rangle}{j}\right)\right).$$

Proof. The proof of [38, Lemma 8.2] carries over with easy modifications. The main new technicality is to verify the bound

$$\sum_{d=1}^{\infty} \frac{(X+\ell d)^{-\beta}}{1+(d\langle \ell\alpha\rangle)^2}\left(1+\log^+\left(\frac{X}{\ell d}\right)\right)$$

$$\ll_\beta \begin{cases} X^{1-\beta}\ell^{-1} & \text{if } 1 \leqslant X/\ell \leqslant \langle \ell\alpha\rangle^{-1} \\ X^{-\beta}\langle \ell\alpha\rangle^{-1}\left(1+\log\left(\frac{X\langle \ell\alpha\rangle}{\ell}\right)\right) & \text{if } \langle \ell\alpha\rangle^{-1} < X/\ell \\ \ell^{-\beta} & \text{if } X/\ell < 1, \end{cases}$$

valid for all $d, \ell \in \mathbb{Z}^+$. □

Using Lemma 8.6 (with $\beta = a - \frac{1}{2}$), it follows that (122), and hence also (109), is

$$\ll_{a,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+3}} S_{\infty,a,1}(h) \|\boldsymbol{\eta}\|^{-m} \left\{ \sum_{j=1}^{\infty} \min\left(\frac{1}{j^2}, \frac{\sqrt{y}}{j\langle j\boldsymbol{q}\boldsymbol{\xi}_2\rangle}\right)\left(1+\log^+\left(\frac{\langle j\boldsymbol{q}\boldsymbol{\xi}_2\rangle}{j\sqrt{y}}\right)\right) + y^{\frac{1}{4}-\frac{\varepsilon}{2}} \right\}.$$

(We replaced '$N\boldsymbol{q}\boldsymbol{\xi}_2$' by $\boldsymbol{q}\boldsymbol{\xi}_2$ through the same type of estimate as in (95).) Adding the last bound over $R \in \overline{\Gamma}'/\Gamma'$ and $\boldsymbol{\eta} \in B_k$, using $\sum_{\boldsymbol{\eta}\in\mathbb{Z}^{2k}\setminus\{\mathbf{0}\}} \|\boldsymbol{\eta}\|^{-m} < \infty$ and $\sum_{\boldsymbol{r}\in\mathbb{Z}^k} \|\binom{\boldsymbol{q}}{\boldsymbol{r}}\|^{-m} \ll \|\boldsymbol{q}\|^{k-m}$ for every $\boldsymbol{q} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, and noticing that a and ε can be taken arbitrarily near 2 and 0, respectively, we obtain the bound in Proposition 8.1. This completes the proof of Proposition 8.1, and also of Theorem 1.2.

Remark 11. We now explain why we had to use Lemma 6.4 in place of Lemma 6.3 in the above proof of Proposition 8.1. One can prove a bound for the L^1-norm of $\partial_{x_1}^{\ell_1}\partial_{x_2}^{\ell_2} F_{c,\theta}$ which is very similar to the bound in Lemma 8.4, and in the case $c\sqrt{y} \leqslant 1$ this leads to a bound

$$\int_{\substack{0<\theta<\pi \\ (y|\cot\theta|\leqslant 1)}} \|\partial_{x_1}^{\ell_1}\partial_{x_2}^{\ell_2} F_{c,\theta}\|_{\mathrm{L}^1} \frac{y\,d\theta}{\sin^2\theta} \ll y(c\sqrt{y})^{-1-\ell_2}.$$

Multiplying this with $\sigma(c)^{3/2}\sqrt{c}$ and adding over c (cf. (109) and (111)) gives (if $\ell_2 > \frac{1}{2}$) a bound $y^{(1-\ell_2)/2}$, which is insufficient. Indeed, Lemma 6.3 requires us to take ℓ_2 as large as 2. Using instead the $\mathrm{L}^{1,2}$-norm and Lemma 6.4 means that we can effectively take ℓ_2 to be as small as $\frac{1}{2}+\varepsilon$, leading to the final bound $y^{\frac{1}{4}-\frac{\varepsilon}{2}}$. (One could sharpen Lemma 6.3 to a bound of the same style as in Lemma 6.4 but only involving the L^1-norm; this would allow us to use '$\ell_2 = 1+\varepsilon$'; however, this would still not be sufficient.)

8.2. *The case* $\boldsymbol{\xi}_2 = \mathbf{0}$

The treatment in this case is quite a bit easier than that for $\boldsymbol{\xi}_1 = \mathbf{0}$. We prove the following bound.

Proposition 8.7. *Let $k \geqslant 2$. Fix a real number $\varepsilon > 0$ and an integer $m \geqslant \max(7, 2k+1)$. For any $f \in \mathrm{C}_0^{3m+2}(X)$, $h \in \mathrm{C}^2(\mathbb{R})$ with $S_{1,0,2}(h) < \infty$, $\boldsymbol{\xi}_1 \in \mathbb{R}^k$ and $0 < y \leqslant 1$, we have*

$$\sum_{\boldsymbol{\eta}\in B_k} \sum_{R\in\overline{\Gamma}'/\Gamma'} \sum_{T\in[R]} e\left({}^{\mathrm{t}}T\boldsymbol{\eta}\begin{pmatrix}\boldsymbol{\xi}_1 \\ \mathbf{0}\end{pmatrix}\right) \int_{\mathbb{R}} \widehat{f}_R\left(T\begin{pmatrix}\sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y}\end{pmatrix}, \boldsymbol{\eta}\right) h(x)\,dx$$

$$\ll_{m,\varepsilon} \|f\|_{\mathrm{C}_0^{3m+2}} S_{1,0,2}(h)\left(\widetilde{\delta}_{m-k,\boldsymbol{\xi}_1}(y^{-\frac{1}{2}}) + y^{\frac{1}{4}-\varepsilon}\right). \tag{124}$$

Note that Theorem 1.3 follows from Proposition 8.7 together with Proposition 7.7 and the relations (49) and (50).

Proof. The beginning of the proof of Proposition 8.1 carries over without changes; the first difference is that in place of (101) we get:

$$\sum_{\substack{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\in[R] \\ c>0}} e((a\boldsymbol{q}+c\boldsymbol{r})\boldsymbol{\xi}_1)\int_0^\pi \widehat{f}_R\Big(\frac{a}{c}-\frac{\sin 2\theta}{2c^2y},\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)h\Big(-\frac{d}{c}+y\cot\theta\Big)\frac{y\,d\theta}{\sin^2\theta}. \tag{125}$$

Interchanging the roles of a and d in the summation, we see that (125) can be alternatively expressed as:

$$\sum_{\substack{c\equiv c_0 \bmod N \\ c>0}} e(c\boldsymbol{r}\boldsymbol{\xi}_1)\int_0^\pi \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(\widetilde{1})} e(d\boldsymbol{q}\boldsymbol{\xi}_1)\widehat{f}_R\Big(\frac{d}{c}-\frac{\sin 2\theta}{2c^2y},\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\nu\Big(-\frac{a}{c}+y\cot\theta\Big)\frac{y\,d\theta}{\sin^2\theta}, \tag{126}$$

where $\nu(x)=\sum_{n\in\mathbb{Z}}h(x+nN)$ (a function on $\mathbb{R}/N\mathbb{Z}$) and where $\sum^{(\widetilde{1})}$ is the same as $\sum^{(1)}$ (cf. p. 16) but using $\widetilde{R}:=\left(\begin{smallmatrix} d_0 & b_0 \\ c_0 & a_0 \end{smallmatrix}\right)$ in place of $R=\left(\begin{smallmatrix} a_0 & b_0 \\ c_0 & d_0 \end{smallmatrix}\right)$. Now by Lemma 6.3 we have, for any c and θ appearing above:

$$\begin{aligned}
&\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}^{(\widetilde{1})} e(d\boldsymbol{q}\boldsymbol{\xi}_1)\widehat{f}_R\Big(\frac{d}{c}-\frac{\sin 2\theta}{2c^2y},\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\nu\Big(-\frac{a}{c}+y\cot\theta\Big) \\
&\ll\Big\{\int_{\mathbb{R}}\Big|\widehat{f}_R\Big(u,\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\Big|\,du+\int_{\mathbb{R}}\Big|\frac{\partial^2}{\partial u^2}\widehat{f}_R\Big(u,\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\Big|\,du\Big\} \\
&\quad\times\Big(\|\nu\|_{\mathrm{L}^1(\mathbb{R}/N\mathbb{Z})}\sum_{\ell\in\mathbb{Z}}\frac{(c,\lfloor cN\boldsymbol{q}\boldsymbol{\xi}_1+\ell\rfloor)}{1+\ell^2}+\|\nu''\|_{\mathrm{L}^1(\mathbb{R}/N\mathbb{Z})}\sigma(c)\sqrt{c}\Big).
\end{aligned}$$

Using (98) and writing $v=\frac{\sin^2\theta}{c^2y}$, for any $\ell\geqslant 0$, we get:

$$\begin{aligned}
\int_{\mathbb{R}}\Big|\frac{\partial^\ell}{\partial u^\ell}\widehat{f}_R\Big(u,\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\Big|\,du &\ll\|f\|_{\mathrm{C}_0^{3m+\ell}}\|\boldsymbol{\eta}\|^{-m}v^{-\ell+\frac{m}{2}}\int_{\mathbb{R}}(u^2+v^2+1)^{-\frac{m}{2}}\,du \\
&\ll\|f\|_{\mathrm{C}_0^{3m+\ell}}\|\boldsymbol{\eta}\|^{-m}\min\left(\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{-2\ell+m},\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{2-2\ell-m}\right),
\end{aligned}$$

and thus

$$\sum_{\ell\in\{0,2\}}\int_{\mathbb{R}}\Big|\frac{\partial^\ell}{\partial u^\ell}\widehat{f}_R\Big(u,\frac{\sin^2\theta}{c^2y},\theta;\boldsymbol{\eta}\Big)\Big|\,du\ll\|f\|_{\mathrm{C}_0^{3m+2}}\|\boldsymbol{\eta}\|^{-m}\begin{cases}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m-4} & \text{if } \frac{|\sin\theta|}{c\sqrt{y}}\leqslant 1\\ \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{2-m} & \text{if } \frac{|\sin\theta|}{c\sqrt{y}}\geqslant 1.\end{cases}$$

Using also

$$\begin{aligned}
\int_0^\pi\left\{\begin{array}{ll}\Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{m-4} & \text{if } \frac{|\sin\theta|}{c\sqrt{y}}\leqslant 1\\ \Big(\frac{|\sin\theta|}{c\sqrt{y}}\Big)^{2-m} & \text{if } \frac{|\sin\theta|}{c\sqrt{y}}\geqslant 1\end{array}\right\}\frac{y\,d\theta}{\sin^2\theta} &\ll y\min\big((c\sqrt{y})^{-1},(c\sqrt{y})^{4-m}\big) \\
&\ll y^{3-\frac{m}{2}}c^{-1}(y^{-\frac{1}{2}}+c)^{5-m},
\end{aligned}$$

we conclude that (125) is

$$\ll \|f\|_{\mathrm{C}_0^{3m+2}} \|\boldsymbol{\eta}\|^{-m} y^{3-\frac{m}{2}} \Big\{ S_{1,0,0}(h) \sum_{c=1}^{\infty} c^{-1} (y^{-\frac{1}{2}} + c)^{5-m} \sum_{\ell \in \mathbb{Z}} \frac{(c, \lfloor cN\boldsymbol{q}\boldsymbol{\xi}_1 + \ell \rfloor)}{1+\ell^2}$$

$$+ S_{1,0,2}(h) \sum_{c=1}^{\infty} (y^{-\frac{1}{2}} + c)^{5-m} \frac{\sigma(c)}{\sqrt{c}} \Big\},$$

and by Lemmas 7.5 and 7.6 (using $m \geqslant 7$), this is

$$\ll \|f\|_{\mathrm{C}_0^{3m+2}} S_{1,0,2}(h) \|\boldsymbol{\eta}\|^{-m} \left\{ \sum_{j=1}^{\infty} \frac{1}{j^2 + y^{-1/2} j \langle j\boldsymbol{q}\boldsymbol{\xi}_1 \rangle} + y^{\frac{1}{4}-\varepsilon} \right\}.$$

Adding this bound over $R \in \overline{\Gamma}'/\Gamma'$ and $\boldsymbol{\eta} \in B_k$, using $\sum_{\boldsymbol{r} \in \mathbb{Z}^k} \|\binom{\boldsymbol{q}}{\boldsymbol{r}}\|^{-m} \ll \|\boldsymbol{q}\|^{k-m}$ for every $\boldsymbol{q} \in \mathbb{Z}^k \setminus \{\boldsymbol{0}\}$, and $\sum_{\boldsymbol{\eta} \in \mathbb{Z}^{2k} \setminus \{\boldsymbol{0}\}} \|\boldsymbol{\eta}\|^{-m} < \infty$ (these hold since $m > 2k$), we obtain the bound in Proposition 8.7. This also completes the proof of Theorem 1.3. □

9. *Application to a quantitative Oppenheim result*

Our goal in this section is to prove Theorem 1.4, by making Marklof's approach from [**24**] effective. This will involve an application of Theorem 1.2 at a key step.

9.1. *Set-up*

Let $\mathbb{H} = \{\tau = u + iv \in \mathbb{C} : v > 0\}$, the Poincaré upper half plane. Let k be a positive integer and let $\mathcal{S}(\mathbb{R}^k)$ be the Schwartz space of functions on $\mathbb{R}^k$ which, together with their derivatives, decrease rapidly at infinity. A central role in the approach of [**24**] is played by the *Jacobi theta sum*, $\Theta_f(\tau, \phi; \boldsymbol{\xi})$. It is defined by the following formula, for any $f \in \mathcal{S}(\mathbb{R}^k)$, $\tau = u + iv \in \mathbb{H}$, $\phi \in \mathbb{R}$ and $\boldsymbol{\xi} = \binom{\boldsymbol{\xi}_1}{\boldsymbol{\xi}_2} \in \mathbb{R}^{2k}$:

$$\Theta_f(\tau, \phi; \boldsymbol{\xi}) = v^{k/4} \sum_{\boldsymbol{m} \in \mathbb{Z}^k} f_\phi((\boldsymbol{m} - \boldsymbol{\xi}_2) v^{1/2})\, e(\tfrac{1}{2}\|\boldsymbol{m} - \boldsymbol{\xi}_2\|^2 u + \boldsymbol{m} \cdot \boldsymbol{\xi}_1), \tag{127}$$

where, for ϕ in any interval $\nu\pi < \phi < (\nu+1)\pi$ ($\nu \in \mathbb{Z}$), f_ϕ is given by the formula

$$f_\phi(\boldsymbol{w}) = \int_{\mathbb{R}^k} G_\phi(\boldsymbol{w}, \boldsymbol{w}') f(\boldsymbol{w}')\, d\boldsymbol{w}',$$

with the integral kernel

$$G_\phi(\boldsymbol{w}, \boldsymbol{w}') = e\left(-\frac{k(2\nu+1)}{8}\right) |\sin\phi|^{-k/2} e\left[\frac{\frac{1}{2}(\|\boldsymbol{w}\|^2 + \|\boldsymbol{w}'\|^2)\cos\phi - \boldsymbol{w} \cdot \boldsymbol{w}'}{\sin\phi}\right], \tag{128}$$

while for $\phi = \nu\pi$ ($\nu \in \mathbb{Z}$) we have $f_\phi(\boldsymbol{w}) = e(-\frac{k\nu}{4}) f((-1)^\nu \boldsymbol{w})$. The operators $U^\phi : f \mapsto f_\phi$ form a 1-parameter group of unitary operators on $\mathrm{L}^2(\mathbb{R}^k)$; in particular, $U^\phi \circ U^{\phi'} = U^{\phi+\phi'}$ for any $\phi, \phi' \in \mathbb{R}$ (see [**24**, Sections 3–4]).

For any $f, g \in \mathcal{S}(\mathbb{R}^k)$, the product $\Theta_f(\tau, \phi; \boldsymbol{\xi})\overline{\Theta_g(\tau, \phi; \boldsymbol{\xi})}$ depends only on $\phi \bmod 2\pi$ and may thus be viewed as a function on $G = \mathrm{SL}(2, \mathbb{R}) \ltimes \mathbb{R}^{2k}$ through the Iwasawa parameterization (cf. (40))

$$(\tau, \phi, \boldsymbol{\xi}) \mapsto \left(\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{v} & 0 \\ 0 & 1/\sqrt{v} \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \boldsymbol{\xi} \right), \quad \text{where } \tau = u + iv.$$

By [24, Proposition 4.9], this function $\Theta_f\overline{\Theta}_g \in \mathrm{C}^\infty(G)$ is in fact left Γ^k invariant, where

$$\Gamma^k = \left\{ \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} ab\boldsymbol{s} \\ cd\boldsymbol{s} \end{pmatrix} + \boldsymbol{m} \right) \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}),\ \boldsymbol{m} \in \mathbb{Z}^{2k} \right\} \tag{129}$$

with $\boldsymbol{s} := {}^{\mathrm{t}}(1/2,\dots,1/2) \in \mathbb{R}^k$. The group Γ^k is a finite index subgroup of $\mathrm{SL}(2,\mathbb{Z}) \ltimes (\frac{1}{2}\mathbb{Z})^{2k}$, and contains $\Gamma_\theta \ltimes \mathbb{Z}^{2k}$ as an index 3 subgroup, where Γ_θ is the theta group, that is,

$$\Gamma_\theta = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \ : \ ab \equiv cd \equiv 0 \bmod 2 \right\}.$$

See [24, Lemmata 4.11 and 4.12].

For the proof of Theorem 1.4, we will eventually specialize to $k = 2$: The starting point for the method developed in [24] is the following identity[†], valid for any $f, g \in \mathcal{S}(\mathbb{R}^2)$, $h \in \mathrm{L}^1(\mathbb{R})$, $T > 0$ and $\boldsymbol{\xi}_2 \in \mathbb{R}^2$:

$$\begin{aligned} &\int_{\mathbb{R}} \Theta_f\left(u + T^{-2}i, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) \overline{\Theta_g\left(u + T^{-2}i, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right)}\, h(u)\, du \\ &\quad = \frac{1}{T^2} \sum_{\boldsymbol{m}_1 \in \mathbb{Z}^2} \sum_{\boldsymbol{m}_2 \in \mathbb{Z}^2} f\big(T^{-1}(\boldsymbol{m}_1 - \boldsymbol{\xi}_2)\big)\, \overline{g(T^{-1}(\boldsymbol{m}_2 - \boldsymbol{\xi}_2))}\, \widehat{h}\left(-\tfrac{1}{2}\, Q\begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \end{pmatrix}\right), \end{aligned} \tag{130}$$

where Q is the inhomogeneous quadratic form on $\mathbb{R}^4$ given by (9) with $\boldsymbol{\xi}_2 = \binom{\alpha}{\beta} \in \mathbb{R}^2$, that is,

$$Q\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix} = \|\boldsymbol{x}_1 - \boldsymbol{\xi}_2\|^2 - \|\boldsymbol{x}_2 - \boldsymbol{\xi}_2\|^2, \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^2. \tag{131}$$

The formula (130) follows by replacing Θ_f and Θ_g by their defining sums (cf. (127)) and changing the order of summation and integration.

The key step in [24] is then to determine the limit of the left-hand side of (130) as $T \to \infty$, by using the invariance properties of the function $\Theta_f\overline{\Theta_g}$ and an equidistribution result as in Theorem 1.1 above (with $\boldsymbol{\xi}_1 = \mathbf{0}$); this is where we will apply our effective result, Theorem 1.2, instead. A central difficulty in [24] comes from the fact that the theta functions Θ_f, Θ_g are unbounded; thus one needs to truncate the function $\Theta_f\overline{\Theta_g}$ in the cusp before the equidistribution result can be applied, and then bound the error caused by the truncation. In fact it turns out that one picks up an explicit extra contribution from the part of the integral in (130) over a tiny interval $|u| \ll T^{-(1+\varepsilon)}$, whereas the error caused by the truncation for the remaining part of the integral can be proved to be appropriately small, provided that $\boldsymbol{\xi}_2$ is Diophantine. The treatment of these matters in [24] is already in principle effective, and so our work concerning the truncation error will essentially only consist in keeping more explicit track on how the bounds in [24] depends on various parameters (see, in particular, Proposition 9.6). Also, for the application of Theorem 1.2, we require precise bounds on derivatives of the function $\Theta_f\overline{\Theta_g}$; this is worked out in Lemma 9.2.

9.2. *Bounds for the derivatives of* $\Theta_f\overline{\Theta_g}$

Although we will eventually specialize to $k = 2$, we will consider a general $k \in \mathbb{Z}^+$ as long as this causes no extra work. We will use the same notation $S_{p,a,n}$ as introduced in Section 1 also

[†]See [24, Section 2.3], where the identity (130) appears in the special case when $f(\boldsymbol{x}) \equiv \psi_1(\|\boldsymbol{x}\|^2)$, $g(\boldsymbol{x}) \equiv \psi_2(\|\boldsymbol{x}\|^2)$ and using a slightly different notation than in (130). Note that we write $\widehat{h}(s) = \int_{\mathbb{R}} h(u)e(-su)\, du$ in (130), in line with previous definitions in our paper, whereas a different normalization of $\widehat{h}$ is used in [24, p. 423(top)].

for the corresponding weighted Sobolev norm of a function $f \in \mathrm{C}^n(\mathbb{R}^k)$ with $k \geqslant 2$; namely

$$S_{p,a,n}(f) = \sum_{|\gamma| \leqslant n} \|(1+\|\boldsymbol{x}\|)^a \, \partial^\gamma f(\boldsymbol{x})\|_{\mathrm{L}^p}. \tag{132}$$

Here, we use standard multi-index notation, that is, γ runs through k-tuples of nonnegative integers, $|\gamma| = \gamma_1 + \cdots + \gamma_k$ and $\partial^\gamma = \partial_{x_1}^{\gamma_1} \cdots \partial_{x_k}^{\gamma_k}$. Note that if $a(\geqslant 0)$ is an integer then (see, for example, [**10**, Chapter 8.1, Exercise 1]),

$$S_{p,a,n}(f) \asymp \sum_{|\beta| \leqslant n} \sum_{|\beta'| \leqslant a} \|\boldsymbol{x}^{\beta'} \partial^\beta f(\boldsymbol{x})\|_{\mathrm{L}^p} \asymp \sum_{|\beta| \leqslant n} \sum_{|\beta'| \leqslant a} \|\partial^\beta (\boldsymbol{x}^{\beta'} f(\boldsymbol{x}))\|_{\mathrm{L}^p}, \tag{133}$$

where the implied constant may depend on p, a, n, k. In particular, for $p = 2$, combining (133) with the Plancherel theorem, we have

$$S_{2,a,n}(f) \asymp S_{2,n,a}(\widehat{f}), \tag{134}$$

where $\widehat{f}(\boldsymbol{y}) = \int_{\mathbb{R}^k} f(\boldsymbol{x}) e(-\boldsymbol{x}\boldsymbol{y}) \, d\boldsymbol{x}$ is the Fourier transform of f. We also record the following basic Sobolev embedding inequality which we will use several times: For any $a \in \mathbb{R}_{\geqslant 0}$, $n \in \mathbb{Z}_{\geqslant 0}$ and $f \in \mathcal{S}(\mathbb{R}^k)$,

$$S_{\infty,a,n}(f) \ll S_{2,a,n+\lfloor k/2 \rfloor + 1}(f), \tag{135}$$

where the implied constant may depend on k, a, n. The proof of (135) is completely elementary: Set $m := \lfloor k/2 \rfloor + 1$; then $(1+\|\boldsymbol{x}\|)^{-m} \in \mathrm{L}^2$, and so by Fourier inversion and Cauchy–Schwarz,

$$\|f\|_{\mathrm{L}^\infty} \leqslant \|\widehat{f}\|_{\mathrm{L}^1} \leqslant \left\|(1+\|\boldsymbol{x}\|)^m \cdot \widehat{f}\right\|_{\mathrm{L}^2} \left\|(1+\|\boldsymbol{x}\|)^{-m}\right\|_{\mathrm{L}^2} \ll S_{2,m,0}(\widehat{f}) \ll S_{2,0,m}(f). \tag{136}$$

This proves (135) when $a = n = 0$. In the case of general a and n, one notes that $S_{\infty,a,n}(f) \ll \sum_{|\gamma| \leqslant n} \|(1+\|\boldsymbol{x}\|^2)^{a/2} \cdot \partial^\gamma f\|_{\mathrm{L}^\infty}$, and then applies (136) to $(1+\|\boldsymbol{x}\|^2)^{a/2} \cdot \partial^\gamma f$.

We will often work with the Sobolev norms $S_{2,a,a}$ on functions in $\mathrm{C}^a(\mathbb{R}^k)$, and we introduce the notation $\|\cdot\|_{\mathrm{L}^2_a}$ for these. Thus, for any integer $a \geqslant 0$ and $f \in \mathrm{C}^a(\mathbb{R}^k)$,

$$\|f\|_{\mathrm{L}^2_a} := S_{2,a,a}(f) = \sum_{|\beta| \leqslant a} \|(1+\|\boldsymbol{x}\|)^a \partial^\beta f(\boldsymbol{x})\|_{\mathrm{L}^2}. \tag{137}$$

Given $f \in \mathcal{S}(\mathbb{R}^k)$, we view $f_\phi(\boldsymbol{w})$ as a function on the space $\mathbb{R}^{k+1}$, given by the co-ordinates $(\boldsymbol{w}, \phi)$. Thus, $\partial^\beta f_\phi(\boldsymbol{w})$ for $\beta \in (\mathbb{Z}_{\geqslant 0})^{k+1}$ denotes $\partial_{w_1}^{\beta_1} \cdots \partial_{w_k}^{\beta_k} \partial_\phi^{\beta_{k+1}} f_\phi(\boldsymbol{w})$. The following lemma corresponds to [**24**, Lemma 4.3], but extended to arbitrary derivatives of f_ϕ and with the implied constant made more precise.

LEMMA 9.1. *Let $A \in \mathbb{Z}_{\geqslant 0}$, $\beta \in (\mathbb{Z}_{\geqslant 0})^{k+1}$ and $a \in \mathbb{Z}$, $a > A + \frac{k}{2} + 4|\beta|$. Then for any $f \in \mathcal{S}(\mathbb{R}^k)$, $\boldsymbol{w} \in \mathbb{R}^k$ and $\phi \in \mathbb{R}$,*

$$|\partial^\beta f_\phi(\boldsymbol{w})| \ll_{A,\beta} \|f\|_{\mathrm{L}^2_a} (1+\|\boldsymbol{w}\|)^{-A}.$$

Proof. For ϕ in any interval $\nu\pi + \frac{1}{100} < \phi < (\nu+1)\pi - \frac{1}{100}$, $\nu \in \mathbb{Z}$, we use

$$\partial^\beta f_\phi(\boldsymbol{w}) = \int_{\mathbb{R}^k} \big(\partial^\beta G_\phi(\boldsymbol{w}, \boldsymbol{w}')\big) f(\boldsymbol{w}') \, d\boldsymbol{w}',$$

with $G_\phi(\boldsymbol{w},\boldsymbol{w}')$ as in (128), and with ∂^β acting on the $k+1$ variables $w_1,\dots,w_k,\phi$. One proves by induction that

$$\partial^\beta G_\phi(\boldsymbol{w},\boldsymbol{w}') = G_\phi(\boldsymbol{w},\boldsymbol{w}')\frac{P_\beta(\boldsymbol{w},\boldsymbol{w}',\sin\phi,\cos\phi)}{(\sin\phi)^{2|\beta|}},$$

where P_β is a polynomial in $2k+2$ variables, with complex coefficients which only depend on k and β, and only containing terms $w_1^{\alpha_1}\cdots w_k^{\alpha_k}{w_1'}^{\alpha_{k+1}}\cdots {w_k'}^{\alpha_{2k}}(\sin\phi)^{\alpha_{2k+1}}(\cos\phi)^{\alpha_{2k+2}}$ with $\sum_1^{2k}\alpha_j \leqslant 2|\beta|$. Integrating by parts $n\geqslant 0$ times with respect to w_j' for some j, it follows that

$$\partial^\beta f_\phi(\boldsymbol{w}) = \frac{e\big(-\frac{1}{8}k(2\nu+1)\big)}{|\sin\phi|^{k/2}}\left(\frac{\sin\phi}{2\pi i w_j}\right)^n\int_{\mathbb{R}^k} K(\boldsymbol{w},\boldsymbol{w}',\phi)\,e\left[-\frac{\boldsymbol{w}\cdot\boldsymbol{w}'}{\sin\phi}\right]d\boldsymbol{w}'$$

where

$$K(\boldsymbol{w},\boldsymbol{w}',\phi) = \left(\frac{\partial}{\partial w_j'}\right)^n\left(e\left[\frac{\frac{1}{2}(\|\boldsymbol{w}\|^2+\|\boldsymbol{w}'\|^2)\cos\phi}{\sin\phi}\right]\frac{P_\beta(\boldsymbol{w},\boldsymbol{w}',\sin\phi,\cos\phi)}{(\sin\phi)^{2|\beta|}}f(\boldsymbol{w}')\right)$$

and so

$$|K(\boldsymbol{w},\boldsymbol{w}',\phi)| \ll_{\beta,n} (1+\|\boldsymbol{w}\|)^{2|\beta|}\sum_{\ell=0}^n(1+\|\boldsymbol{w}'\|)^{2|\beta|+\ell}\left|\left(\frac{\partial}{\partial w_j'}\right)^{n-\ell}f(\boldsymbol{w}')\right|$$

for all $\boldsymbol{w},\boldsymbol{w}'\in\mathbb{R}^k$ and $\phi\in(\nu\pi+\frac{1}{100},(\nu+1)\pi-\frac{1}{100})$. If $\|\boldsymbol{w}\|\geqslant 1$, then we apply the above with $n=A+2|\beta|$ and j being the index for which $|w_j|=\max(|w_1|,\dots,|w_k|)$; if $\|\boldsymbol{w}\|<1$, then we instead use $n=0$. The desired bound follows using the Cauchy–Schwarz inequality combined with the fact that $(1+\|\boldsymbol{w}'\|)^{-\frac{k+\varepsilon}{2}}\in \mathrm{L}^2(\mathbb{R}^k)$ for any $\varepsilon>0$.

To treat the remaining values of ϕ, we use the fact that $f_{\phi+\frac{\pi}{2}} = e^{-\frac{1}{4}\pi k i}U^\phi\widehat{f}$; hence by what we have already proved, $|\partial^\beta f_\phi(\boldsymbol{w})|\ll \|\widehat{f}\|_{\mathrm{L}^2_a}(1+\|\boldsymbol{w}\|)^{-A}$ for ϕ in any interval $(\nu-\frac{1}{2})\pi+\frac{1}{100}<\phi<(\nu+\frac{1}{2})\pi-\frac{1}{100}$, $\nu\in\mathbb{Z}$, and the desired bound follows using (134). □

Using Lemma 9.1, we now obtain bounds on arbitrary derivatives of the function $\Theta_f\overline{\Theta}_g\in \mathrm{C}^\infty(G)$. Recall that we write $\sum_{\mathrm{ord}(D)\leqslant m}$ to denote a sum over all monomials D of degree $\leqslant m$ in the fixed basis $X_1,\dots,X_{3+2k}$ of $\mathfrak{g}$ (cf. (18)).

LEMMA 9.2. *Let $f,g\in\mathcal{S}(\mathbb{R}^k)$. Let m and a be integers satisfying $m\geqslant 0$ and $a>\frac{3}{2}k+6m+1$. Then for any $(\tau,\phi,\boldsymbol{\xi})\in G$ with $v=\operatorname{Im}\tau\geqslant\frac{1}{2}$,*

$$\sum_{\mathrm{ord}(D)\leqslant m}\big|\big(D(\Theta_f\overline{\Theta}_g)\big)(\tau,\phi;\boldsymbol{\xi})\big| \ll_m \|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a}\,v^{m+\frac{1}{2}k}. \tag{138}$$

Next let A and a be integers satisfying $A\geqslant 1$ and $a>\frac{3}{2}k+2A$. Then for any $(\tau,\phi,\boldsymbol{\xi})\in G$ with $v=\operatorname{Im}\tau\geqslant\frac{1}{2}$,

$$\left|(\Theta_f\overline{\Theta}_g)(\tau,\phi;\boldsymbol{\xi}) - v^{k/2}\sum_{\boldsymbol{m}\in\mathbb{Z}^k} f_\phi((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})}\right| \ll_A \|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a}v^{-A}, \tag{139}$$

and if furthermore $\boldsymbol{\xi}_2\in\boldsymbol{n}+[-\frac{1}{2},\frac{1}{2}]^k$ and $\boldsymbol{n}\in\mathbb{Z}^k$, then

$$\left|(\Theta_f\overline{\Theta}_g)(\tau,\phi;\boldsymbol{\xi}) - v^{k/2}f_\phi((\boldsymbol{n}-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{n}-\boldsymbol{\xi}_2)v^{1/2})}\right| \ll_A \|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a}v^{-A}. \tag{140}$$

Proof. Recall that we write $\tau = u + iv$. We have

$$\begin{aligned}
\Theta_f\overline{\Theta}_g(\tau,\phi;\boldsymbol{\xi}) &= v^{k/2}\sum_{\boldsymbol{m}_1,\boldsymbol{m}_2\in\mathbb{Z}^k} f_\phi((\boldsymbol{m}_1-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{m}_2-\boldsymbol{\xi}_2)v^{1/2})}\\
&\quad\times e\big(\tfrac12(\|\boldsymbol{m}_1-\boldsymbol{\xi}_2\|^2-\|\boldsymbol{m}_2-\boldsymbol{\xi}_2\|^2)u+(\boldsymbol{m}_1-\boldsymbol{m}_2)\cdot\boldsymbol{\xi}_1\big)\\
&= v^{k/2}\sum_{\boldsymbol{m}_1,\boldsymbol{m}_2\in\mathbb{Z}^k} f_\phi((\boldsymbol{m}_1-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{m}_2-\boldsymbol{\xi}_2)v^{1/2})}\\
&\quad\times e\big(\tfrac12(\boldsymbol{m}_1-\boldsymbol{m}_2)((\boldsymbol{m}_1+\boldsymbol{m}_2-2\boldsymbol{\xi}_2)u+2\boldsymbol{\xi}_1)\big)\\
&= \sum_{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{Z}^k} v^{k/2}f_\phi((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{m}-\boldsymbol{m}'-\boldsymbol{\xi}_2)v^{1/2})}\,e\big(\boldsymbol{m}'\cdot((\boldsymbol{m}-\tfrac12\boldsymbol{m}'-\boldsymbol{\xi}_2)u+\boldsymbol{\xi}_1)\big)\\
&:= \sum_{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{Z}^k} F_{\boldsymbol{m},\boldsymbol{m}'}(\tau,\phi;\boldsymbol{\xi}),
\end{aligned}$$

say. Note that $F_{\boldsymbol{m},0}(\tau,\phi;\boldsymbol{\xi}) = v^{k/2}f_\phi((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})\overline{g_\phi((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})}$.

In the $(u+iv,\phi;\boldsymbol{\xi})$-co-ordinates with $\boldsymbol{\xi} = {}^{\mathrm{t}}(\xi_1,\dots,\xi_{2k})$, the Lie derivatives $X_1,\dots,X_{3+2k}$ are given by

$$X_1 = v(\cos 2\phi)\partial_u - v(\sin 2\phi)\partial_v - (\sin\phi)^2\partial_\phi;$$

$$X_2 = v(\cos 2\phi)\partial_u - v(\sin 2\phi)\partial_v + (\cos\phi)^2\partial_\phi;$$

$$X_3 = 2v(\sin 2\phi)\partial_u + 2v(\cos 2\phi)\partial_v + (\sin 2\phi)\partial_\phi;$$

$$\begin{cases} X_{3+\ell} = \left(\frac{v\cos\phi+u\sin\phi}{\sqrt{v}}\right)\partial_{\xi_\ell} + \left(\frac{\sin\phi}{\sqrt{v}}\right)\partial_{\xi_{k+\ell}} \\ X_{3+k+\ell} = \left(\frac{-v\sin\phi+u\cos\phi}{\sqrt{v}}\right)\partial_{\xi_\ell} + \left(\frac{\cos\phi}{\sqrt{v}}\right)\partial_{\xi_{k+\ell}} \end{cases} \qquad (\ell\in\{1,2,\dots,k\}).$$

(The formulae for X_1, X_2, X_3 are standard and may, for example, easily be derived using the formulae in the proof of [38, Lemma 6.1] and $X_2 - X_1 = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right) = \partial_\phi$. Regarding $X_{3+\ell}$ and $X_{3+k+\ell}$, cf. (39).) Using the automorphy of $\Theta_f\overline{\Theta}_g$, it is enough to prove (138) when $|u|\leqslant\frac12$ and $\|\boldsymbol{\xi}\|\ll 1$. Using the above formulae we then get, for any monomial D in $X_1,\dots,X_{3+2k}$ with d_1 factors in $\{X_1,X_2,X_3\}$ and d_2 factors in $\{X_4,\dots,X_{3+2k}\}$:

$$|DF_{\boldsymbol{m},\boldsymbol{m}'}(\tau,\phi;\boldsymbol{\xi})| \ll v^{d_1+\frac12 d_2}\sum_{|\alpha|\leqslant d_1}\sum_{|\beta|\leqslant d_2}\left|\partial_u^{\alpha_1}\partial_v^{\alpha_2}\partial_\phi^{\alpha_3}\partial_{\xi_1}^{\beta_1}\cdots\partial_{\xi_{2k}}^{\beta_{2k}}F_{\boldsymbol{m},\boldsymbol{m}'}(u+iv,\phi;\boldsymbol{\xi})\right|, \quad (141)$$

where α runs through multi-indices in $(\mathbb{Z}_{\geqslant 0})^3$ and β runs through multi-indices in $(\mathbb{Z}_{\geqslant 0})^{2k}$. Next, from the definition of $F_{\boldsymbol{m},\boldsymbol{m}'}(\tau,\phi;\boldsymbol{\xi})$, by a standard computation, we obtain

$$\begin{aligned}
&\left|\partial_u^{\alpha_1}\partial_v^{\alpha_2}\partial_\phi^{\alpha_3}\partial_{\xi_1}^{\beta_1}\cdots\partial_{\xi_{2k}}^{\beta_{2k}}F_{\boldsymbol{m},\boldsymbol{m}'}(u+iv,\phi;\boldsymbol{\xi})\right| \ll (1+\|\boldsymbol{m}\|+\|\boldsymbol{m}'\|)^{2|\alpha|+|\beta|}v^{(k+|\beta|)/2}\\
&\quad\times\sum_{|\beta'|+|\beta''|\leqslant|\alpha|+|\beta|}\left|(\partial^{\beta'}f_\phi)((\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2})(\partial^{\beta''}g_\phi)((\boldsymbol{m}-\boldsymbol{m}'-\boldsymbol{\xi}_2)v^{1/2})\right|,
\end{aligned}$$

where β' and β'' run through multi-indices in $(\mathbb{Z}_{\geqslant 0})^{k+1}$, with $\partial^{\beta'}$ and $\partial^{\beta''}$ having the same meaning as in Lemma 9.1. Applying now Lemma 9.1, with any fixed integers A and a subject to $A\geqslant 0$ and $a > A + \frac{k}{2} + 4(d_1+d_2)$, we get

$$\begin{aligned}
|DF_{\boldsymbol{m},\boldsymbol{m}'}(\tau,\phi;\boldsymbol{\xi})| &\ll \|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a}v^{d_1+d_2+\frac12 k}(1+\|\boldsymbol{m}\|+\|\boldsymbol{m}'\|)^{2d_1+d_2}\\
&\quad\times(1+\|\boldsymbol{m}-\boldsymbol{\xi}_2\|v^{1/2})^{-A}(1+\|\boldsymbol{m}-\boldsymbol{m}'-\boldsymbol{\xi}_2\|v^{1/2})^{-A}.
\end{aligned}$$

Let $d = d_1 + d_2 = \deg(D)$. Using $1 + \|\boldsymbol{m}\| + \|\boldsymbol{m}'\| \ll (1 + \|\boldsymbol{m}\|)(1 + \|\boldsymbol{m} - \boldsymbol{m}'\|)$ we obtain

$$\big|\big(D(\Theta_f \overline{\Theta}_g)\big)(\tau, \phi; \boldsymbol{\xi})\big| \ll \|f\|_{\mathrm{L}^2_a} \|g\|_{\mathrm{L}^2_a} v^{d+\frac{1}{2}k} \sum_{\boldsymbol{m}, \boldsymbol{m}' \in \mathbb{Z}^k} (1 + \|\boldsymbol{m}\|)^{2d-A} (1 + \|\boldsymbol{m} - \boldsymbol{m}'\|)^{2d-A}.$$

Taking here $A = 2d + k + 1$, we obtain (138).

For the remaining bounds, we apply Lemma 9.1 with fixed integers $A' \geqslant 0$ and $a > A' + \frac{k}{2}$. For any $(\boldsymbol{m}, \boldsymbol{m}')$ satisfying $\|\boldsymbol{m} - \boldsymbol{\xi}_2\| + \|\boldsymbol{m} - \boldsymbol{m}' - \boldsymbol{\xi}_2\| \geqslant \frac{1}{2}$ this leads to

$$\begin{aligned} |F_{\boldsymbol{m},\boldsymbol{m}'}(\tau, \phi; \boldsymbol{\xi})| &= v^{k/2} \Big| f_\phi((\boldsymbol{m} - \boldsymbol{\xi}_2) v^{1/2}) g_\phi((\boldsymbol{m} - \boldsymbol{m}' - \boldsymbol{\xi}_2) v^{1/2}) \Big| \\ &\ll \|f\|_{\mathrm{L}^2_a} \|g\|_{\mathrm{L}^2_a} v^{(k-A')/2} (1 + \|\boldsymbol{m} - \boldsymbol{\xi}_2\|)^{-A'} (1 + \|\boldsymbol{m} - \boldsymbol{m}' - \boldsymbol{\xi}_2\|)^{-A'}. \end{aligned}$$

In particular, this holds for all $(\boldsymbol{m}, \boldsymbol{m}')$ with $\boldsymbol{m}' \neq \boldsymbol{0}$. Hence, taking $A' = k + 2A$, we obtain (139). Similarly, if $\boldsymbol{\xi}_2 \in \boldsymbol{n} + [-\frac{1}{2}, \frac{1}{2}]^k$, then we note that $\|\boldsymbol{m} - \boldsymbol{\xi}_2\| + \|\boldsymbol{m} - \boldsymbol{m}' - \boldsymbol{\xi}_2\| \geqslant \frac{1}{2}$ holds for all $(\boldsymbol{m}, \boldsymbol{m}')$ except $(\boldsymbol{m}, \boldsymbol{m}') = (\boldsymbol{n}, 0)$, and we thus obtain (140). □

9.3. *Bounds on the truncation error*

Let us fix, once and for all, a C^∞ function $g_1 : \mathbb{R}_{>0} \to [0, 1]$ satisfying $g_{1|(0,1]} \equiv 0$ and $g_{1|[2,\infty)} \equiv 1$. For any $Y \geqslant 1$, we define $g_Y : \mathbb{R}_{>0} \to [0, 1]$, $g_Y(y) := g_1(y/Y)$, so that $g_{Y|(0,Y]} \equiv 0$ and $g_{Y|[2Y,\infty)} \equiv 1$. Next, we define the function $\mathcal{X}_Y : G \to \mathbb{R}_{\geqslant 0}$ through

$$\mathcal{X}_Y(\tau, \phi; \boldsymbol{\xi}) = \mathcal{X}_Y(\tau) = \sum_{\gamma \in (\pm\overline{\Gamma}'_\infty) \backslash \mathrm{SL}(2,\mathbb{Z})} g_Y(\operatorname{Im} \gamma\tau). \tag{142}$$

Note here that $(\pm\overline{\Gamma}'_\infty) = \big\{\pm\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right) : n \in \mathbb{Z}\big\}$ (cf. (33)). The function $\mathcal{X}_Y$ is smooth and $\mathrm{SL}(2,\mathbb{Z})$-invariant. For any $\tau \in \mathbb{H}$, there is (since $Y \geqslant 1$) at most one term in the sum in (142) which gives a non-zero contribution. In particular, $\mathcal{X}_Y(g) \in [0, 1]$ for all $g \in G$. Also, in terms of the cuspidal height function $\mathcal{Y}$ (cf. (3)), we have $\mathcal{X}_Y(g) = 0$ whenever $\mathcal{Y}(g) \leqslant Y$ and $\mathcal{X}_Y(g) = 1$ whenever $\mathcal{Y}(g) \geqslant 2Y$.

LEMMA 9.3. *For any $Y \geqslant 1$ and any monomial D in $X_1, \dots, X_{3+2k}$ of degree $\leqslant m$, $D\mathcal{X}_Y$ is a bounded function on G with $\|D\mathcal{X}_Y\|_{\mathrm{L}^\infty} \ll_m 1$.*

Proof. Since $\mathcal{X}_Y$ (and thus $(D\mathcal{X}_Y)$) is $\mathrm{SL}(2,\mathbb{Z})$-invariant, it suffices to consider points $(\tau, \phi, \boldsymbol{\xi})$ with $\tau = u + iv$ belonging to the standard fundamental domain for $\mathrm{SL}(2,\mathbb{Z})$, that is, $|u| \leqslant \frac{1}{2}$ and $|\tau| \geqslant 1$. Then we may in fact assume $v > 1$, since otherwise $(\tau, \phi, \boldsymbol{\xi})$ is not in the support of $\mathcal{X}_Y$. However, for $v > 1$ we have

$$\mathcal{X}_Y(\tau, \phi; \boldsymbol{\xi}) = g_Y(v) = g_1\left(\operatorname{Im}\begin{pmatrix} Y^{-1/2} & 0 \\ 0 & Y^{1/2} \end{pmatrix}(\tau)\right).$$

Since D is left invariant, this implies that $\|D\mathcal{X}_Y\|_{\mathrm{L}^\infty} = \|D\widetilde{\mathcal{X}}_1\|_{\mathrm{L}^\infty}$, where $\widetilde{\mathcal{X}}_1$ is the function $\widetilde{\mathcal{X}}_1 : G \to [0, 1]$, $(\tau, \phi, \boldsymbol{\xi}) \mapsto g_1(\operatorname{Im}\tau)$. This L^∞-norm is clearly finite, and independent of Y. □

For $\boldsymbol{\xi} = \left(\begin{smallmatrix} \xi_1 \\ \xi_2 \end{smallmatrix}\right) \in \mathbb{R}^{2k}$ and $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$, we introduce the short-hand notation $\boldsymbol{\xi}_\gamma := c\boldsymbol{\xi}_1 + d\boldsymbol{\xi}_2$. We also write $v_\gamma := \operatorname{Im}\gamma\tau$ when $\tau = u + iv \in \mathbb{H}$. Given $Y \geqslant 1$ and $f \in \mathrm{C}(\mathbb{R}^k)$ with $S_{\infty,A,0}(f) < \infty$ for some $A > k$, we define the function $F_{f,Y} : G \to \mathbb{C}$ by (cf. [**24**, 6.2])

$$F_{f,Y}(\tau, \phi; \boldsymbol{\xi}) = F_{f,Y}(\tau; \boldsymbol{\xi}) := \sum_{\gamma \in \overline{\Gamma}'_\infty \backslash \mathrm{SL}(2,\mathbb{Z})} \sum_{\boldsymbol{m} \in \mathbb{Z}^k} f\Big((\boldsymbol{\xi}_\gamma + \boldsymbol{m}) v_\gamma^{1/2}\Big)\, v_\gamma^{k/2}\, g_Y(v_\gamma). \tag{143}$$

This series is absolutely convergent, and $F_{f,Y}$ is left $\overline{\Gamma}$ invariant. In fact, we will only use $F_{f,Y}$ for functions $f \geqslant 0$; then of course $F_{f,Y} \geqslant 0$.

As in [23, 6.3; 24, 6.4], we have the explicit formula

$$\begin{aligned} F_{f,Y}(\tau;\boldsymbol{\xi}) = &\sum_{\boldsymbol{m}\in\mathbb{Z}^k} \Big\{ f\Big((\boldsymbol{\xi}_2+\boldsymbol{m})v^{1/2}\Big) + f\Big((-\boldsymbol{\xi}_2+\boldsymbol{m})v^{1/2}\Big)\Big\} v^{k/2} g_Y(v) \\ &+ \sum_{\boldsymbol{m}\in\mathbb{Z}^k} \left\{ f\left((\boldsymbol{\xi}_1+\boldsymbol{m})\frac{v^{1/2}}{|\tau|}\right) + f\left((-\boldsymbol{\xi}_1+\boldsymbol{m})\frac{v^{1/2}}{|\tau|}\right)\right\} \frac{v^{k/2}}{|\tau|^k} g_Y\left(\frac{v}{|\tau|^2}\right) \\ &+ \sum_{\substack{(c,d)\in\mathbb{Z}^2\\ \gcd(c,d)=1\\ c,d\neq 0}} \sum_{\boldsymbol{m}\in\mathbb{Z}^k} f\left((c\boldsymbol{\xi}_1+d\boldsymbol{\xi}_2+\boldsymbol{m})\frac{v^{1/2}}{|c\tau+d|}\right) \frac{v^{k/2}}{|c\tau+d|^k} g_Y\left(\frac{v}{|c\tau+d|^2}\right). \end{aligned} \tag{144}$$

The following lemma shows that for an appropriate choice of f^*, the function $F_{f^*,Y}$ controls the error when truncating $\Theta_f\overline{\Theta_g}$ at height $\asymp Y$.

LEMMA 9.4. *Let $f, g \in \mathcal{S}(\mathbb{R}^k)$ and let A and a be integers satisfying $A \geqslant 1$ and $a > \frac{3}{2}k + 2A$. Set $f^*(\boldsymbol{w}) = \sup_{\phi\in\mathbb{R}} \big|f_\phi(\frac{1}{2}\boldsymbol{w}) g_\phi(\frac{1}{2}\boldsymbol{w})\big|$. Then for any $Y \geqslant 1$,*

$$\mathcal{X}_Y(\tau)\, \big|(\Theta_f\overline{\Theta}_g)(\tau,\phi;\boldsymbol{\xi})\big| \leqslant F_{f^*,Y}(\tau;2\boldsymbol{\xi}) + O_A\big(\|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a} Y^{-A}\big), \quad \forall (\tau,\phi,\boldsymbol{\xi}) \in G. \tag{145}$$

(Here, $F_{f^*,Y}(\tau;2\boldsymbol{\xi})$ is well defined, since $S_{\infty,A',0}(f^*) < \infty$ for all $A' > 0$ by Lemma 9.1.)

Proof. See [24, 8.4.3]. For any $(\tau,\phi,\boldsymbol{\xi})$, with τ lying in the standard fundamental domain $\mathcal{F}$ for $\mathrm{SL}(2,\mathbb{Z})$, $\mathcal{F} = \{\tau = u+iv \in \mathbb{H} : |u| \leqslant \frac{1}{2},\ |\tau| \geqslant 1\}$, it follows from the definition of f^* together with (139) in Lemma 9.2 that

$$\big|(\Theta_f\overline{\Theta}_g)(\tau,\phi;\boldsymbol{\xi})\big| \leqslant v^{k/2} \sum_{\boldsymbol{m}\in\mathbb{Z}^k} f^*\Big(2(\boldsymbol{m}-\boldsymbol{\xi}_2)v^{1/2}\Big) + O_A\big(\|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a} v^{-A}\big).$$

Multiplying this inequality with $g_Y(v)$ and comparing with (144), we obtain that (145) holds for all $(\tau,\phi,\boldsymbol{\xi}) \in G$ with $\tau \in \mathcal{F}$, since $\mathcal{X}_Y(\tau) = g_Y(v)$ for all such τ. But both sides in (145) are functions of $(\tau,\phi,\boldsymbol{\xi}) \in G$ which are Γ^k left invariant (for the function $\widehat{F}_{f^*,Y}(\tau;\boldsymbol{\xi}) := F_{f^*,Y}(\tau;2\boldsymbol{\xi})$ this is noted in [23, 6.9-10; 24, 7.5-6]); hence the inequality holds for *all* $(\tau,\phi,\boldsymbol{\xi}) \in G$. □

The following lemma is a more explicit version of [23, Lemma 6.5]. Recall that our κ corresponds to '$\kappa - 1$' in [23].

LEMMA 9.5. *Let $A > k$. Then for any $[\kappa;c]$-Diophantine $\boldsymbol{\alpha} \in \mathbb{R}^k$, and any $D, T \geqslant 1$,*

$$\sum_{d=1}^{D} \sum_{\boldsymbol{m}\in\mathbb{Z}^k} (1+T\|d\boldsymbol{\alpha}+\boldsymbol{m}\|)^{-A} \ll_{k,A} \begin{cases} D^{A\kappa+1}(cT)^{-A} & \text{if } D^{\kappa+A^{-1}} \leqslant cT \\ 1 & \text{if } D^{\kappa} \leqslant cT \leqslant D^{\kappa+A^{-1}} \\ D(cT)^{-1/\kappa} & \text{if } cT \leqslant D^{\kappa}. \end{cases} \tag{146}$$

Proof. Since $\boldsymbol{\alpha}$ is $[\kappa;c]$-Diophantine, $\|d\boldsymbol{\alpha}+\boldsymbol{m}\| \geqslant cd^{-\kappa}$ for all integers $d \geqslant 1$ and $\boldsymbol{m} \in \mathbb{Z}^k$. Also for each fixed d, there is at most one $\boldsymbol{m} \in \mathbb{Z}^k$ in the box $-d\boldsymbol{\alpha} + (-\frac{1}{2},\frac{1}{2})^k$, and in particular

there is at most one $\boldsymbol{m} \in \mathbb{Z}^k$ with $\|d\boldsymbol{\alpha} + \boldsymbol{m}\| < \frac{1}{2}$. Hence, for $d \in \{1, \dots, D\}$,

$$\sum_{\boldsymbol{m} \in \mathbb{Z}^k} (1 + T\|d\boldsymbol{\alpha} + \boldsymbol{m}\|)^{-A} \ll_{k,A} (1 + Tcd^{-\kappa})^{-A} + T^{-A} \ll_A \left(\frac{D^\kappa}{cT}\right)^A,$$

where in the last inequality, we use the fact that $D^\kappa/c > 1$ (note that $\boldsymbol{\alpha}$ being $[\kappa; c]$-Diophantine implies $c \leqslant \frac{1}{2}$). Adding the above bound over $d = 1, \dots, D$, we obtain that the left-hand side of (146) is $\ll D^{A\kappa+1}(cT)^{-A}$.

To prove another bound on the same sum, for any fixed $b \in \mathbb{Z}$, we start by considering the set

$$M_b = \Big\{ T((b+d)\boldsymbol{\alpha} + \boldsymbol{m}) \ : \ d \in \mathbb{Z},\ 0 \leqslant d \leqslant (cT)^{1/\kappa},\ \boldsymbol{m} \in \mathbb{Z}^k \Big\}. \tag{147}$$

The distance between any two distinct points in this set is bounded from below by

$$\min \Big\{ T\|q\boldsymbol{\alpha} + \boldsymbol{n}\| \ : \ q \in \mathbb{Z},\ \boldsymbol{n} \in \mathbb{Z}^k,\ |q| \leqslant (cT)^{1/\kappa},\ [q \neq 0 \text{ or } \boldsymbol{n} \neq \boldsymbol{0}] \Big\}$$

$$\geqslant \min \left(T, \min_{0 < q \leqslant (cT)^{1/\kappa}} Tcq^{-\kappa} \right) \geqslant 1,$$

where the first inequality holds since $\boldsymbol{\alpha}$ is $[\kappa; c]$-Diophantine. (Note also that there is no double representation in (147), that is, $T((b+d)\boldsymbol{\alpha} + \boldsymbol{m})$ is an injective function of $\langle d, \boldsymbol{m} \rangle \in \mathbb{Z} \times \mathbb{Z}^k$.) Hence, for any $R \geqslant 1$, M_b contains $\ll_k R^k$ points with $\|\boldsymbol{x}\| \leqslant R$, and so by a standard dyadic decomposition we have

$$\sum_{\boldsymbol{x} \in M_b} (1 + \|\boldsymbol{x}\|)^{-A} \ll_{k,A} 1. \tag{148}$$

Now by appropriate choices of b, the sum in (146) can be majorized by $1 + D(cT)^{-1/\kappa}$ sums as in (148).

We have thus proved that the left-hand side of (146) is always $\ll D^{A\kappa+1}(cT)^{-A}$, and also that it is $\ll 1 + D(cT)^{-1/\kappa}$. Splitting into cases depending on which bound is strongest, we obtain the statement in (146). □

The following proposition is an effective version of [**24**, Proposition 6.5], and is the central result needed to bound the error caused by truncating the function $\Theta_f \overline{\Theta_g}$ in the integral (130). We here specialize to the case $k = 2$; the case $k \geqslant 3$ involves in principle the same computations, however, there are several differences in the detailed analysis (cf. [**23**, Proposition 6.4]).

PROPOSITION 9.6. *Let $k = 2$ and let $\boldsymbol{\xi}_2 \in \mathbb{R}^2$ be $[\kappa; c_0]$-Diophantine. Let $Y \geqslant 1$, $0 < v \leqslant Y$, $A > 2$, $B \geqslant 1$, and $H \geqslant 1$. Then for any $f \in \mathrm{C}(\mathbb{R}^2)$ with $S_{\infty,A,0}(f) < \infty$, and any bounded function $h : \mathbb{R} \to \mathbb{R}$ with support contained in $[-H, H]$,*

$$\int_{|u| > Bv} F_{f,Y}\left(u + iv; \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) h(u)\, du$$

$$\ll_A S_{\infty,A,0}(f)\|h\|_{\mathrm{L}^\infty} \left\{ B^{-1} + H\left(H^{-1}c_0^{-1}v^{\frac{1}{2}}\right)^{\frac{1}{\kappa+1+A^{-1}}} + H\kappa c_0^{-\frac{1}{\kappa}} Y^{-\frac{1}{2\kappa}} \right\}. \tag{149}$$

We remark that the integral in (149) *vanishes* if $Y^{-1} \leqslant v \leqslant Y$; hence the bound in (149) is mainly relevant when $v < Y^{-1}$.

Proof. Without loss of generality, let us assume that f is positive and even, that is, $f \geqslant 0$ and $f(-\boldsymbol{w}) = f(\boldsymbol{w})$. Recall the expansion (144), and note that the terms with $g_Y(v)$ vanish

since $v \leqslant Y$; hence we are left with

$$F_{f,Y}\left(\tau; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) = 2 \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left(\boldsymbol{m}\frac{v^{1/2}}{|\tau|}\right)\frac{v}{|\tau|^2} g_Y\left(\frac{v}{|\tau|^2}\right)$$
$$+ 2 \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ \gcd(c,d)=1 \\ c>0, d\neq 0}} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left((d\boldsymbol{\xi}_2 + \boldsymbol{m})\frac{v^{1/2}}{|c\tau+d|}\right)\frac{v}{|c\tau+d|^2} g_Y\left(\frac{v}{|c\tau+d|^2}\right). \tag{150}$$

The contribution from the first sum in (150) to the integral in (149) is

$$\int_{|u|>Bv} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left(\boldsymbol{m}\frac{v^{1/2}}{|\tau|}\right)\frac{v}{|\tau|^2} g_Y\left(\frac{v}{|\tau|^2}\right) h(u)\, du$$
$$= \int_{|t|>B} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left(\frac{\boldsymbol{m}}{v^{1/2}(t^2+1)^{1/2}}\right)\frac{1}{t^2+1} g_Y\left(\frac{1}{v(t^2+1)}\right) h(vt)\, dt. \tag{151}$$

Using $|f(\boldsymbol{x})| \leqslant S_{\infty,A,0}(f)(1+\|\boldsymbol{x}\|)^{-A}$ and $A > 2$, we have

$$\sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left(\frac{\boldsymbol{m}}{v^{1/2}(t^2+1)^{1/2}}\right) \ll_A S_{\infty,A,0}(f)$$

uniformly over all v, t subject to $v^{-1/2}(t^2+1)^{-1/2} \geqslant 1$; and for all other pairs v, t the factor $g_Y\left(\frac{1}{v(t^2+1)}\right)$ vanishes (since $Y \geqslant 1$). Hence, (151) is

$$\ll_A S_{\infty,A,0}(f)\|h\|_{\mathrm{L}^\infty} \int_{|t|>B} \frac{dt}{t^2+1} \ll S_{\infty,A,0}(f)\|h\|_{\mathrm{L}^\infty} B^{-1}.$$

The contribution from the remaining double sum in (150) to the integral in (149) is bounded above by (we drop the condition $|u| > Bv$ in the integral):

$$\int_{\mathbb{R}} \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ \gcd(c,d)=1 \\ c>0, d\neq 0}} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left((d\boldsymbol{\xi}_2 + \boldsymbol{m})\frac{v^{1/2}}{|c\tau+d|}\right)\frac{v}{|c\tau+d|^2} g_Y\left(\frac{v}{|c\tau+d|^2}\right) h(u)\, du$$
$$= \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ \gcd(c,d)=1 \\ c>0, d\neq 0}} \frac{1}{c^2} \int_{\mathbb{R}} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} f\left(\frac{d\boldsymbol{\xi}_2 + \boldsymbol{m}}{\sqrt{c^2 v(t^2+1)}}\right) g_Y\left(\frac{1}{c^2 v(t^2+1)}\right) h\left(vt - \frac{d}{c}\right)\frac{dt}{t^2+1}, \tag{152}$$

where we changed the order of integration and summation and then substituted $u = vt - \frac{d}{c}$. The g_Y-factor in the above expression vanishes unless t belongs to the set

$$I_c = \left\{t \in \mathbb{R} \,:\, \frac{1}{c^2 v(t^2+1)} > Y\right\} = \left\{t \in \mathbb{R} \,:\, \sqrt{t^2+1} < (vY)^{-\frac{1}{2}} c^{-1}\right\}.$$

Furthermore, the factor $h(vt - \frac{d}{c})$ vanishes unless $|vt - \frac{d}{c}| \leqslant H$, and for $t \in I_c$ this forces

$$\left|\frac{d}{c}\right| \leqslant H + \frac{\sqrt{v}}{c\sqrt{Y}} \leqslant H + 1 \leqslant 2H.$$

Also using $f(\boldsymbol{x}) \leqslant S_{\infty,A,0}(f)(1+\|\boldsymbol{x}\|)^{-A}$, we conclude that the expression in (152) is

$$\leqslant S_{\infty,A,0}(f)\|h\|_{\mathrm{L}^\infty} \sum_{c=1}^{\infty} \frac{1}{c^2} \int_{I_c} \sum_{0<|d|\leqslant 2Hc} \sum_{\boldsymbol{m}\in\mathbb{Z}^2} \left(1+\frac{\|d\boldsymbol{\xi}_2+\boldsymbol{m}\|}{\sqrt{c^2 v(t^2+1)}}\right)^{-A} \frac{dt}{t^2+1}.$$

Applying Lemma 9.5 with

$$D = 2Hc \quad \text{and} \quad T = \frac{1}{\sqrt{c^2 v(t^2+1)}},$$

and both $\boldsymbol{\alpha} = \boldsymbol{\xi}_2$ and $\boldsymbol{\alpha} = -\boldsymbol{\xi}_2$, we get

$$\ll_A S_{\infty,A,0}(f)\|h\|_{\mathrm{L}^\infty} \sum_{c=1}^{\infty} c^{-2} \Bigg(\int_{I_{1,c}} D^{A\kappa+1}(c_0 T)^{-A} \frac{dt}{t^2+1} + \int_{I_{2,c}} \frac{dt}{t^2+1} + \int_{I_{3,c}} D(c_0 T)^{-\delta} \frac{dt}{t^2+1} \Bigg), \tag{153}$$

where $\delta := 1/\kappa$ and

$$I_{1,c} = \left\{ t \in I_c \,:\, \sqrt{t^2+1} \leqslant c_0 (v^{\frac{1}{2}} c)^{-1} D^{-(\kappa+A^{-1})} \right\};$$

$$I_{2,c} = \left\{ t \in I_c \,:\, c_0 (v^{\frac{1}{2}} c)^{-1} D^{-(\kappa+A^{-1})} < \sqrt{t^2+1} \leqslant c_0 (v^{\frac{1}{2}} c)^{-1} D^{-\kappa} \right\};$$

$$I_{3,c} = \left\{ t \in I_c \,:\, c_0 (v^{\frac{1}{2}} c)^{-1} D^{-\kappa} \leqslant \sqrt{t^2+1} \right\}.$$

We discuss the three integrals in (153) one by one. First, note that $I_{1,c} \neq \emptyset$ implies $c \leqslant C_1 := \left((2H)^{-(\kappa+A^{-1})} c_0 v^{-\frac{1}{2}}\right)^{1/(\kappa+1+A^{-1})}$, and for each such c, $t \in I_{1,c}$ implies $\sqrt{t^2+1} \leqslant (C_1/c)^{\kappa+1+A^{-1}}$. Hence,

$$\begin{aligned}
&\sum_{c=1}^{\infty} c^{-2} \int_{I_{1,c}} D^{A\kappa+1}(c_0 T)^{-A} \frac{dt}{t^2+1} \\
&\leqslant \sum_{1\leqslant c\leqslant C_1} c^{-2} \cdot D^{A\kappa+1} c_0^{-A} v^{\frac{A}{2}} c^A \int_{I_{1,c}} (t^2+1)^{\frac{A}{2}-1}\, dt \\
&\ll_A \sum_{1\leqslant c\leqslant C_1} c^{-2} \cdot D^{A\kappa+1} c_0^{-A} v^{\frac{A}{2}} c^A (C_1/c)^{(\kappa+1+A^{-1})(A-1)} \\
&= (2H)^{\kappa+A^{-1}} c_0^{-1} v^{\frac{1}{2}} \sum_{1\leqslant c\leqslant C_1} c^{\kappa+A^{-1}-1} \ll (2H)^{\kappa+A^{-1}} c_0^{-1} v^{\frac{1}{2}} C_1^{\kappa+A^{-1}} = C_1^{-1}.
\end{aligned}$$

Turning to the integral over $I_{2,c}$, the fact that $t \in I_{2,c}$ forces $\sqrt{t^2+1} > (C_1/c)^{\kappa+1+A^{-1}}$, with C_1 as above. Therefore, we see that

$$\sum_{c=1}^{\infty} c^{-2} \int_{I_{2,c}} \frac{dt}{t^2+1} \leqslant \sum_{1\leqslant c\leqslant C_1} c^{-2} (C_1/c)^{-(\kappa+1+A^{-1})} + \sum_{c>C_1} c^{-2} \ll C_1^{-1}.$$

Finally, for the integral over $I_{3,c}$ we have, using only $I_{3,c} \subset I_c$,

$$\sum_{c=1}^{\infty} c^{-2} \int_{I_{3,c}} D(c_0 T)^{-\delta} \frac{dt}{t^2+1} \leqslant 2H c_0^{-\delta} v^{\frac{\delta}{2}} \int_{\mathbb{R}} \left(\sum_{1 \leqslant c < (vY(t^2+1))^{-1/2}} c^{\delta-1} \right) (t^2+1)^{\frac{\delta}{2}-1} \, dt$$

$$\ll H c_0^{-\delta} v^{\frac{\delta}{2}} \int_{\mathbb{R}} \frac{(vY(t^2+1))^{-\frac{\delta}{2}}}{\delta} (t^2+1)^{\frac{\delta}{2}-1} \, dt \ll H \kappa c_0^{-\delta} Y^{-\frac{\delta}{2}}.$$

Hence, we obtain the bound in (149). □

Next, we note that Proposition 9.6 can be extended in a straightforward manner to the case of functions h which do not have compact support but decay appropriately at infinity:

COROLLARY 9.7. *Let $\boldsymbol{\xi}_2 \in \mathbb{R}^2$ be $[\kappa; c_0]$-Diophantine. Let $Y \geqslant 1$, $0 < v \leqslant Y$, $A > 2$ and $B \geqslant 1$. Then for any $f \in \mathrm{C}(\mathbb{R}^2)$ with $S_{\infty,A,0}(f) < \infty$ and any function $h : \mathbb{R} \to \mathbb{R}$ with $S_{\infty,2,0}(h) < \infty$,*

$$\int_{|u|>Bv} F_{f,Y}\left(u + iv; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix} \right) h(u) \, du$$

$$\ll_A S_{\infty,A,0}(f) S_{\infty,2,0}(h) \left\{ B^{-1} + \left(c_0^{-1} v^{\frac{1}{2}} \right)^{\frac{1}{\kappa+1+A^{-1}}} + \kappa c_0^{-1/\kappa} Y^{-\frac{1}{2\kappa}} \right\}. \tag{154}$$

Proof. Decompose the function h as $h = h_0 + h_1 + \cdots$ where $h_0 = h \cdot \chi_{[-1,1]}$ and $h_j = h \cdot \left(\chi_{[-2^j,-2^{j-1})} + \chi_{(2^{j-1},2^j]} \right)$ for $j \geqslant 1$; then apply Proposition 9.7 to bound the contribution from each function h_j, using $\operatorname{supp} h_j \subset [-2^j, 2^j]$ and $\|h_j\|_{\mathrm{L}^\infty} \ll S_{\infty,2,0}(h) 2^{-2j}$. □

9.4. *Proof of Theorem 1.4*

We are now ready to give the proof of Theorem 1.4. The first step is to give an effective rate for the convergence of the integral in (130) to its limit; this is obtained in Proposition 9.10. The proof of this proposition is divided into two lemmas, Lemma 9.8 which concerns the part of the integral where u is not very near zero, and Lemma 9.9 which concerns the remaining part. These two lemmas are (in principle) effective versions of [**24**, Corollary 7.4] and [**24**, Lemma 8.3], respectively.

Throughout this section, we let $\Gamma = \Gamma(2) \ltimes \mathbb{Z}^4$, and $G = \mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus 2}$. Recall that $\Theta_f \overline{\Theta_g}$ is a left Γ^2 invariant function on G, with Γ^2 as in (129); thus in particular, it is left invariant under $\Gamma = \Gamma(2) \ltimes \mathbb{Z}^4$. As always, we let μ be the probability measure on $\Gamma \backslash G$ induced by an appropriately normalized Haar measure on G (which we also denote by μ).

LEMMA 9.8. *Let $f, g \in \mathcal{S}(\mathbb{R}^2)$ and $h \in \mathrm{C}^2(\mathbb{R})$, and assume $S_{\infty,3,2}(h) < \infty$. Let $\boldsymbol{\xi}_2 \in \mathbb{R}^2$ be $[\kappa; c]$-Diophantine. Then for any $v \in (0,1]$ and any real number B subject to*

$$1 \leqslant B \leqslant \tfrac{1}{2} v^{-\frac{1}{2}} \left(\tfrac{1}{4} \delta_{6,\boldsymbol{\xi}_2}\left(v^{-\frac{1}{2}} \right)^{\frac{1}{2}} \right)^{\frac{\kappa}{1+61\kappa}}, \tag{155}$$

we have

$$\left| \int_{|u|>Bv} \Theta_f\left(u+iv, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix} \right) \overline{\Theta_g\left(u+iv, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix} \right)} \, h(u) \, du - \int_{\Gamma\backslash G} \Theta_f \overline{\Theta_g} \, d\mu \int_{\mathbb{R}} h \, du \right|$$

$$\ll \|f\|_{\mathrm{L}^2_{167}} \|g\|_{\mathrm{L}^2_{167}} S_{\infty,3,2}(h) \left(\kappa c^{-\frac{1}{\kappa}} \delta_{6,\boldsymbol{\xi}_2}\left(v^{-\frac{1}{2}} \right)^{\frac{1}{127\kappa}} + B^{-1} \right). \tag{156}$$

REMARK 12. As will be seen in the proof, the (quite small) power $\frac{1}{127\kappa}$ which we obtain in (156) depends strongly on which C_a^m-norm of the test function (that is, $\widetilde{F}$ below) is required to bound in the effective equidistribution result of Theorem 1.2. Since we did not make any effort to optimize the a and m in Theorem 1.2, we do not attempt to optimize the decay rate with respect to v in Lemma 9.8 nor in Lemma 9.9 or Proposition 9.10. We instead focus on giving results which are simple to state, yet explicit.

Proof. Let $F = \Theta_f \overline{\Theta}_g$. Also, for $Y \geqslant 1$ a real number which we will choose below (cf. (161)), let $\widetilde{F} = (1 - \mathcal{X}_Y) \cdot F$, that is,

$$\widetilde{F}(\tau, \phi, \boldsymbol{\xi}) = (1 - \mathcal{X}_Y(\tau)) \cdot (\Theta_f \overline{\Theta}_g)(\tau, \phi; \boldsymbol{\xi}).$$

Then both F and $\widetilde{F}$ are Γ^2 left invariant functions on G; in particular, they are left invariant under $\Gamma = \Gamma(2) \ltimes \mathbb{Z}^4$. Our choice of Y will be such that

$$Y \leqslant \frac{1}{2(B^2+1)v}. \tag{157}$$

Then for all u with $|u| \leqslant Bv$, we have

$$\operatorname{Im}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}(u+iv)\right) = \frac{v}{u^2+v^2} \geqslant \frac{1}{(1+B^2)v} \geqslant 2Y \geqslant 2, \quad \text{and thus } \mathcal{X}_Y(u+iv) = 1$$

(cf. the discussion below (142)). Hence, $\int_{|u|>Bv} \widetilde{F}(u+iv, 0; \boldsymbol{\xi})\, h(u)\, du = \int_{\mathbb{R}} \widetilde{F}(u+iv, 0; \boldsymbol{\xi})\, h(u)\, du$, and so by Theorem 1.2, applied with $\beta = 6$, $m = 27$ and $a = 5/2$, we have

$$\begin{aligned} &\int_{|u|>Bv} \widetilde{F}\left(u+iv, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) h(u)\, du \\ &\quad = \int_{\Gamma\backslash G} \widetilde{F}\, d\mu \int_{\mathbb{R}} h(u)\, du + O_\varepsilon\Big(\|\widetilde{F}\|_{\mathrm{C}^{27}_{5/2}} S_{\infty,3,2}(h)\left(\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac{1}{2}}) + v^{\frac{1}{4}-\varepsilon}\right)\Big). \end{aligned} \tag{158}$$

In order to bound $\|\widetilde{F}\|_{\mathrm{C}^{27}_{5/2}}$, we apply Lemma 9.3 and (138) in Lemma 9.2, together with the fact that F and $\widetilde{F}$ are Γ^2-invariant. It follows that for any integer $a > 4 + 6 \cdot 27 = 166$,

$$\sum_{\operatorname{ord}(D)\leqslant 27} \left|(D\widetilde{F})(M, \boldsymbol{\xi})\right| \ll \|f\|_{\mathrm{L}^2_a} \|g\|_{\mathrm{L}^2_a}\, \mathcal{Y}(M)^{28}, \quad \forall (M, \boldsymbol{\xi}) \in G.$$

Hence, since the support of $\widetilde{F}$ is contained in $\{\mathcal{Y}(M) \leqslant 2Y\}$,

$$\|\widetilde{F}\|_{\mathrm{C}^{27}_{5/2}} \ll \|f\|_{\mathrm{L}^2_a} \|g\|_{\mathrm{L}^2_a}\, Y^{61/2}.$$

Next, we bound the error caused by replacing $\widetilde{F}$ by F in the two integrals $\int_{\Gamma\backslash G} \widetilde{F}\, d\mu$ and $\int_{|u|>Bv} \widetilde{F}\big(u+iv, 0; \left(\begin{smallmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{smallmatrix}\right)\big)\, h(u)\, du$ in (158). First note that by Lemma 9.4, we have

$$\left|F(\tau, \phi; \boldsymbol{\xi}) - \widetilde{F}(\tau, \phi; \boldsymbol{\xi})\right| \leqslant F_{f^*,Y}(\tau; 2\boldsymbol{\xi}) + O\Big(\|f\|_{\mathrm{L}^2_6} \|g\|_{\mathrm{L}^2_6} Y^{-1}\Big), \quad \forall (\tau, \phi; \boldsymbol{\xi}) \in G, \tag{159}$$

with $f^*(\boldsymbol{w}) := \sup_{\phi\in\mathbb{R}} \left|f_\phi(\frac{1}{2}\boldsymbol{w}) g_\phi(\frac{1}{2}\boldsymbol{w})\right|$. Hence,

$$\begin{aligned} &\left|\int_{|u|>Bv} \widetilde{F}\left(u+iv, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) h(u)\, du - \int_{|u|>Bv} F\left(u+iv, 0; \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\xi}_2 \end{pmatrix}\right) h(u)\, du\right| \\ &\quad \leqslant \int_{|u|>Bv} F_{f^*,Y}\left(u+iv; \begin{pmatrix} \mathbf{0} \\ 2\boldsymbol{\xi}_2 \end{pmatrix}\right) |h(u)|\, du + O\Big(\|f\|_{\mathrm{L}^2_6} \|g\|_{\mathrm{L}^2_6} \|h\|_{\mathrm{L}^1}\, Y^{-1}\Big). \end{aligned}$$

Note that since $\boldsymbol{\xi}_2$ is $[\kappa; c]$-Diophantine, $2\boldsymbol{\xi}_2$ is $[\kappa; 2^{-\kappa}c]$-Diophantine. Hence, applying Corollary 9.7 with $A = 3$, and noticing that $S_{\infty,3,0}(f^*) \ll \|f\|_{\mathrm{L}^2_4}\|g\|_{\mathrm{L}^2_4}$ by Lemma 9.1, we get

$$\int_{|u|>Bv} F_{f^*,Y}\left(u+iv; \begin{pmatrix}\mathbf{0}\\ 2\boldsymbol{\xi}_2\end{pmatrix}\right)|h(u)|\,du$$
$$\ll \|f\|_{\mathrm{L}^2_4}\|g\|_{\mathrm{L}^2_4} S_{\infty,2,0}(h)\Big\{B^{-1} + \Big(c^{-1}v^{\frac12}\Big)^{1/(\kappa+\frac43)} + \kappa\, c^{-\frac{1}{\kappa}} Y^{-\frac{1}{2\kappa}}\Big\}.$$

Also by (159) we have

$$\int_{\Gamma\backslash G}\Big|F-\widetilde{F}\Big|\,d\mu \leqslant \int_{\Gamma\backslash G} F_{f^*,Y}(\tau; 2\boldsymbol{\xi})\,d\mu(\tau,\phi;\boldsymbol{\xi}) + O\Big(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}Y^{-1}\Big).$$

Here, one computes, by a standard unfolding argument (cf. [**23**, 6.2]),

$$\int_{\Gamma\backslash G} F_{f^*,Y}(\tau;2\boldsymbol{\xi})\,d\mu = \int_{\Gamma\backslash G} F_{f^*,Y}(\tau;\boldsymbol{\xi})\,d\mu = \frac{3}{\pi Y}\int_0^\infty g_1(y)\,\frac{dy}{y^2}\int_{\mathbb{R}^2} f^*\,d\boldsymbol{w} \ll \|f\|_{\mathrm{L}^2_4}\|g\|_{\mathrm{L}^2_4}Y^{-1}.$$

We combine the above bounds with (158), where we also use the fact that $\delta_{\beta,\boldsymbol{\xi}_2}(T) \geqslant T^{-1}$ ($\forall T \geqslant 1$), which follows by just considering the terms corresponding to $\boldsymbol{r} = \pm\boldsymbol{e}_1$ and $j = 1$ in (5). We then get, with $a = 167$:

$$\left|\int_{|u|>Bv} F\left(u+iv, 0; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\right)h(u)\,du - \int_{\Gamma\backslash G} F\,d\mu\int_{\mathbb{R}} h\,du\right|$$
$$\ll_\varepsilon \|f\|_{\mathrm{L}^2_a}\|g\|_{\mathrm{L}^2_a} S_{\infty,3,2}(h)\,\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac12-2\varepsilon}\,Y^{\frac{61}{2}} + \|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}\|h\|_{\mathrm{L}^1}Y^{-1} \tag{160}$$
$$+ \|f\|_{\mathrm{L}^2_4}\|g\|_{\mathrm{L}^2_4} S_{\infty,2,0}(h)\Big\{B^{-1} + \Big(c^{-1}v^{\frac12}\Big)^{1/(\kappa+\frac43)} + \kappa\, c^{-\frac{1}{\kappa}} Y^{-\frac{1}{2\kappa}}\Big\}.$$

In order to minimize the order of magnitude of the maximum of $\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac12-2\varepsilon}Y^{\frac{61}{2}}$ and $Y^{-\frac{1}{2\kappa}}$, we now make the choice

$$Y := \Big(\tfrac13\,\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac12-2\varepsilon}\Big)^{-\frac{2\kappa}{1+61\kappa}}. \tag{161}$$

Because of the factor $\frac13$ in this expression, we are guaranteed to have $Y \geqslant 1$, as required above. Indeed, one verifies $\sum_{\boldsymbol{r}\in\mathbb{Z}^2\setminus\{\mathbf{0}\}}\|\boldsymbol{r}\|^{-6}\sum_{j=1}^\infty j^{-2} < 9$; hence $\delta_{6,\boldsymbol{\xi}_2}(T) < 9$ for all $T \geqslant 1$ (cf. (6)). Furthermore, our assumption (155) ensures that (157) is fulfilled, so long as $9^{2\varepsilon} \leqslant \frac43$. Note also that $Y^{-\frac{1}{2\kappa}} \geqslant Y^{-1}$ (since $\kappa \geqslant \frac12$) and $\big(c^{-1}v^{\frac12}\big)^{1/(\kappa+\frac43)} \ll \kappa\, c^{-\frac{1}{\kappa}}Y^{-\frac{1}{2\kappa}}$ (since $\kappa \geqslant \frac12$, $0 < c < 2^{-1/2} < 1$ and $\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac12-2\varepsilon} \geqslant v^{\frac14-\varepsilon} \geqslant v^{\frac14}$). Finally, we take $\varepsilon = \frac{1}{508}$ and note that we then have $(\frac12-2\varepsilon)\frac{1}{1+61\kappa} \geqslant \frac{1}{127\kappa}$, since $\kappa \geqslant \frac12$; also $9^{2\varepsilon} \leqslant \frac43$ as required above. Hence, the bound (156) now follows from (160). □

LEMMA 9.9. *Suppose $f, g \in \mathcal{S}(\mathbb{R}^2)$. Let $h \in \mathrm{C}^1(\mathbb{R})$ with h and h' bounded. Then for any $\boldsymbol{\xi} = \left(\begin{smallmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{smallmatrix}\right)$, $v \in (0,1]$ and $B \in [1, v^{-1/2}]$, we have*

$$\int_{|u|<Bv} \Theta_f(u+iv,0;\boldsymbol{\xi})\overline{\Theta_g(u+iv,0;\boldsymbol{\xi})}\,h(u)\,du = \lambda_{f,\overline{g}}\,h(0) + O\Big(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}S_{\infty,0,1}(h)B^{-1}\Big),$$

where

$$\lambda_{f,g} = \int_0^\infty \left(\int_0^{2\pi} f(r\cos\zeta, r\sin\zeta)\,d\zeta\right)\left(\int_0^{2\pi} g(r\cos\zeta, r\sin\zeta)\,d\zeta\right) r\,dr. \tag{162}$$

Proof. Recall that the function $\Theta_f\overline{\Theta}_g$ is left Γ^2 invariant; in particular, it is invariant under left multiplication by $\left(\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right), \mathbf{0}\right) \in \Gamma^2$, and so

$$\Theta_f(\tau,0;\boldsymbol{\xi})\overline{\Theta_g(\tau,0;\boldsymbol{\xi})} = \Theta_f\left(-\frac{1}{\tau}, \arg\tau; \begin{pmatrix}-\boldsymbol{\xi}_2\\ 0\end{pmatrix}\right)\overline{\Theta_g\left(-\frac{1}{\tau}, \arg\tau; \begin{pmatrix}-\boldsymbol{\xi}_2\\ 0\end{pmatrix}\right)},$$

for all $\tau = u+iv \in \mathbb{H}$. By (140) in Lemma 9.2 (applied with $A=1$), if $\operatorname{Im}(-1/\tau) \geqslant \frac{1}{2}$, then the last expression equals

$$f_{\arg\tau}(0)\overline{g_{\arg\tau}(0)}\frac{v}{|\tau|^2} + O\left(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}\frac{|\tau|^2}{v}\right).$$

Note that $|u| \leqslant v^{1/2} \leqslant 1$ implies $\operatorname{Im}(-1/\tau) \geqslant \frac{1}{2}$, that is, the above holds for all $\tau \in \mathbb{H}$ with $|u| \leqslant v^{1/2} \leqslant 1$. Hence, we get

$$\begin{aligned}&\int_{|u|<Bv} \Theta_f(u+iv,0;\boldsymbol{\xi})\overline{\Theta_g(u+iv,0;\boldsymbol{\xi})}h(u)\,du\\ &\quad= \int_{|u|<Bv}\frac{v}{|\tau|^2}f_{\arg\tau}(0)\overline{g_{\arg\tau}(0)}\,h(u)\,du + O\Big(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}\|h\|_{\mathrm{L}^\infty}B^3v^2\Big).\end{aligned}$$

Using polar co-ordinates we get (cf. [**24**, p. 457])

$$f_{\arg\tau}(0)\overline{g_{\arg\tau}(0)} = \frac{|\tau|^2}{v^2}\,\pi^2\,\widehat{\psi}_1\Big(\frac{u}{2v}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2v}\Big)},$$

where $\psi_1(r) := \frac{1}{2\pi}\int_0^{2\pi} f\big(\sqrt{r}(\cos\zeta,\sin\zeta)\big)\,d\zeta$, $\psi_2(r) := \frac{1}{2\pi}\int_0^{2\pi} g\big(\sqrt{r}(\cos\zeta,\sin\zeta)\big)\,d\zeta$, and $\widehat{\psi}_j(u) := \int_0^\infty e(ur)\psi_j(r)\,dr$. Therefore, using also $h(u) = h(0) + O\big(\|h'\|_{\mathrm{L}^\infty}|u|\big)$,

$$\begin{aligned}\int_{|u|<Bv}\frac{v}{|\tau|^2}f_{\arg\tau}(0)\overline{g_{\arg\tau}(0)}h(u)\,du &= \frac{\pi^2h(0)}{v}\int_{|u|<Bv}\widehat{\psi}_1\Big(\frac{u}{2v}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2v}\Big)}\,du\\ &+ O\left(\frac{\|h'\|_{\mathrm{L}^\infty}}{v}\int_{|u|<Bv}|u|\left|\widehat{\psi}_1\Big(\frac{u}{2v}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2v}\Big)}\right|du\right).\end{aligned} \tag{163}$$

To bound the last error term, first replace the integration variable u by $2vu$; then use the fact that by integration by parts we have $|\widehat{\psi}_j(u)| \ll \int_0^\infty \big(|\psi_j| + |\psi_j'|\big)\,dr \cdot \min(1, |u|^{-1})$. Here, $\int_0^\infty |\psi_1|\,dr \leqslant \frac{1}{\pi}\|f\|_{\mathrm{L}^1}$, while

$$\begin{aligned}\int_0^\infty |\psi_1'(r)|\,dr &\leqslant \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi}\big(\big|(\partial_{x_1}f)\big(\sqrt{r}(\cos\zeta,\sin\zeta)\big)\big| + \big|(\partial_{x_2}f)\big(\sqrt{r}(\cos\zeta,\sin\zeta)\big)\big|\big)\,\frac{d\zeta\,dr}{2\sqrt{r}}\\ &= \frac{1}{2\pi}\int_{\mathbb{R}^2}\big(|(\partial_{x_1}f)(\boldsymbol{x})| + |(\partial_{x_2}f)(\boldsymbol{x})|\big)\,\frac{d\boldsymbol{x}}{\|\boldsymbol{x}\|} \ll S_{\infty,0,1}(f) + S_{1,0,1}(f) \ll \|f\|_{\mathrm{L}^2_6}.\end{aligned}$$

(The next to last bound follows by splitting the domain of integration into the two parts $\{\|\boldsymbol{x}\| \leqslant 1\}$ and $\{\|\boldsymbol{x}\| > 1\}$, and the last bound follows from (135).) Similarly for ψ_2. Hence, the error term in (163) is

$$\ll \|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}\|h'\|_{\mathrm{L}^\infty}\,v\log(B+2).$$

Finally, we are left with

$$\begin{aligned}\frac{\pi^2h(0)}{v}\int_{|u|<Bv}\widehat{\psi}_1\Big(\frac{u}{2v}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2v}\Big)}\,du &= \pi^2h(0)\int_{|u|<B}\widehat{\psi}_1\Big(\frac{u}{2}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2}\Big)}\,du\\ &= \pi^2h(0)\left(\int_{-\infty}^{\infty}\widehat{\psi}_1\Big(\frac{u}{2}\Big)\overline{\widehat{\psi}_2\Big(\frac{u}{2}\Big)}\,du + O\left(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}\int_{|u|>B}|u|^{-2}\,du\right)\right)\end{aligned}$$

$$= 2\pi^2 h(0) \int_0^\infty \psi_1(r)\overline{\psi_2(r)}\, dr + O\Big(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}|h(0)|B^{-1}\Big)$$

$$= \lambda_{f,\overline{g}} h(0) + O\Big(\|f\|_{\mathrm{L}^2_6}\|g\|_{\mathrm{L}^2_6}|h(0)|B^{-1}\Big),$$

where in the next to last equality we used Parseval's identity.

Collecting the above results, and noticing that $v\log(B+2) \ll B^{-1}$ and $B^3 v^2 \leqslant B^{-1}$ because of $1 \leqslant B \leqslant v^{-1/2}$, we obtain the statement of the lemma. □

PROPOSITION 9.10. *Let $f, g \in \mathcal{S}(\mathbb{R}^2)$ and $h \in \mathrm{C}^2(\mathbb{R})$, and assume $S_{\infty,3,2}(h) < \infty$. Let $\boldsymbol{\xi}_2 \in \mathbb{R}^2$ be $[\kappa; c]$-Diophantine. Then for any $v \in (0,1]$,*

$$\begin{aligned}\int_{\mathbb{R}} \Theta_f\Big(u+iv, 0; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\Big)\, \overline{\Theta_g\Big(u+iv, 0; \begin{pmatrix}\mathbf{0}\\ \boldsymbol{\xi}_2\end{pmatrix}\Big)}\, h(u)\, du \\ = \int_{\mathbb{R}^2} f(\boldsymbol{x})\overline{g(\boldsymbol{x})}\, d\boldsymbol{x} \int_{\mathbb{R}} h\, du + \lambda_{f,\overline{g}}\, h(0) \\ + O\Big(\|f\|_{\mathrm{L}^2_{167}}\|g\|_{\mathrm{L}^2_{167}} S_{\infty,3,2}(h)\, \kappa c^{-\frac{1}{\kappa}}\, \delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac{1}{127\kappa}}\Big).\end{aligned} \tag{164}$$

Proof. By [24, Lemma 8.2] (see also [23, Lemma 7.2]),

$$\int_{\Gamma\backslash G} \Theta_f \overline{\Theta_g}\, d\mu = \int_{\mathbb{R}^2} f(\boldsymbol{x})\overline{g(\boldsymbol{x})}\, d\boldsymbol{x}.$$

Therefore, the proposition follows from Lemmas 9.8 and 9.9, applied with

$$B = \tfrac12 v^{-\frac12}\Big(\tfrac14 \delta_{6,\boldsymbol{\xi}_2}\big(v^{-\frac12}\big)^{\frac12}\Big)^{\frac{\kappa}{1+61\kappa}}, \tag{165}$$

as long as this number satisfies $B \geqslant 1$. Indeed, from the observations below (161) we see that the number B in (165) satisfies $B \leqslant v^{-\frac12}$, as is required in Lemma 9.9. Furthermore, using $\delta_{6,\boldsymbol{\xi}_2}(T) \geqslant T^{-1}\ \forall T \geqslant 1$ (as noted in the proof of Lemma 9.8) and $\frac{\kappa}{2(1+61\kappa)} + \frac{1}{127\kappa} < 1$, it follows that $B^{-1} \ll \delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12})^{\frac{1}{127\kappa}}$, so that we indeed obtain the error bound in the last line of (164).

It remains to consider the case when the number B in (165) is less than 1. Using $\delta_{6,\boldsymbol{\xi}_2}(v^{-\frac12}) \geqslant v^{\frac12}$ it then follows that v is bounded below by some positive absolute constant. Hence, by (138) in Lemma 9.2 and the Γ^k invariance of $\Theta_f\overline{\Theta}_g$, we have $\Big|\Theta_f\big(u+iv, 0; \binom{0}{\boldsymbol{\xi}_2}\big)\overline{\Theta_g\big(u+iv, 0; \binom{0}{\boldsymbol{\xi}_2}\big)}\Big| \ll \|f\|_{\mathrm{L}^2_5}\|g\|_{\mathrm{L}^2_5}$ for all $u \in \mathbb{R}$, and so the left-hand side of (164) is $\ll \|f\|_{\mathrm{L}^2_5}\|g\|_{\mathrm{L}^2_5} S_{\infty,2,0}(h)$. Furthermore, from (162) together with $\big|\int_0^{2\pi} f(r\cos\zeta, r\sin\zeta)\, d\zeta\big| \leqslant S_{\infty,2,0}(f)\cdot(1+r)^{-2} \ll \|f\|_{\mathrm{L}^2_2}\cdot(1+r)^{-2}$ (cf. (135)) and the corresponding bound for $\overline{g}$, it follows that $|\lambda_{f,\overline{g}} h(0)| \ll \|f\|_{\mathrm{L}^2_2}\|g\|_{\mathrm{L}^2_2}\|h\|_{\mathrm{L}^\infty}$. Finally, $\big|\int_{\mathbb{R}^2} f(\boldsymbol{x})\overline{g(\boldsymbol{x})}\, d\boldsymbol{x} \int_{\mathbb{R}} h\, du\big| \ll \|f\|_{\mathrm{L}^2}\|g\|_{\mathrm{L}^2} S_{\infty,2,0}(h)$. Hence, (164) holds trivially in this case. □

With Proposition 9.10 established, the proof of Theorem 1.4 can now be completed by a sequence of approximation steps.

Proof of Theorem 1.4. Let $(\alpha, \beta) \in \mathbb{R}^2$ be given as in the statement of the theorem, and set $\boldsymbol{\xi}_2 = \binom{\alpha}{\beta}$. By (130) and Proposition 9.10, writing g_1 and $\overline{g_2}$ in place of f and g, respectively,

we have for any $g_1, g_2 \in \mathcal{S}(\mathbb{R}^2)$, $h \in \mathrm{C}^2(\mathbb{R})$ with $S_{\infty,3,2}(h) < \infty$, and $T \geqslant 1$:

$$\begin{aligned}\frac{1}{T^2} \sum_{\boldsymbol{m}_1 \in \mathbb{Z}^2} \sum_{\boldsymbol{m}_2 \in \mathbb{Z}^2} g_1\big(T^{-1}(\boldsymbol{m}_1 - \boldsymbol{\xi}_2)\big)\, g_2\big(T^{-1}(\boldsymbol{m}_2 - \boldsymbol{\xi}_2)\big)\, \widehat{h}\Big(-\tfrac{1}{2}\, Q\begin{pmatrix}\boldsymbol{m}_1\\ \boldsymbol{m}_2\end{pmatrix}\Big)\\ = \int_{\mathbb{R}^2} g_1(\boldsymbol{x}) g_2(\boldsymbol{x})\, d\boldsymbol{x} \cdot \widehat{h}(0) + \lambda_{g_1,g_2} h(0)\\ + O\Big(\|g_1\|_{\mathrm{L}^2_{167}} \|g_2\|_{\mathrm{L}^2_{167}} S_{\infty,3,2}(h)\, \kappa c^{-\frac{1}{\kappa}}\, \delta^{\frac{1}{127\kappa}}\Big),\end{aligned} \tag{166}$$

where we use the short-hand notation $\delta := \delta_{6,\boldsymbol{\xi}_2}(T)$. Let us consider the contribution from all terms with $\boldsymbol{m}_2 = \boldsymbol{m}_1$ in sum in the left-hand side. Set $G := g_1 g_2 \in \mathcal{S}(\mathbb{R}^2)$. Note that

$$G(T^{-1}\boldsymbol{x}) = \int_{\boldsymbol{x}+[0,1]^2} G(T^{-1}\boldsymbol{y})\, d\boldsymbol{y} + O\Big(\frac{S_{\infty,3,1}(G)}{T(1+T^{-1}\|\boldsymbol{x}\|)^3}\Big), \quad \forall \boldsymbol{x} \in \mathbb{R}^2.$$

Adding this relation over all $\boldsymbol{x} = \boldsymbol{m}_1 - \boldsymbol{\xi}_2$ $(\boldsymbol{m}_1 \in \mathbb{Z}^2)$, and noticing $\sum(1 + T^{-1}\|\boldsymbol{m}_1 - \boldsymbol{\xi}_2\|)^{-3} \ll T^2$, we get

$$\frac{1}{T^2} \sum_{\boldsymbol{m}_1 \in \mathbb{Z}^2} (g_1 g_2)(T^{-1}(\boldsymbol{m}_1 - \boldsymbol{\xi}_2)) = \int_{\mathbb{R}^2} g_1 g_2\, d\boldsymbol{x} + O\big(S_{\infty,3,1}(g_1 g_2)\, T^{-1}\big).$$

We multiply this identity by $\widehat{h}(0)$, and note that the error term is then subsumed by the error term in (166), since $S_{\infty,3,1}(g_1 g_2) \leqslant \sum_{m=0}^{1} S_{\infty,\frac{3}{2},m}(g_1) S_{\infty,\frac{3}{2},1-m}(g_2) \ll \|g_1\|_{L^2_3}\, \|g_2\|_{L^2_3}$ (cf. (135)). Hence, subtracting the resulting identity from (166), we obtain

$$\begin{aligned}\frac{1}{T^2} \sum_{\boldsymbol{m}_1 \in \mathbb{Z}^2} \sum_{\substack{\boldsymbol{m}_2 \in \mathbb{Z}^2\\ \boldsymbol{m}_2 \neq \boldsymbol{m}_1}} g_1\big(T^{-1}(\boldsymbol{m}_1 - \boldsymbol{\xi}_2)\big)\, g_2\big(T^{-1}(\boldsymbol{m}_2 - \boldsymbol{\xi}_2)\big)\, \widehat{h}\Big(-\tfrac{1}{2}\, Q\begin{pmatrix}\boldsymbol{m}_1\\ \boldsymbol{m}_2\end{pmatrix}\Big)\\ = \lambda_{g_1,g_2} h(0) + O\Big(\|g_1\|_{\mathrm{L}^2_{167}} \|g_2\|_{\mathrm{L}^2_{167}} S_{\infty,3,2}(h)\, \kappa c^{-\frac{1}{\kappa}}\, \delta^{\frac{1}{127\kappa}}\Big).\end{aligned} \tag{167}$$

Next, we take g_1, g_2 in (167) to be given by $g_j(\boldsymbol{x}) := f_j(\boldsymbol{x} + T^{-1}\boldsymbol{\xi}_2)$ for some $f_1, f_2 \in \mathcal{S}(\mathbb{R}^2)$. Recall that $\boldsymbol{\xi}_2 = \binom{\alpha}{\beta}$, and by assumption in Theorem 1.4 this vector lies in $[-1,1]^2$. Hence, $\|g_j\|_{\mathrm{L}^2_{167}} \ll \|f_j\|_{\mathrm{L}^2_{167}}$ and $\|g_j - f_j\|_{\mathrm{L}^\infty} \ll S_{\infty,0,1}(f_j)\, T^{-1}$, for $j = 1, 2$. Inspecting the definition of $\lambda_{f,g}$ in (162) it follows that

$$\begin{aligned}|\lambda_{g_1,g_2} - \lambda_{f_1,f_2}| &\ll \big(S_{\infty,3,0}(g_1) S_{\infty,0,1}(f_2) + S_{\infty,0,1}(f_1) S_{\infty,3,0}(f_2)\big)\, T^{-1} \int_0^\infty (1+r)^{-3} r\, dr\\ &\ll \|f_1\|_{\mathrm{L}^2_3} \|f_2\|_{\mathrm{L}^2_3}\, T^{-1}.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{T^2} \sum_{\boldsymbol{m}_1 \in \mathbb{Z}^2} \sum_{\substack{\boldsymbol{m}_2 \in \mathbb{Z}^2\\ \boldsymbol{m}_2 \neq \boldsymbol{m}_1}} f_1(T^{-1}\boldsymbol{m}_1)\, f_2(T^{-1}\boldsymbol{m}_2)\, \widehat{h}\Big(-\tfrac{1}{2}\, Q\begin{pmatrix}\boldsymbol{m}_1\\ \boldsymbol{m}_2\end{pmatrix}\Big)\\ = \lambda_{f_1,f_2} h(0) + O\Big(\|f_1\|_{\mathrm{L}^2_{167}} \|f_2\|_{\mathrm{L}^2_{167}} S_{\infty,3,2}(h)\, \kappa c^{-\frac{1}{\kappa}}\, \delta^{\frac{1}{127\kappa}}\Big).\end{aligned} \tag{168}$$

Next, take h to be given by $h(u) = \frac{1}{2}\widehat{g}(\frac{1}{2}u)$, where g is any function in $\mathrm{C}^3(\mathbb{R})$ with $S_{1,2,3}(g) < \infty$. Then $h^{(j)}(u) = \frac{1}{2}\int_{-\infty}^{\infty} g(s)\,(-\pi i s)^j e(-\frac{1}{2}us)\, ds$ $(j \in \{0,1,2\})$, and so, by integration by parts, $(\pi i u)^m h^{(j)}(u) = \frac{1}{2}\int_{-\infty}^{\infty} \big(\frac{d^m}{ds^m}\big\{g(s)\,(-\pi i s)^j\big\}\big)\, e(-\frac{1}{2}us)\, ds$ $(j \in \{0,1,2\},\ m \in \{0,1,2,3\})$. It follows from these identities that $S_{\infty,3,2}(h) \ll S_{1,2,3}(g)$. Furthermore, by Fourier inversion,

$g(s) = \widehat{h}(-\frac{1}{2}s)$. Let us also write $f_1 \otimes f_2$ for the function in $\mathcal{S}(\mathbb{R}^4)$ given by $(f_1 \otimes f_2)(\substack{\boldsymbol{x}_1 \\ \boldsymbol{x}_2}) = f_1(\boldsymbol{x}_1) f_2(\boldsymbol{x}_2)$. Comparing (11) and (162) we then have $\lambda_{f_1,f_2} = 2\lambda_{f_1 \otimes f_2}$. Comparing also with (10), we obtain

$$N_{\alpha,\beta}(f_1 \otimes f_2, g, T) = \lambda_{f_1 \otimes f_2}\, \widehat{g}(0) + O\Big(\|f_1\|_{\mathrm{L}^2_{167}} \|f_2\|_{\mathrm{L}^2_{167}} S_{1,2,3}(g)\, \kappa c^{-\frac{1}{\kappa}}\, \delta^{\frac{1}{127\kappa}} \Big). \tag{169}$$

It will be useful to note the following consequence of (169).

LEMMA 9.11. *For any $[\kappa; c]$-Diophantine vector $(\alpha, \beta) \in [-1,1]^2$, any $g \in \mathrm{C}^1(\mathbb{R})$ with $S_{1,2,1}(g) < \infty$, and any $R \geqslant 1$,*

$$\sum_{\substack{\boldsymbol{m} \in \mathbb{Z}^4 \setminus \Delta \\ \|\boldsymbol{m}\| \leqslant R}} |g(Q(\boldsymbol{m}))| \ll S_{1,2,1}(g) \Big(1 + \kappa c^{-\frac{1}{\kappa}}\, \delta_{6,\boldsymbol{\xi}_2}(R)^{\frac{1}{127\kappa}} \Big) R^2, \tag{170}$$

where the implied constant is absolute.

Proof. A standard construction shows that there exists a function $\widetilde{g} \in \mathrm{C}^\infty(\mathbb{R})$ satisfying $\widetilde{g} \geqslant |g|$ and $S_{1,2,3}(\widetilde{g}) \ll S_{1,2,1}(g)$, with an absolute implied constant. Fix a choice of a non-negative function $f \in \mathrm{C}^\infty_c(\mathbb{R}^2)$ satisfying $f(\boldsymbol{x}) = 1$ for all $\boldsymbol{x}$ with $\|\boldsymbol{x}\| \leqslant 1$. Applying (169) with $T = R$, $f_1 = f_2 = f$ and $\widetilde{g}$ in place of g gives (cf. (10))

$$\frac{1}{R^2} \sum_{\boldsymbol{m} \in \mathbb{Z}^4 \setminus \Delta} [f \otimes f](R^{-1}\boldsymbol{m}) \widetilde{g}(Q(\boldsymbol{m})) \ll \|\widetilde{g}\|_{\mathrm{L}^1} + S_{1,2,3}(\widetilde{g}) \kappa c^{-\frac{1}{\kappa}}\, \delta_{6,\boldsymbol{\xi}_2}(R)^{\frac{1}{127\kappa}}.$$

Using $\|\widetilde{g}\|_{\mathrm{L}^1} \leqslant S_{1,2,3}(\widetilde{g}) \ll S_{1,2,1}(g)$ and the fact that $[f \otimes f](R^{-1}\boldsymbol{m}) \geqslant 1$ whenever $\|\boldsymbol{m}\| \leqslant R$, we obtain (170). □

REMARK 13. By contrast, if (α, β) is not Diophantine then the left-hand side of (170) may grow more rapidly than R^2 as $R \to \infty$ (cf. [**24**, Section 9]).

We now continue with the proof of Theorem 1.4. Take $f \in \mathrm{C}^1_c(\mathbb{R}^4)$ with support contained in the unit ball centered at the origin. We wish to go from (169) to an asymptotic formula for $N_{\alpha,\beta}(f, g, T)$. Fix, once and for all, a function $\phi \in \mathrm{C}^\infty_c(\mathbb{R}^2)$ with support contained in the unit ball centered at the origin and satisfying $\int_{\mathbb{R}^2} \phi(\boldsymbol{x})\, d\boldsymbol{x} = 1$. Then for an appropriate number $0 < \eta < 1$ (to be fixed below) we define $\phi_\eta \in \mathrm{C}^\infty_c(\mathbb{R}^2)$ by $\phi_\eta(\boldsymbol{x}) := \eta^{-2}\phi(\eta^{-1}\boldsymbol{x})$, and set

$$\widetilde{f} := f * (\phi_\eta \otimes \phi_\eta).$$

Note that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^4$ with $(\phi_\eta \otimes \phi_\eta)(\boldsymbol{x} - \boldsymbol{y}) \neq 0$ one has $\|\boldsymbol{y} - \boldsymbol{x}\| \leqslant \sqrt{2}\,\eta$ and thus $|f(\boldsymbol{y}) - f(\boldsymbol{x})| \ll S_{\infty,0,1}(f) \cdot \eta$; hence

$$\Big| f(\boldsymbol{x}) - \widetilde{f}(\boldsymbol{x}) \Big| \ll S_{\infty,0,1}(f) \cdot \eta, \quad \forall \boldsymbol{x} \in \mathbb{R}^4. \tag{171}$$

Therefore, by (11) and using the fact that the supports of both f and $\widetilde{f}$ are contained in the ball $\{\|\boldsymbol{x}\| \leqslant 3\}$,

$$|\lambda_f - \lambda_{\widetilde{f}}| \ll S_{\infty,0,1}(f) \cdot \eta,$$

and also, by (10),

$$\Big|N_{\alpha,\beta}(f,g,T) - N_{\alpha,\beta}(\widetilde{f},g,T)\Big| \ll \frac{S_{\infty,0,1}(f)\,\eta}{T^2} \sum_{\substack{\boldsymbol{m}\in\mathbb{Z}^4\setminus\Delta\\ \|\boldsymbol{m}\|\leqslant 3T}} |g(Q(\boldsymbol{m}))|$$

$$\ll S_{\infty,0,1}(f) S_{1,2,1}(g)\Big(1+\kappa c^{-\frac{1}{\kappa}}\,\delta_{6,\boldsymbol{\xi}_2}(T)^{\frac{1}{127\kappa}}\Big)\eta,$$

where the last bound follows from Lemma 9.11 and the fact that $\delta_{\beta,\boldsymbol{\xi}}(T)$ is essentially a decreasing function of T, in the sense that

$$\delta_{\beta,\boldsymbol{\xi}}(T') < 2\delta_{\beta,\boldsymbol{\xi}}(T) \quad \text{for any } \ 0<T\leqslant T' \tag{172}$$

(this follows from (6) and the fact that $\frac{1+\log^+ y}{1+y} < 2\frac{1+\log^+ x}{1+x}$ whenever $0<x\leqslant y$).

Next, using $\widetilde{f} := f * (\phi_\eta \otimes \phi_\eta)$ we have

$$N_{\alpha,\beta}(\widetilde{f},g,T) = \int_{\mathbb{R}^4} f(\boldsymbol{y}) N_{\alpha,\beta}\big(\phi_{\eta,\boldsymbol{y}_1}\otimes\phi_{\eta,\boldsymbol{y}_2},g,T\big)\,d\boldsymbol{y},$$

where $\phi_{\eta,\boldsymbol{a}}(\boldsymbol{x}) = \phi_\eta(\boldsymbol{x}-\boldsymbol{a})$ for $\boldsymbol{x},\boldsymbol{a}\in\mathbb{R}^2$, and as usual we write $\boldsymbol{y} = \left(\begin{smallmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{smallmatrix}\right) \in \mathbb{R}^4$ with $\boldsymbol{y}_1,\boldsymbol{y}_2\in\mathbb{R}^2$. Hence, by (169),

$$N_{\alpha,\beta}(\widetilde{f},g,T) = \lambda_{\widetilde{f}}\,\widehat{g}(0) + O\bigg(\int_{\mathbb{R}^4} |f(\boldsymbol{y})| \|\phi_{\eta,\boldsymbol{y}_1}\|_{\mathrm{L}^2_{167}} \|\phi_{\eta,\boldsymbol{y}_2}\|_{\mathrm{L}^2_{167}} S_{1,2,3}(g)\,\kappa c^{-\frac{1}{\kappa}}\,\delta^{\frac{1}{127\kappa}}\,d\boldsymbol{y}\bigg). \tag{173}$$

Here, we have $\|\phi_{\eta,\boldsymbol{b}}\|_{L^2_a} \asymp_a (1+\|\boldsymbol{b}\|)^a \eta^{-a-1}$ $(\forall \boldsymbol{b}\in\mathbb{R}^2)$; hence, using also the assumption about the support of f,

$$\int_{\mathbb{R}^4} |f(\boldsymbol{y})| \|\phi_{\eta,\boldsymbol{y}_1}\|_{\mathrm{L}^2_{167}} \|\phi_{\eta,\boldsymbol{y}_2}\|_{\mathrm{L}^2_{167}}\,d\boldsymbol{y} \ll \|f\|_{\mathrm{L}^\infty}\eta^{-336}.$$

Combining the above bounds, we obtain

$$|N_{\alpha,\beta}(f,g,T) - \lambda_f\,\widehat{g}(0)| \ll \|f\|_{\mathrm{L}^\infty} S_{1,2,3}(g)\kappa c^{-\frac{1}{\kappa}}\,\delta^{\frac{1}{127\kappa}}\eta^{-336}$$

$$+ S_{\infty,0,1}(f) S_{1,2,3}(g)\Big(1+\kappa c^{-\frac{1}{\kappa}}\,\delta^{\frac{1}{127\kappa}}\Big)\eta$$

$$\ll S_{\infty,0,1}(f) S_{1,2,3}(g)\kappa c^{-\frac{1}{\kappa}}\Big(\delta^{\frac{1}{127\kappa}}\eta^{-336}+\eta\Big).$$

Note also that $S_{\infty,0,1}(f) \ll \sum_{j=1}^4 \big\|\partial_{x_j} f\big\|_{\mathrm{L}^\infty}$, since we assume that the support of f is contained in the unit ball. Choosing $\eta = (\frac{1}{9}\,\delta)^{\frac{1}{337\cdot 127\kappa}}$ (this number satisfies $0<\eta<1$, by an observation which we made below (161)), we now obtain the bound in Theorem 1.4 with $B = 42799$. □

9.5. *Consequences of Theorem 1.4*

Let us start by showing that the assumptions in Theorem 1.4 on f having a fixed compact support and (α,β) satisfying $|\alpha|,|\beta|\leqslant 1$, can both be weakened by simple *a posteriori* arguments:

COROLLARY 9.12. *Let $B>0$ be as in Theorem 1.4, and let $\varepsilon>0$. Then for any $[\kappa;c]$-Diophantine vector $\boldsymbol{\xi} = (\alpha,\beta)\in\mathbb{R}^2$, any $f\in \mathrm{C}^1(\mathbb{R}^4)$ with $S_{\infty,3+\varepsilon,1}(f)<\infty$, any $g\in\mathrm{C}^3(\mathbb{R})$ with $S_{1,2,3}(g)<\infty$, and any $T\geqslant \max(1,\varepsilon\|\boldsymbol{\xi}\|)$,*

$$\bigg|N_{\alpha,\beta}(f,g,T) - \lambda_f\int_{\mathbb{R}} g(s)\,ds\bigg| \ll_\varepsilon S_{\infty,3+\varepsilon,1}(f)\,S_{1,2,3}(g)\kappa c^{-\frac{1}{\kappa}}\left(\delta_{6,\boldsymbol{\xi}}(T)^{1/(B\kappa)} + \frac{\|\boldsymbol{\xi}\|}{T}\right). \tag{174}$$

We stress that the implied constant in (174) depends only on ε.

Proof. Let us first keep f as in Theorem 1.4, but allow $\boldsymbol{\xi} = (\alpha, \beta) \in \mathbb{R}^2$ to be outside $[-1,1]^2$. Choose $\boldsymbol{k} \in \mathbb{Z}^2$, so that the vector $(\alpha', \beta') := \boldsymbol{\xi} - \boldsymbol{k}$ lies in $[-1,1]^2$, and so that $\boldsymbol{k} = \mathbf{0}$ if already $\boldsymbol{\xi} \in [-1,1]^2$. Of course (α', β') is $[\kappa; c]$-Diophantine just like $\boldsymbol{\xi}$, and $\delta_{6,(\alpha',\beta')}(T) = \delta_{6,\boldsymbol{\xi}}(T)$ for all T. Recall that the inhomogeneous form Q is defined by (9); let Q' be the corresponding form coming from (α', β'), that is, $Q'(\boldsymbol{x}) \equiv Q(\boldsymbol{x} + (\boldsymbol{k}, \boldsymbol{k}))$. Then

$$N_{\alpha,\beta}(f,g,T) - N_{\alpha',\beta'}(f,g,T) = \frac{1}{T^2} \sum_{\boldsymbol{m} \in \mathbb{Z}^4 \setminus \Delta} \big(f\big(T^{-1}(\boldsymbol{m} + (\boldsymbol{k}, \boldsymbol{k}))\big) - f(T^{-1}\boldsymbol{m}) \big) g(Q'(\boldsymbol{m})).$$

Here,

$$\big| f\big(T^{-1}(\boldsymbol{m} + (\boldsymbol{k}, \boldsymbol{k}))\big) - f(T^{-1}\boldsymbol{m}) \big| \ll \sum_{j=1}^{4} \big\| \partial_{x_j} f \big\|_{\mathrm{L}^\infty} \frac{\|\boldsymbol{k}\|}{T}. \tag{175}$$

Furthermore, since f is supported in the unit ball, the difference in (175) vanishes whenever $\|\boldsymbol{m}\| \geqslant T + \sqrt{2}\|\boldsymbol{k}\|$. Hence, using Lemma 9.11 and $1 + \kappa c^{-\frac{1}{\kappa}} \, \delta_{6,\boldsymbol{\xi}_2}(R) \, \frac{1}{127\kappa} \ll \kappa c^{-\frac{1}{\kappa}}$ (the latter holds because of (6), and since $\kappa \geqslant \frac{1}{2}$ and $0 < c < 2^{-1/2} < 1$), we get

$$|N_{\alpha,\beta}(f,g,T) - N_{\alpha',\beta'}(f,g,T)| \ll \sum_{j=1}^{4} \big\| \partial_{x_j} f \big\|_{\mathrm{L}^\infty} S_{1,2,1}(g) \kappa c^{-\frac{1}{\kappa}} \left(\frac{T + \|\boldsymbol{k}\|}{T} \right)^2 \frac{\|\boldsymbol{k}\|}{T}.$$

Note that $\|\boldsymbol{k}\| \ll \|\boldsymbol{\xi}\|$, and $\|\boldsymbol{\xi}\| \leqslant \varepsilon^{-1} T$ by assumption; thus $\big(\frac{T+\|\boldsymbol{k}\|}{T}\big)^2 \frac{\|\boldsymbol{k}\|}{T} \ll_\varepsilon \frac{\|\boldsymbol{\xi}\|}{T}$. Combining the above with Theorem 1.4 applied to (α', β'), we conclude that

$$\left| N_{\alpha,\beta}(f,g,T) - \lambda_f \int_{\mathbb{R}} g(s)\, ds \right| \ll_\varepsilon \sum_{j=1}^{4} \big\| \partial_{x_j} f \big\|_{\mathrm{L}^\infty} S_{1,2,3}(g) \kappa c^{-\frac{1}{\kappa}} \left(\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)} + \frac{\|\boldsymbol{\xi}\|}{T} \right), \tag{176}$$

for all $T \geqslant \max(1, \varepsilon\|\boldsymbol{\xi}\|)$.

We next wish to extend the bound to more general functions f, as in the statement of the corollary. To achieve this, we will use the fact that both $N_{\alpha,\beta}(f,g,T)$ and λ_f transform in an obvious manner under scaling of the function f. Indeed, introducing the scaling operator δ_R (for any $R > 0$) acting on $\mathrm{C}_c(\mathbb{R}^4)$ through $[\delta_R f](\boldsymbol{x}) := f(R\boldsymbol{x})$, we have by immediate inspection in (10) and (11):

$$N_{\alpha,\beta}(\delta_R f, g, T) = R^{-2} N_{\alpha,\beta}(f, g, T/R) \quad (T > 0) \tag{177}$$

and

$$\lambda_{\delta_R f} = R^{-2} \lambda_f. \tag{178}$$

Now let $f \in \mathrm{C}^1(\mathbb{R}^4)$ with $S_{\infty,3+\varepsilon,1}(f) < \infty$ be given. We will decompose f dyadically radially, using a partition of unity. Fix a C^∞ function $\varphi : \mathbb{R} \to [0,1]$ satisfying $\varphi(r) = 0$ for $r \leqslant 0.1$ and $\varphi(r) = 1$ for $r \geqslant 0.9$, and then define the C^∞-functions $\varphi_0, \varphi_1, \ldots : \mathbb{R} \to [0,1]$ through $\varphi_0(r) = 1 - \varphi(r-1)$ and

$$\varphi_j(r) = \begin{cases} \varphi(r - 2^{j-1}) & \text{if } r < 2^j \\ 1 - \varphi(r - 2^j) & \text{if } r \geqslant 2^j \end{cases} \quad (j = 1, 2, 3, \ldots,\ r \in \mathbb{R}).$$

Then $\operatorname{supp} \varphi_0 \subset (-\infty, 2)$ and $\operatorname{supp} \varphi_j \subset (2^{j-1}, 2^j + 1)$ for all $j \geqslant 1$; furthermore

$$\sum_{j=0}^{\infty} \varphi_j(r) = 1 \quad (\forall r \in \mathbb{R}), \text{ and } \|\varphi_j'\|_{\mathrm{L}^\infty} = \|\varphi'\|_{\mathrm{L}^\infty} \quad (j = 1, 2, 3, \ldots).$$

Then define $f_j \in \mathrm{C}_c^1(\mathbb{R}^4)$ through $f_j(\boldsymbol{x}) := \varphi_j(\|\boldsymbol{x}\|)f(\boldsymbol{x})$. Then $f(\boldsymbol{x}) = \sum_{j=0}^{\infty} f_j(\boldsymbol{x})$, and it follows that

$$N_{\alpha,\beta}(f,g,T) = \sum_{j=0}^{\infty} N_{\alpha,\beta}(f_j,g,T) \quad \text{and} \quad \lambda_f = \sum_{j=1}^{\infty} \lambda_{f_j}. \tag{179}$$

(To prove the first relation one uses (10); the change of order of summation is justified since we have absolute convergence; $\sum_{j=0}^{\infty}\sum_{\boldsymbol{m}\in\mathbb{Z}^4\setminus\Delta}\big|f_j(T^{-1}\boldsymbol{m})g(Q(\boldsymbol{m}))\big| < \infty$. This absolute convergence follows from $\|f_j\|_{\mathrm{L}^\infty} \ll S_{\infty,3,0}(f)2^{-3j}$ and the fact that the support of f_j is contained in the ball of radius 2^{j+1} about the origin, combined with the bound

$$\sum_{\substack{\boldsymbol{m}\in\mathbb{Z}^4\setminus\Delta\\ \|\boldsymbol{m}\|\leqslant S}} \big|g(Q(\boldsymbol{m}))\big| \ll S^2$$

for S large, which follows from (176) by the argument in the proof of Lemma 9.11. The justification of the second relation in (179) is similar but easier.)

We also set

$$\widetilde{f}_j := \delta_{2^{j+1}} f_j.$$

Then $\widetilde{f}_j \in \mathrm{C}_c^1(\mathbb{R}^4)$ and the support of $\widetilde{f}_j$ is contained in the unit ball centered at $\mathbf{0}$. Hence, (176) applies to $\widetilde{f}_j$, yielding

$$\left|N_{\alpha,\beta}(\widetilde{f}_j,g,T) - \lambda_{\widetilde{f}_j}\int_{\mathbb{R}} g(s)\,ds\right| \ll_\varepsilon \sum_{k=1}^{4}\left\|\partial_{x_k}\widetilde{f}_j\right\|_{\mathrm{L}^\infty} S_{1,2,3}(g)\kappa c^{-\frac{1}{\kappa}}\left(\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)} + \frac{\|\boldsymbol{\xi}\|}{T}\right) \tag{180}$$

for all $T \geqslant \max(1, \varepsilon\|\boldsymbol{\xi}\|)$. Here,

$$\sum_{k=1}^{4}\left\|\partial_{x_k}\widetilde{f}_j\right\|_{\mathrm{L}^\infty} = 2^{j+1}\sum_{k=1}^{4}\|\partial_{x_k} f_j\|_{\mathrm{L}^\infty} \ll S_{\infty,3+\varepsilon,1}(f)\cdot 2^{-(2+\varepsilon)j}. \tag{181}$$

Indeed, for $\boldsymbol{x}\neq\mathbf{0}$ we have $\big|\partial_{x_k}\varphi_j(\|\boldsymbol{x}\|)\big| = \big|\varphi_j'(\|\boldsymbol{x}\|)x_k\big|/\|\boldsymbol{x}\| \leqslant \|\varphi'\|_{\mathrm{L}^\infty}$, while at $\boldsymbol{x}=\mathbf{0}$ $\partial_{x_k}\varphi_j(\|\boldsymbol{x}\|)$ vanishes; hence

$$\sum_{k=1}^{4}\|\partial_{x_k} f_j\|_{\mathrm{L}^\infty} \ll \sup\left\{|f(\boldsymbol{x})| + \sum_{k=1}^{4}|\partial_{x_k} f(\boldsymbol{x})| \;:\; \boldsymbol{x}\in\mathbb{R}^4,\ \|\boldsymbol{x}\|\in\operatorname{supp}\varphi_j\right\},$$

and (181) follows since $\|\boldsymbol{x}\|\in\operatorname{supp}\varphi_j$ implies $1+\|\boldsymbol{x}\| \asymp 2^j$. Combining (180) and (181) with (177) and (178), we obtain

$$\begin{aligned}&\left|N_{\alpha,\beta}(f_j,g,2^{-j-1}T) - \lambda_{f_j}\int_{\mathbb{R}} g(s)\,ds\right|\\ &\quad\ll 2^{-\varepsilon j} S_{\infty,3+\varepsilon,1}(f)\,S_{1,2,3}(g)\kappa c^{-\frac{1}{\kappa}}\left(\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)} + \frac{\|\boldsymbol{\xi}\|}{T}\right).\end{aligned} \tag{182}$$

This holds for all $T \geqslant \max(1,\varepsilon\|\boldsymbol{\xi}\|)$. We replace T by $2^{j+1}T$ in (182), use (172), and finally add over all j, using (179). This gives (174). □

Finally, we give the proofs of Corollaries 1.5 and 1.6.

Proof of Corollary 1.5. Let $\chi : \mathbb{R}^4 \to \{0,1\}$ be the characteristic function of the unit ball and let $\chi_{(a,b)} : \mathbb{R}\to\{0,1\}$ be the characteristic function of the interval (a,b). For $\eta,\eta' > 0$

(two constants which we will fix below) we choose $f_\pm \in \mathrm{C}_c^\infty(\mathbb{R}^4)$, so that $0 \leqslant f_- \leqslant \chi \leqslant f_+ \leqslant 1$ and $f_-(\boldsymbol{x}) = 1$ whenever $\|\boldsymbol{x}\| \leqslant 1-\eta$ and $f_+(\boldsymbol{x}) = 0$ whenever $\|\boldsymbol{x}\| \geqslant 1+\eta$, and we choose $g_\pm \in \mathrm{C}_c^\infty(\mathbb{R})$, so that $0 \leqslant g_- \leqslant \chi_{(a,b)} \leqslant g_+ \leqslant 1$ and $g_-(s) = 1$ whenever $a+\eta' \leqslant s \leqslant b-\eta'$ and $g_+(s) = 0$ whenever $s \leqslant a-\eta'$ or $s \geqslant b+\eta'$. (Thus, if $\eta > 1$, we may take $f_- \equiv 0$ and if $\eta' > \frac{1}{2}(b-a)$ we may take $g_- \equiv 0$.) These functions can be chosen, so that $S_{\infty,4,1}(f_\pm) \ll \eta^{-1}$ and $S_{1,2,3}(g_\pm) \ll \big(1+|a|+|b|\big)^2\big(b-a+\eta'^{-2}\big)$, so long as $\eta, \eta' \ll 1$. By construction, we have

$$N_{\alpha,\beta}(f_-, g_-, T) \leqslant N_{\alpha,\beta}(a,b,T) \leqslant N_{\alpha,\beta}(f_+, g_+, T).$$

We also have $\big|\lambda_{f_\pm} - \lambda_\chi\big| \ll \eta$ and $\lambda_\chi = \frac{\pi^2}{2}$, thus $\lambda_{f_\pm} = \frac{\pi^2}{2} + O(\eta)$. Hence, by Corollary 9.12,

$$\begin{aligned}
&\Big|N_{\alpha,\beta}(a,b,T) - \tfrac{\pi^2}{2}(b-a)\Big| \\
&\quad\ll (b-a)\eta + \eta' + \eta^{-1}(1+|a|+|b|)^2\Big(b-a+\eta'^{-2}\Big)\kappa c^{-\frac{1}{\kappa}}\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)} \\
&\quad\ll (1+|a|+|b|)^3\kappa c^{-\frac{1}{\kappa}}\Big(\eta + \eta' + \eta^{-1}\eta'^{-2}\delta_{6,(\alpha,\beta)}(T)^{1/(B\kappa)}\Big). \qquad (183)
\end{aligned}$$

Choosing $\eta = \eta' = \delta_{6,(\alpha,\beta)}(T)^{1/(4B\kappa)}$, we obtain (14), with $B' = 4B$. □

Remark 14. Of course, the bound in (183) is often wasteful regarding the dependence on a, b. However, recall that we have to keep $\eta, \eta' \ll 1$ in order for the first bound in (183) to be valid, and our main aim in Corollary 1.5 was to give a reasonably simple statement of a general bound with an absolute implied constant, and with a power rate decay with respect to T for any fixed (α, β) subject to a Diophantine condition.

Proof of Corollary 1.6. This can again be derived from Theorem 1.4 by an approximation argument; however, it is easier to argue directly from (167), since there $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ appear shifted by $\boldsymbol{\xi}_2$, which is exactly what we need. Indeed, let $\chi : \mathbb{R}^2 \to \{0,1\}$ be the characteristic function of the open unit ball centered at the origin and let $\chi_{(-b/2,-a/2)}$ be the characteristic function of the interval $(-b/2, -a/2)$; then for $g_1 = g_2 = \chi$ and $\widehat{h} = \chi_{(-b/2,-a/2)}$, the left-hand side of (167) is exactly equal to $\pi R_2[a,b](T^2)$ (cf. (15)). Now the corollary follows by a similar approximation argument as in the proof of Corollary 1.5. □

Acknowledgement. We would like to thank Sanju Velani for helpful discussions regarding Lemma 3.1.

References

1. M. V. Berry and M. Tabor, 'Level clustering in the regular spectrum', *Proc. Roy. Soc. A* 356 (1977) 375–394.
2. J. D. Bovey and M. M. Dodson, 'The Hausdorff dimension of systems of linear forms', *Acta Arith.* 45 (1986) 337–358.
3. T. Browning and I. Vinogradov, 'Effective Ratner theorem for $\mathrm{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$ and gaps in $\sqrt{n}$ modulo 1', *J. Lond. Math. Soc.* 94 (2016) 61–84.
4. M. Burger, 'Horocycle flow on geometrically finite surfaces', *Duke Math. J.* 61 (1990) 779–803.
5. C. Dettmann, J. Marklof and A. Strömbergsson, 'Universal hitting time statistics for integrable flows', *J. Stat. Phys.* 166 (2017) 714–749.
6. M. Einsiedler, G. Margulis and A. Venkatesh, 'Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces', *Invent. Math.* 177 (2009) 137–212.
7. A. Eskin, G. Margulis and S. Mozes, 'Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture', *Ann. of Math.* 147 (1998) 93–141.
8. A. Eskin, G. Margulis and S. Mozes, 'Quadratic forms of signature (2,2) and eigenvalue spacings on rectangular 2-tori', *Ann. of Math.* 161 (2005) 679–725.
9. L. Flaminio and G. Forni, 'Invariant distributions and time averages for horocycle flows', *Duke Math. J.* 119 (2003) 465–526.
10. G. Folland, *Real analysis* (John Wiley & Sons Inc., New York, 1999).

11. F. Götze and G. Margulis, 'Distribution of values of quadratic forms at integral points', Preprint, 2010, arXiv:1004.5123.
12. L. Grafakos, *Classical Fourier analysis* (Springer, Berlin, 2008).
13. B. Green and T. Tao, 'The quantitative behaviour of polynomial orbits on nilmanifolds', *Ann. of Math.* 175 (2012) 465–540.
14. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers* (Clarendon Press, Oxford, 1938).
15. H. Iwaniec and E. Kowalski, *Analytic number theory* (American Mathematical Society, Providence, RI, 2004).
16. A. Khintchine, 'Zwei Bemerkungen zu einer Arbeit des Herrn Perron', *Math. Z.* 22 (1925) 274–284.
17. A. Khintchine, 'Zur metrischen Theorie der diophantischen Approximationen', *Math. Z.* 24 (1926) 706–714.
18. H. H. Kim, 'Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2', *J. Amer. Math. Soc.* 16 (2003) 139–183 (electronic). (With appendix 1 by D.Ramakrishnan and appendix 2 by Kim and P. Sarnak.)
19. A. W. Knapp, *Lie groups beyond an introduction*, Progress in Mathematics 140 (Birkhäuser Boston Inc., Boston, MA, 2002).
20. E. Lindenstrauss and G. Margulis, 'Effective estimates on indefinite ternary forms', *Israel J. Math.* 203 (2014) 445–499.
21. G. A. Margulis, 'Indefinite quadratic forms and unipotent flows on homogeneous spaces', *Dynamical systems and ergodic theory*, vol. 23 (Banach Center Publ., PWN, Warsaw, 1989) 399–409.
22. G. A. Margulis and A. Mohammadi, 'Quantitative version of the Oppenheim conjecture for inhomogeneous quadratic forms', *Duke Math. J.* 158 (2011) 121–160.
23. J. Marklof, 'Pair correlation densities of inhomogeneous quadratic forms. II', *Duke Math. J.* 115 (2002) 409–434.
24. J. Marklof, 'Pair correlation densities of inhomogeneous quadratic forms', *Ann. of Math.* 158 (2003) 419–471.
25. J. Marklof, 'Mean square value of exponential sums related to the representation of integers as sums of squares', *Acta Arith.* 117 (2005) 353–370.
26. T. Miyake, *Modular forms* (Springer, Berlin, 1989).
27. A. Mohammadi, 'A special case of effective equidistribution with explicit constants', *Ergodic Theory Dynam. Systems* 32 (2012) 237–247.
28. M. B. Nathanson, *Additive number theory*, Graduate Texts in Mathematics 164 (Springer, Berlin, 1996).
29. O. Perron, 'Über diophantische Approximationen', *Math. Ann.* 83 (1921) 77–84.
30. T. Prinyasart, 'An effective equidistribution of diagonal translates of certain orbits in $\mathrm{ASL}(3,\mathbb{Z})\backslash\mathrm{ASL}(3,\mathbb{R})$', PhD Thesis, University of California San Diego, San Diego, CA, 2018.
31. M. Ratner, 'On Raghunathan's measure conjecture', *Ann. of Math.* 134 (1991) 545–607.
32. M. Ratner, 'Raghunathan's topological conjecture and distributions of unipotent flows', *Duke Math. J.* 63 (1991) 235–280.
33. W. M. Schmidt, 'On simultaneous approximations of two algebraic numbers by rationals', *Acta Math.* 119 (1967) 27–50.
34. W. M. Schmidt, 'Simultaneous approximation to algebraic numbers by rationals', *Acta Math.* 125 (1970) 189–201.
35. A. Selberg, 'On the estimation of Fourier coefficients of modular forms', Proceedings of Symposia in Pure Mathematics 8 (American Mathematical Society, Providence, RI, 1965) 1–15.
36. N. Shah, 'Limit distributions of expanding translates of certain orbits on homogeneous spaces', *Proc. Indian Acad. Sci. Math. Sci.* 106 (1996) 105–125.
37. A. Strömbergsson, 'On the deviation of ergodic averages for horocycle flows', *J. Mod. Dyn.* 7 (2013) 291–328.
38. A. Strömbergsson, 'Effective Ratner equidistribution for $\mathrm{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$', *Duke Math. J.* 164 (2015) 843–902.
39. A. Weil, 'On some exponential sums', *Proc. Natl. Acad. Sci. USA* 34 (1948) 204–207.

Andreas Strömbergsson
Department of Mathematics
Uppsala University
Box 480
Uppsala SE-75106
Sweden

astrombe@math.uu.se

Pankaj Vishe
Department of Mathematical Sciences
Durham University
Durham DH1 3LE
United Kingdom

pankaj.vishe@durham.ac.uk

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.