Acta Math., 198 (2007), 231–298
DOI: 10.1007/s11511-007-0016-x

Discrete Radon transforms and applications to ergodic theory

by

Alexandru D. Ionescu
University of Wisconsin, Madison
Madison, WI, U.S.A.

Akos Magyar
University of Georgia, Athens
Athens, GA, U.S.A.

Elias M. Stein
Princeton University
Princeton, NJ, U.S.A.

Stephen Wainger
University of Wisconsin, Madison
Madison, WI, U.S.A.

Contents

A. Ionescu was supported in part by an NSF grant, a Sloan fellowship and a Packard fellowship. A. Magyar was supported in part by an NSF grant. E. M. Stein was supported in part by an NSF grant. S. Wainger was supported in part by an NSF grant.

1. Introduction

In this paper we are concerned with L^p estimates for discrete operators in certain non-translation-invariant settings, and the applications of such estimates to ergodic theorems for certain families of non-commuting operators. We first describe the type of operators that we consider in the translation-invariant setting. Assume that $P: \mathbf{Z}^{d_1} \to \mathbf{Z}^{d_2}$ is a polynomial mapping and $K: \mathbf{R}^{d_1} \setminus B(1) \to \mathbf{C}$ is a Calderón–Zygmund kernel (see formulas (1.3) and (1.4) for precise definitions). For (compactly supported) functions $f: \mathbf{Z}^{d_2} \to \mathbf{C}$, we define the maximal operator

$$\widetilde{M}(f)(m) = \sup_{r>0} \left| \frac{1}{|B(r) \cap \mathbf{Z}^{d_1}|} \sum_{n \in B(r) \cap \mathbf{Z}^{d_1}} f(m - P(n)) \right|,$$

and the singular integral operator

$$\widetilde{T}(f)(m) = \sum_{n \in \mathbf{Z}^{d_1} \setminus \{0\}} K(n) f(m - P(n)).$$

The maximal operator $\widetilde{M}(f)$ was considered by Bourgain [3], [4], [5], who showed that

$$\|\widetilde{M}(f)\|_{L^p(\mathbf{Z}^{d_2})} \leqslant C_p \|f\|_{L^p(\mathbf{Z}^{d_2})}, \quad p \in (1, \infty], \quad \text{if } d_1 = d_2 = 1. \tag{1.1}$$

Maximal inequalities such as (1.1) have applications to pointwise and L^p, $p \in (1, \infty)$, ergodic theorems; see [3], [4] and [5]. A typical theorem is the following: assume that $P: \mathbf{Z} \to \mathbf{Z}$ is a polynomial mapping, (X, μ) is a finite measure space and $T: X \to X$ is a measure-preserving invertible transformation. For $F \in L^p(X)$, $p \in (1, \infty)$, let

$$\tilde{A}_r(F)(x) = \frac{1}{2r+1} \sum_{|n| \leqslant r} F(T^{P(n)} x) \quad \text{for any } r \in \mathbf{Z}_+.$$

Then there is a function $F_* \in L^p(X)$ with the property that

$$\lim_{r \to \infty} \tilde{A}_r(F) = F_* \quad \text{almost everywhere and in } L^p.$$

In addition, $F_* = \mu(X)^{-1} \int_X F(x)\, d\mu$ if T^q is ergodic for $q = 1, 2, \ldots$.

The related singular integral operator $\widetilde{T}(f)$ was considered first by Arkhipov and Oskolkov [1] and by Stein and Wainger [15]. Following earlier work of [1], [15] and [17], Ionescu and Wainger [8] proved that

$$\|\widetilde{T}(f)\|_{L^p(\mathbf{Z}^{d_2})} \leqslant C_p \|f\|_{L^p(\mathbf{Z}^{d_2})}, \quad p \in (1, \infty). \tag{1.2}$$

A more complete description of the results leading to the bound (1.2) can be found in the introduction of [8].

In this paper, we start the systematic study of the suitable analogues of the operators $\widetilde{M}$ and $\widetilde{T}$ in discrete settings which are not translation-invariant.(1) As before, the maximal function estimate has applications to ergodic theorems involving families of non-commuting operators.

Motivated by models involving actions of nilpotent groups, we consider a special class of non-translation-invariant Radon transforms, called the "quasi-translation" invariant Radon transforms. Assume that $d, d' \geqslant 1$ and let $P: \mathbf{Z}^d \times \mathbf{Z}^d \to \mathbf{Z}^{d'}$ be a polynomial mapping. For any $r>0$ let $B(r)$ denote the ball $\{x \in \mathbf{R}^d : |x| < r\}$. Let $K: \mathbf{R}^d \setminus B(1) \to \mathbf{C}$ denote a Calderón–Zygmund kernel, i.e.

$$|x|^d |K(x)| + |x|^{d+1} |\nabla K(x)| \leqslant 1, \quad |x| \geqslant 1, \tag{1.3}$$

and

$$\left| \int_{|x| \in [1,N]} K(x)\, dx \right| \leqslant 1 \quad \text{for any } N \geqslant 1. \tag{1.4}$$

For (compactly supported) functions $f: \mathbf{Z}^d \times \mathbf{Z}^{d'} \to \mathbf{C}$ we define the discrete maximal Radon transform

$$M(f)(m_1, m_2) = \sup_{r>0} \left| \frac{1}{|B(r) \cap \mathbf{Z}^d|} \sum_{n \in B(r) \cap \mathbf{Z}^d} f(m_1 - n, m_2 - P(m_1, n)) \right|, \tag{1.5}$$

and the discrete singular Radon transform

$$T(f)(m_1, m_2) = \sum_{n \in \mathbf{Z}^d \setminus \{0\}} K(n) f(m_1 - n, m_2 - P(m_1, n)). \tag{1.6}$$

The operator T was considered by Stein and Wainger [16], who proved that

$$\|T\|_{L^2(\mathbf{Z}^d \times \mathbf{Z}^{d'}) \to L^2(\mathbf{Z}^d \times \mathbf{Z}^{d'})} \leqslant C. \tag{1.7}$$

In this paper, we prove estimates like (1.7) in the full range of exponents p for both the singular integral operator T and the maximal operator M, in the special case in which

$$\text{the polynomial } P \text{ has degree at most 2.} \tag{1.8}$$

THEOREM 1.1. *Assuming condition* (1.8), *the discrete maximal Radon transform M extends to a bounded* (*subadditive*) *operator on $L^p(\mathbf{Z}^d \times \mathbf{Z}^{d'})$, $p \in (1, \infty]$, with*

$$\|M\|_{L^p(\mathbf{Z}^d \times \mathbf{Z}^{d'}) \to L^p(\mathbf{Z}^d \times \mathbf{Z}^{d'})} \leqslant C_p.$$

The constant C_p depends only on the exponent p and the dimension d.

---

(1) Such operators, called Radon transforms, have been studied extensively in the continuous setting; see [6] and the references therein.

THEOREM 1.2. *Assuming condition* (1.8), *the discrete singular Radon transform* T *extends to a bounded operator on* $L^p(\mathbf{Z}^d\times\mathbf{Z}^{d'})$, $p\in(1,\infty)$, *with*

$$\|T\|_{L^p(\mathbf{Z}^d\times\mathbf{Z}^{d'})\to L^p(\mathbf{Z}^d\times\mathbf{Z}^{d'})}\leqslant C_p.$$

The constant C_p *depends only on the exponent* p *and the dimension* d.

See also Theorems 2.1–2.4 and 5.2 for equivalent versions of Theorems 1.1 and 1.2 in the setting of nilpotent groups. In the special case $d=d'=1$ and $P(m_1,n)=n^2$, Theorem 1.1 gives

$$\left\|\sup_{r>0}\frac{1}{|B(r)\cap\mathbf{Z}|}\sum_{|n|\leqslant r}|f(m_1-n,m_2-n^2)|\right\|_{L^p(\mathbf{Z}^2)}\leqslant C_p\|f\|_{L^p(\mathbf{Z}^2)} \tag{1.9}$$

for any $p\in(1,\infty]$ and $f\in L^p(\mathbf{Z}^2)$. We consider functions f of the form

$$f(m_1,m_2)=g(m_2)\,\mathbf{1}_{[-M,M]}(m_1);$$

by letting $M\to\infty$, it follows from (1.9) that

$$\left\|\sup_{r>0}\frac{1}{|B(r)\cap\mathbf{Z}|}\sum_{|n|\leqslant r}|g(m-n^2)|\right\|_{L^p(\mathbf{Z})}\leqslant C_p\|g\|_{L^p(\mathbf{Z})},$$

which is Bourgain's theorem [5] in the case $P(n)=n^2$.

We now state our main ergodic theorem. Let (X,μ) denote a finite measure space, and let $T_1,...,T_d,S_1,...,S_{d'}$ denote a family of measure-preserving invertible transformations on X satisfying the commutator relations

$$[T_j,S_k]=[S_j,S_k]=I \quad\text{and}\quad [[T_j,T_k],T_l]=I \quad \text{for all } j,\ k \text{ and } l. \tag{1.10}$$

Here I denotes the identity transformation and $[T,S]=T^{-1}S^{-1}TS$ the commutator of T and S. For a polynomial mapping

$$Q=(Q_1,...,Q_{d'})\colon\mathbf{Z}^d\to\mathbf{Z}^{d'} \quad \text{of degree at most 2,} \tag{1.11}$$

and $F\in L^p(X)$, $p\in(1,\infty)$, we define the averages

$$A_r(F)(x)=\frac{1}{|B(r)\cap\mathbf{Z}^d|}\sum_{n=(n_1,...,n_d)\in B(r)\cap\mathbf{Z}^d}F(T_1^{n_1}\,...\,T_d^{n_d}S_1^{Q_1(n)}\,...\,S_{d'}^{Q_{d'}(n)}\,x). \tag{1.12}$$

THEOREM 1.3. *Assume that* $T_1, ..., T_d, S_1, ..., S_{d'}$ *satisfy* (1.10) *and let* Q *be as in* (1.11). *Then, for every* $F \in L^p(X)$, $p \in (1, \infty)$, *there exists* $F_* \in L^p(X)$ *such that*

$$\lim_{r\to\infty} A_r(F) = F_* \quad \textit{almost everywhere and in } L^p. \tag{1.13}$$

Moreover, if the family of transformations $\{T_j^q, S_k^q : 1 \leqslant j \leqslant d \text{ and } 1 \leqslant k \leqslant d'\}$ *is ergodic for every integer* $q \geqslant 1$, *then*

$$F_* = \frac{1}{\mu(X)} \int_X F \, d\mu. \tag{1.14}$$

See also Theorem 5.1 for an equivalent version formulated in terms of the action of a discrete nilpotent group of step 2.

It would be desirable to remove the restrictions on the degrees of the polynomials P and Q in (1.8) and (1.11), and allow more general commutator relations in (1.10).([2]) These two issues are related. In this paper we exploit the restriction (1.8) to connect the Radon transforms M and T to certain group translation-invariant Radon transforms on discrete nilpotent groups of step 2. We then analyze the resulting Radon transforms using Fourier analysis techniques. The analogue of this construction for higher degree polynomials P leads to nilpotent Lie groups of higher step, for which it is not clear whether the Fourier transform method can be applied. We hope to return to this in the future.

We describe now some of the ingredients in the proofs of Theorems 1.1–1.3. In §2 we use a transference principle and reduce Theorems 1.1 and 1.2 to Lemmas 2.7 and 2.8 on the discrete nilpotent group $\mathbf{G}_0^\#$.

In §3 we prove four technical lemmas concerning oscillatory integrals on $L^2(\mathbf{Z}_q^d)$ and $L^2(\mathbf{Z}^d)$. These bounds correspond to estimates for fixed θ after using the Fourier transform in the central variable of the group $\mathbf{G}_0^\#$. We remark that natural scalar-valued objects, such as the Gauss sums, become operator-valued objects in our non-commutative setting. For example, the bound $\|\mathcal{S}^{a/q}\|_{L^2(\mathbf{Z}_q^d) \to L^2(\mathbf{Z}_q^d)} \leqslant q^{-1/2}$ in Lemma 3.1 is the natural analogue of the standard scalar bound on Gauss sums $|S^{a/q}| \leqslant Cq^{-1/2}$.

In §4 we prove Lemma 2.7 (which implies Theorem 1.1). In §4.1 we prove certain strong L^2 bounds (see Lemma 4.1); the proof of these L^2 bounds is based on a variant of the "circle method", adapted to our non-translation-invariant setting. In §4.2 we prove a restricted L^p bound, $p>1$, with a logarithmic loss. The idea of using such restricted L^p estimates as an ingredient for proving the full L^p estimates originates in Bourgain's paper [5]. Finally, in §4.3 we prove Lemma 2.7, by combining the strong L^2 bounds in §4.1, and the restricted L^p bounds in §4.2.

([2]) A possible setting for the pointwise ergodic theorem would be that of polynomial sequences in nilpotent groups; compare with [2] and [9].

In §5 we prove Theorem 1.3. First we restate Theorem 1.3 in terms of actions of discrete nilpotent groups of step 2, see Theorem 5.1. Then we use a maximal ergodic theorem, which follows by transference from Theorem 1.1, to reduce matters to proving almost everywhere convergence for functions F in a dense subset of $L^p(X)$. For this we adapt a limiting argument of Bourgain [5].

In §6 we prove Lemma 2.8 (which implies Theorem 1.2). In §6.1 we prove strong L^2 bounds, using only Plancherel's theorem and the fixed θ estimates in §3. In §6.2 we recall (without proofs) a partition of the integers and a square function estimate used by Ionescu and Wainger [8]. In §6.3 we complete the proof of Lemma 2.8. First we reduce matters to proving a suitable square function estimate for a more standard oscillatory singular integral operator (see Lemma 6.6). Then we use the equivalence between square function estimates and weighted inequalities (cf. [7, Chapter V]) to further reduce to proving a weighted inequality for an (essentially standard) oscillatory singular integral operator. This weighted inequality is proved in §7.

In §7, which is self-contained, we collect several estimates related to the real-variable theory on the group $\mathbf{G}_0^\#$. We prove weighted L^p estimates for maximal averages and oscillatory singular integrals, in which the relevant underlying balls have eccentricity $N\gg 1$. The main issue is to prove these L^p bounds with only logarithmic losses of the type $(\log N)^C$. These logarithmic losses can then be combined with the gains of $N^{-\bar{c}}$ in the L^2 estimates in Lemmas 4.1 and 6.1 to obtain the theorems in the full range of exponents p. The proofs in this section are essentially standard real-variable proofs (compare with [14]); we provide all the details for the sake of completeness.

2. Preliminary reductions: a transference principle

In this section we reduce Theorems 1.1 and 1.2 to Lemmas 2.7 and 2.8 on the discrete free group $\mathbf{G}_0^\#$ defined below. This is based on the "method of transference" (see, for example, [11]). Since the polynomial mapping P in Theorems 1.1 and 1.2 has degree at most 2 (see condition (1.8)), we can write

$$P(m_1,n)=R(n,m_1-n)+A(m_1-n)+B(m_1), \tag{2.1}$$

for some polynomial mappings $A,B\colon\mathbf{Z}^d\to\mathbf{Z}^{d'}$ and a bilinear mapping $R\colon\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$. The representation (2.1) follows simply by setting

$$\begin{aligned}B(m)&=P(m,m),\\ A(m)&=P(m,0)-P(m,m),\\ R(m,m')&=P(m+m',m)+P(m',m')-P(m+m',m+m')-P(m',0).\end{aligned}$$

Since $R(m,0)=R(0,m')=0$ for any $m,m'\in\mathbf{Z}^d$, it follows from (1.8) that R is bilinear.

Definitions (1.5) and (1.6) show that

$$M(f)(m_1,m_2)=\widetilde{M}(f_A)(m_1,m_2-B(m_1)),$$
$$T(f)(m_1,m_2)=\widetilde{T}(f_A)(m_1,m_2-B(m_1)),$$

where $f_A(m_1,m_2)=f(m_1,m_2-A(m_1))$, and $\widetilde{M}$ and $\widetilde{T}$ are defined in the same way as M and T, by replacing $P(m_1,n)$ with $R(n,m_1-n)$. Therefore, in proving Theorems 1.1 and 1.2 we may assume that $P(m_1,n)=R(n,m_1-n)$, where R is a bilinear mapping. In this case, the operators M and T can be viewed as group translation-invariant operators on certain nilpotent Lie groups, which we define below.

Assume that $d,d'\geqslant 1$ are integers and $R:\mathbf{R}^d\times\mathbf{R}^d\to\mathbf{R}^{d'}$ is a bilinear map. We define the nilpotent Lie group

$$\mathbf{G}=\{(x,s)\in\mathbf{R}^d\times\mathbf{R}^{d'}:(x,s)\cdot(y,t)=(x+y,s+t+R(x,y))\},\tag{2.2}$$

with the standard unimodular Haar measure $dx\,ds$. In addition, if

$$R(\mathbf{Z}^d\times\mathbf{Z}^d)\subseteq\mathbf{Z}^{d'},\tag{2.3}$$

then the set

$$\mathbf{G}^{\#}=\mathbf{Z}^d\times\mathbf{Z}^{d'}\subseteq\mathbf{G}\tag{2.4}$$

is a discrete subgroup of $\mathbf{G}$, equipped with the counting Haar measure.

For any (bounded compactly supported) function $F:\mathbf{G}\to\mathbf{C}$ we define the discrete maximal Radon transform

$$\mathcal{M}(F)(x,s)=\sup_{r>0}\left|\frac{1}{|B(r)\cap\mathbf{Z}^d|}\sum_{n\in B(r)\cap\mathbf{Z}^d}F((n,0)^{-1}\cdot(x,s))\right|,\tag{2.5}$$

and the discrete singular Radon transform

$$\mathcal{T}(F)(x,s)=\sum_{n\in\mathbf{Z}^d\setminus\{0\}}K(n)F((n,0)^{-1}\cdot(x,s)).\tag{2.6}$$

Assuming condition (2.3), for (compactly supported) functions $f:\mathbf{G}^{\#}\to\mathbf{C}$, we define

$$\mathcal{M}^{\#}(f)(m,u)=\sup_{r>0}\left|\frac{1}{|B(r)\cap\mathbf{Z}^d|}\sum_{n\in B(r)\cap\mathbf{Z}^d}f((n,0)^{-1}\cdot(m,u))\right|\tag{2.7}$$

and

$$\mathcal{T}^{\#}(f)(m,u)=\sum_{n\in\mathbf{Z}^d\setminus\{0\}}K(n)f((n,0)^{-1}\cdot(m,u)).\tag{2.8}$$

In view of equation (2.1), Theorems 1.1 and 1.2 follow from Theorems 2.1 and 2.2 below.

THEOREM 2.1. *Assume that* $R:\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$ *is a bilinear map satisfying condition* (2.3). *Then the discrete maximal Radon transform* $\mathcal{M}^\#$ *extends to a bounded* (*subadditive*) *operator on* $L^p(\mathbf{G}^\#)$, $p\in(1,\infty]$, *with*

$$\|\mathcal{M}^\#(f)\|_{L^p(\mathbf{G}^\#)}\leqslant C_p\|f\|_{L^p(\mathbf{G}^\#)}.$$

The constant C_p *depends only on the exponent* p *and the dimension* d.

THEOREM 2.2. *Assume that* $R:\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$ *is a bilinear map satisfying condition* (2.3). *Then the discrete singular Radon transform* $\mathcal{T}^\#$ *extends to a bounded operator on* $L^p(\mathbf{G}^\#)$, $p\in(1,\infty)$, *with*

$$\|\mathcal{T}^\#(f)\|_{L^p(\mathbf{G}^\#)}\leqslant C_p\|f\|_{L^p(\mathbf{G}^\#)}.$$

The constant C_p *depends only on the exponent* p *and the dimension* d.

Theorems 2.1 and 2.2 can be restated as theorems on the Lie group $\mathbf{G}$.

THEOREM 2.3. *Assume that* $R:\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$ *is a bilinear map. Then the discrete maximal Radon transform* $\mathcal{M}$ *extends to a bounded* (*subadditive*) *operator on* $L^p(\mathbf{G})$, $p\in(1,\infty]$, *with*

$$\|\mathcal{M}(F)\|_{L^p(\mathbf{G})}\leqslant C_p\|F\|_{L^p(\mathbf{G})}.$$

The constant C_p *may depend only on the exponent* p *and the dimension* d.

THEOREM 2.4. *Assume that* $R:\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$ *is a bilinear map. Then the discrete singular Radon transform* $\mathcal{T}$ *extends to a bounded operator on* $L^p(\mathbf{G})$, $p\in(1,\infty)$, *with*

$$\|\mathcal{T}(F)\|_{L^p(\mathbf{G})}\leqslant C_p\|F\|_{L^p(\mathbf{G})}.$$

The constant C_p *may depend only on the exponent* p *and the dimension* d.

Assuming condition (2.3), we now justify the equivalence of Theorems 2.3 and 2.1 and Theorems 2.4 and 2.2. We notice that the map $\Phi:\mathbf{G}^\#\times[0,1)^d\times[0,1)^{d'}\to\mathbf{G}$,

$$\Phi((m,u),(\mu,\alpha))=(m,u)\cdot(\mu,\alpha)=(m+\mu,u+\alpha+R(m,\mu)),$$

establishes a measure-preserving bijection between $\mathbf{G}^\#\times[0,1)^d\times[0,1)^{d'}$ and $\mathbf{G}$. For any (compactly supported) function $f:\mathbf{G}^\#\to\mathbf{C}$ we define

$$F:\mathbf{G}\longrightarrow\mathbf{C},\quad F(\Phi((m,u),(\mu,\alpha)))=f(m,u).$$

The definitions show that for any $(\mu,\alpha)\in[0,1)^d\times[0,1)^{d'}$,

$$\begin{aligned}\mathcal{M}^\#(f)(m,u)&=\mathcal{M}(F)(\Phi((m,u),(\mu,\alpha))),\\ \mathcal{T}^\#(f)(m,u)&=\mathcal{T}(F)(\Phi((m,u),(\mu,\alpha))).\end{aligned}$$

Thus Theorem 2.3 implies Theorem 2.1 and Theorem 2.4 implies Theorem 2.2.

For the converse, assume that $F:\mathbf{G}\to\mathbf{C}$ is given. For any $(\mu,\alpha)\in[0,1)^d\times[0,1)^{d'}$ we define

$$f_{(\mu,\alpha)}:\mathbf{G}^{\#}\longrightarrow\mathbf{C},\quad f_{(\mu,\alpha)}(m,u)=F(\Phi((m,u),(\mu,\alpha))).$$

The definitions show that

$$\mathcal{M}(F)(\Phi((m,u),(\mu,\alpha)))=\mathcal{M}^{\#}(f_{(\mu,\alpha)})(m,u),$$
$$\mathcal{T}(F)(\Phi((m,u),(\mu,\alpha)))=\mathcal{T}^{\#}(f_{(\mu,\alpha)})(m,u),$$

so Theorem 2.1 implies Theorem 2.3 and Theorem 2.2 implies Theorem 2.4.

We further reduce Theorems 2.3 and 2.4 to a special "universal" case. We define the bilinear map $R_0:\mathbf{R}^d\times\mathbf{R}^d\to\mathbf{R}^{d^2}$ by

$$R_0(x,y)=\sum_{l_1,l_2=1}^{d} x_{l_1}y_{l_2}e_{l_1l_2}, \tag{2.9}$$

where $\{e_{l_1l_2}:l_1,l_2=1,...,d\}$ denotes the standard orthonormal basis of R^{d^2}. Using the bilinear map R_0, we define the nilpotent Lie group $\mathbf{G}_0$ as in (2.2). For any (bounded compactly supported) function $F:\mathbf{G}_0\to\mathbf{C}$, we define $\mathcal{M}_0(F)$ and $\mathcal{T}_0(F)$ as in (2.5) and (2.6).

LEMMA 2.5. *The discrete maximal Radon transform $\mathcal{M}_0$ extends to a bounded operator on $L^p(\mathbf{G}_0)$, $p\in(1,\infty]$.*

LEMMA 2.6. *The discrete singular Radon transform $\mathcal{T}_0$ extends to a bounded operator on $L^p(\mathbf{G}_0)$, $p\in(1,\infty)$.*

We now show that Lemmas 2.5 and 2.6 imply Theorems 2.3 and 2.4, respectively. Assume that the bilinear map R in the definition of the group $\mathbf{G}$ is

$$R(x,y)=\sum_{l_1,l_2=1}^{d} x_{l_1}y_{l_2}v_{l_1l_2},$$

for some vectors $v_{l_1l_2}\in\mathbf{R}^{d'}$. We define the linear map $L:\mathbf{R}^{d^2}\to\mathbf{R}^{d'}$ by $L(e_{l_1l_2})=v_{l_1l_2}$ (so $L(R_0(x,y))=R(x,y)$ for any $x,y\in\mathbf{R}^d$) and the group morphism

$$\tilde{L}:\mathbf{G}_0\longrightarrow\mathbf{G},\quad \tilde{L}(x,s)=(x,L(s)).$$

We define the isometric representation π of $\mathbf{G}_0$ on $L^p(\mathbf{G})$, $p\in[1,\infty]$, by

$$\pi(g_0)(F)(g)=F(\tilde{L}(g_0^{-1})\cdot g),\quad g_0\in\mathbf{G}_0, F\in L^p(\mathbf{G}), g\in\mathbf{G}.$$

For $r>0$ we define the generalized measures μ_r and ν_r on $C_c(\mathbf{G}_0)$ by

$$\mu_r(F_0)=\frac{1}{|B(r)\cap\mathbf{Z}^d|}\sum_{n\in B(r)\cap\mathbf{Z}^d}F_0(n,0),$$
$$\nu_r(F_0)=\sum_{n\in B(r)\cap\mathbf{Z}^d\setminus\{0\}}K(n)F_0(n,0).$$

Clearly, for any (bounded compactly supported) function $F_0\colon\mathbf{G}_0\to\mathbf{C}$,

$$\mathcal{M}_0(F_0)(g_0)=\sup_{r>0}|F_0*\mu_r(g_0)|,$$
$$\mathcal{T}_0(F_0)(g_0)=\lim_{r\to\infty}F_0*\nu_r(g_0).$$

Moreover, the definitions show that for any (bounded compactly supported) function $F\colon\mathbf{G}\to\mathbf{C}$,

$$\mathcal{M}(F)(g)=\sup_{r>0}\left|\int_{G_0}[\pi(g_0)(F)](g)\,d\mu_r(g_0)\right|,$$
$$\mathcal{T}(F)(g)=\lim_{r\to\infty}\int_{G_0}[\pi(g_0)(F)](g)\,d\nu_r(g_0).$$

By [12, Proposition 5.1], we have that Theorems 2.3 and 2.4 follow from Lemmas 2.5 and 2.6, respectively.

Finally, we define the discrete subgroup $\mathbf{G}_0^{\#}=\mathbf{Z}^d\times\mathbf{Z}^{d^2}\subseteq\mathbf{G}_0$. Then we define the operators $\mathcal{M}_0^{\#}$ and $\mathcal{T}_0^{\#}$ as in (2.7) and (2.8):

$$\mathcal{M}_0^{\#}(f)(m,u)=\sup_{r>0}\left|\frac{1}{|B(r)\cap\mathbf{Z}^d|}\sum_{n\in B(r)\cap\mathbf{Z}^d}f((n,0)^{-1}\cdot(m,u))\right|$$

and

$$\mathcal{T}_0^{\#}(f)(m,u)=\sum_{n\in\mathbf{Z}^d\setminus\{0\}}K(n)f((n,0)^{-1}\cdot(m,u)),$$

for (compactly supported) functions $f\colon\mathbf{G}_0^{\#}\to\mathbf{C}$. In view of the equivalence discussed earlier (since R_0 clearly satisfies condition (2.3)), it suffices to prove the following two lemmas.

LEMMA 2.7. *The discrete maximal Radon transform $\mathcal{M}_0^{\#}$ extends to a bounded operator on $L^p(\mathbf{G}_0^{\#})$, $p\in(1,2]$.*

LEMMA 2.8. *The discrete singular Radon transform $\mathcal{T}_0^{\#}$ extends to a bounded operator on $L^p(\mathbf{G}_0^{\#})$, $p\in(1,\infty)$.*

We remark that in Lemma 2.8 it suffices to prove the estimate for $p\in[2,\infty)$. Indeed, assume that $p\in(1,2]$, $p'=p/(p-1)\in[2,\infty)$, and let $\widetilde{K}(n,v)=K(n)\mathbf{1}_{\{0\}}(v)$, $\widetilde{K}\colon\mathbf{G}_0^\#\to\mathbf{C}$. Then $\mathcal{T}_0^\#(f)=f*\widetilde{K}$ and, by duality,

$$\|\mathcal{T}_0^\#\|_{L^p(\mathbf{G}_0^\#)\to L^p(\mathbf{G}_0^\#)}=\sup_{\|f\|_{L^{p'}(\mathbf{G}_0^\#)}=1}\biggl\|\int_{\mathbf{G}_0^\#}f(h\cdot g)\widetilde{K}(h)\,dh\biggr\|_{L_g^{p'}(\mathbf{G}_0^\#)}. \tag{2.10}$$

We define now the "dual" group $\mathbf{G'}_0^\#$:

$$\mathbf{G'}_0^\#=\{(m,u)\in\mathbf{Z}^d\times\mathbf{R}^{d^2}:(m,u)\cdot(n,v)=(m+n,u+v+R_0'(m,n))\},$$

where $R_0'(m,n)=R_0(n,m)=\sum_{l_1,l_2=1}^d m_{l_1}n_{l_2}e_{l_2l_1}$. The right-hand side of equation (2.10) is equal to

$$\sup_{\|f\|_{L^{p'}(\mathbf{G'}_0^\#)}=1}\biggl\|\int_{\mathbf{G'}_0^\#}f(g\cdot h)\widetilde{K}(h)\,dh\biggr\|_{L_g^{p'}(\mathbf{G'}_0^\#)}=\sup_{\|f\|_{L^{p'}(\mathbf{G'}_0^\#)}=1}\|f*_{\mathbf{G'}_0^\#}\widetilde{K}\|_{L^{p'}(\mathbf{G'}_0^\#)}. \tag{2.11}$$

We use now the bijection

$$\mathbf{G}_0^\#\longleftrightarrow\mathbf{G'}_0^\#,\quad\Bigl(m,\sum_{l_1,l_2}u_{l_1l_2}e_{l_1l_2}\Bigr)\longleftrightarrow\Bigl(m,\sum_{l_1,l_2}u_{l_1l_2}e_{l_2l_1}\Bigr).$$

Since $p'\in[2,\infty)$, it follows from Lemma 2.8 that

$$\|f*_{\mathbf{G'}_0^\#}\widetilde{K}\|_{L^{p'}(\mathbf{G'}_0^\#)}\leqslant C_{p'}\|f\|_{L^{p'}(\mathbf{G'}_0^\#)}.$$

Using (2.10) and (2.11), it follows that $\|\mathcal{T}_0^\#\|_{L^p(\mathbf{G}_0^\#)\to L^p(\mathbf{G}_0^\#)}\leqslant C_p$, as desired.

3. Oscillatory integrals on $L^2(\mathbf{Z}_q^d)$ and $L^2(\mathbf{Z}^d)$

In this section we prove four lemmas concerning oscillatory integrals on L^2. The bounds in these lemmas depend on a fixed parameter θ in the Fourier space corresponding to taking the Fourier transform in the central variable of the group $\mathbf{G}_0^\#$. In Lemma 3.1, $\theta=a/q$ (the Gauss sum operator). In Lemma 3.2, θ is close to a/q, q large. In Lemma 3.3, θ is close to a/q, q small. Finally, Lemma 3.4 is an estimate for a singular integral. The main issue in all these lemmas is to have a quantitative gain over the trivial $L^2\to L^2$ estimates with bound 1. Lemmas of this type have been used in [10] and [16].

We assume throughout this section that $d'=d^2$, and that $\mathbf{G}_0^\#$ is the discrete nilpotent group defined in §2. For any $\mu\geqslant1$ let $\mathbf{Z}_\mu=\mathbf{Z}\cap[1,\mu]$. If $a=(a_{l_1l_2})_{l_1,l_2=1,\dots,d}\in\mathbf{Z}^{d'}$ is a vector and $q\geqslant1$ is an integer, then we denote by (a,q) the greatest common divisor of a and q, i.e. the largest integer $q'\geqslant 1$ that divides q and all the components $a_{l_1l_2}$. Any number in $\mathbf{Q}^{d'}$ can be written uniquely in the form

$$a/q, \quad q\in\{1,2,...\}, a\in\mathbf{Z}^{d'}, (a,q)=1. \tag{3.1}$$

A number as in (3.1) will be called an *irreducible d'-fraction*. For any irreducible d'-fraction a/q and $g\colon\mathbf{Z}_q^d\to\mathbf{C}$ we consider the (Gauss sum) operator

$$\mathcal{S}^{a/q}(g)(m)=q^{-d}\sum_{n\in\mathbf{Z}_q^d} g(n)e^{-2\pi iR_0(m-n,n)\cdot a/q}. \tag{3.2}$$

LEMMA 3.1. (Gauss sum estimate) *With the notation above*,

$$\|\mathcal{S}^{a/q}(g)\|_{L^2(\mathbf{Z}_q^d)}\leqslant q^{-1/2}\|g\|_{L^2(\mathbf{Z}_q^d)}. \tag{3.3}$$

Proof. We consider the operator $\mathcal{S}^{a/q}(\mathcal{S}^{a/q})^*$; the kernel of this operator is

$$L(m,n)=q^{-2d}\sum_{w\in\mathbf{Z}_q^d} e^{-2\pi iR_0(m-n,w)\cdot a/q}=q^{-2d}\prod_{l_2=1}^{d}\delta_q\biggl(\sum_{l_1=1}^{d}(m_{l_1}-n_{l_1})\cdot a_{l_1l_2}\biggr), \tag{3.4}$$

where $\delta_q\colon\mathbf{Z}\to\{0,q\}$,

$$\delta_q(m)=\begin{cases} q, & \text{if } m/q\in\mathbf{Z},\\ 0, & \text{if } m/q\notin\mathbf{Z}.\end{cases} \tag{3.5}$$

We have to show that $\sum_{m\in\mathbf{Z}_q^d}|L(m,n)|$ and $\sum_{n\in\mathbf{Z}_q^d}|L(m,n)|$ are bounded uniformly by q^{-1}. In view of equation (3.4), it suffices to prove that the number of solutions $(m_1,...,m_d)\in\mathbf{Z}_q^d$ of the system

$$\sum_{l_1=1}^{d} m_{l_1}a_{l_1l_2}=0 \pmod q \quad \text{for any } l_2=1,...,d, \tag{3.6}$$

is at most q^{d-1}.

Assume that $q=p_1^{\alpha_1}...p_k^{\alpha_k}$ is the unique decomposition of q as a product of powers of distinct primes. Any integer m can be written uniquely in the form

$$m=\sum_{j=1}^{k} m^j\cdot(q/p_j^{\alpha_j}) \pmod q, \quad m^j\in\mathbf{Z}_{p_j^{\alpha_j}}. \tag{3.7}$$

We write $a_{l_1l_2}$ and m_{l_1} as in (3.7). Since the primes p_j are distinct, the system (3.6) is equivalent to the system

$$\sum_{l_1=1}^{d} m_{l_1}^j a_{l_1l_2}^j=0 \pmod{p_j^{\alpha_j}} \quad \text{for any } l_2=1,...,d \text{ and } j=1,...,k. \tag{3.8}$$

We now use the fact that a/q is an irreducible d'-fraction. Thus for any $j=1,...,k$ there are some $l_1(j), l_2(j)\in\{1,...,d\}$ with the property that $(a_{l_1(j)l_2(j)}, p_j)=1$. For any $j=1,...,k$ we consider only the equation in the system (3.8) corresponding to $l_2=l_2(j)$. Since $a_{l_1(j)l_2(j)}$ is invertible in the ring $\mathbf{Z}/p_j^{\alpha_j}\mathbf{Z}$, for any fixed j the system (3.8) can have at most $[p_j^{\alpha_j}]^{d-1}$ solutions $(m_1^j,...,m_d^j)\in\mathbf{Z}^d_{p_j^{\alpha_j}}$. The lemma follows. □

Assume now that $j\geqslant 0$ is an integer and $\Phi_j\colon\mathbf{R}^d\to\mathbf{C}$ is a function supported in the set $\{x:|x|\leqslant 2^{j+1}\}$ such that

$$2^{dj}|\Phi_j(x)|+2^{(d+1)j}|\nabla\Phi_j(x)|\leqslant 1, \quad x\in\mathbf{R}^d. \tag{3.9}$$

For $\theta\in\mathbf{R}^{d'}$ and (compactly supported) functions $g\colon\mathbf{Z}^d\to\mathbf{C}$ we define

$$\mathcal{U}_j^\theta(g)(m)=\sum_{n\in\mathbf{Z}^d}\Phi_j(m-n)g(n)e^{-2\pi i R_0(m-n,n)\cdot\theta}. \tag{3.10}$$

We prove two L^2 bounds for the operators $\mathcal{U}_j^\theta(g)$.

Lemma 3.2. (Minor arcs) *Assume that a/q is an irreducible d'-fraction, $\delta>0$ and $\theta\in\mathbf{R}^{d'}$. Assume also that there are some indices $k_1,k_2\in\{1,...,d\}$ with the property that*

$$\begin{gathered} a_{k_1k_2}/q=\bar{a}_{k_1k_2}/\bar{q}_{k_1k_2}, \quad (\bar{a}_{k_1k_2},\bar{q}_{k_1k_2})=1, \\ 2^{\delta j}\leqslant\bar{q}_{k_1k_2}\leqslant 2^{(2-\delta)j} \quad \textit{and} \quad |\theta_{k_1k_2}-\bar{a}_{k_1k_2}/\bar{q}_{k_1k_2}|\leqslant 2^{-2j}. \end{gathered} \tag{3.11}$$

Then

$$\|\mathcal{U}_j^\theta(g)\|_{L^2(\mathbf{Z}^d)}\leqslant C2^{-\delta' j}\|g\|_{L^2(\mathbf{Z}^d)}, \quad \delta'>0. \tag{3.12}$$

Proof. Clearly, we may assume that $j\geqslant C$. The kernel of the operator $\mathcal{U}_j^\theta(\mathcal{U}_j^\theta)^*$ is

$$L_j^\theta(m,n)=\sum_{w\in\mathbf{Z}^d}\Phi_j(m-w)\bar{\Phi}_j(n-w)e^{-2\pi i R_0(m-n,w)\cdot\theta}. \tag{3.13}$$

Notice that the kernel L_j^θ is supported in the set $\{(m,n):|m-n|\leqslant 2^{j+2}\}$ and the sum in equation (3.13) is taken over $|w-m|\leqslant 2^{j+1}$. Let $A_{l_2}(m)=\sum_{l_1=1}^d m_{l_1}\theta_{l_1l_2}$. We write $w=(w_{k_2},w')$. It follows from equation (3.13) that

$$|L_j^\theta(m,n)|\leqslant\sum_{w'\in\mathbf{Z}^{d-1}}\left|\sum_{w_{k_2}\in\mathbf{Z}}\Phi_j(m-(w_{k_2},w'))\bar{\Phi}_j(n-(w_{k_2},w'))e^{-2\pi i w_{k_2}\cdot A_{k_2}(m-n)}\right|. \tag{3.14}$$

By summation by parts, it is easy to see that

$$\left|\sum_{v\in\mathbf{Z}}e^{-2\pi i v\xi}h(v)\right|\leqslant C\varrho(\xi)^{-1}\|h'\|_{L^1}$$

for any $h\in C^1(\mathbf{R})$, where $\varrho(\xi)$ denotes the distance from the real number ξ to $\mathbf{Z}$. Using inequality (3.9), it follows that

$$|L_j^\theta(m,n)|\leqslant C2^{-dj}\mathbf{1}_{[0,2^{j+2}]}(|m-n|)[1+2^j\varrho(A_{k_2}(m-n))]^{-1}. \tag{3.15}$$

We estimate $\sum_{n\in\mathbf{Z}^d}|L_j^\theta(m,n)|$ and $\sum_{m\in\mathbf{Z}^d}|L_j^\theta(m,n)|$. We write $m=(m_{k_1},m')$ and $n=(n_{k_1},n')$. Using the bound (3.15),

$$\sum_{n\in\mathbf{Z}^d}|L_j^\theta(m,n)|+\sum_{m\in\mathbf{Z}^d}|L_j^\theta(m,n)|\leqslant C2^{-j}\sup_{\mu\in\mathbf{R}}\sum_{v=-2^{j+2}}^{2^{j+2}}[1+2^j\varrho(\theta_{k_1k_2}v+\mu)]^{-1}. \tag{3.16}$$

Thus, for the bound (3.12), it suffices to prove that for some constants $C\geqslant 1$ and $\delta'>0$,

$$\#\{v\in[-2^{j+2},2^{j+2}]\cap\mathbf{Z}:\varrho(\theta_{k_1k_2}v+\mu)\leqslant C^{-1}2^{-(1-\delta')j}\}\leqslant C2^{(1-\delta')j} \tag{3.17}$$

for any $\mu\in\mathbf{R}$ and $j\geqslant C$. Since $|\theta_{k_1k_2}-\bar{a}_{k_1k_2}/\bar{q}_{k_1k_2}|\leqslant 2^{-2j}$ (see conditions (3.11)), we may replace $\theta_{k_1k_2}$ by $\bar{a}_{k_1k_2}/\bar{q}_{k_1k_2}$ in the bound (3.17). We have two cases: if $\bar{q}_{k_1k_2}\geqslant 2^{j+4}$, then the set of points $\{\bar{a}_{k_1k_2}v/\bar{q}_{k_1k_2}:v\in[-2^{j+2},2^{j+2}]\cap\mathbf{Z}\}$ is a subset of the set $\{b/\bar{q}_{k_1k_2}:b\in\mathbf{Z}\}$ and $\bar{a}_{k_1k_2}v/\bar{q}_{k_1k_2}-\bar{a}_{k_1k_2}v'/\bar{q}_{k_1k_2}\notin\mathbf{Z}$ if $v\neq v'\in[-2^{j+2},2^{j+2}]\cap\mathbf{Z}$. Using coniditions (3.11), $\bar{q}_{k_1k_2}\leqslant 2^{(2-\delta)j}$. Thus the number of points in $\{b/\bar{q}_{k_1k_2}:b\in\mathbf{Z}/\bar{q}_{k_1k_2}\mathbf{Z}\}$ that lie in an interval of length $C^{-1}2^{-(1-\delta')j}$ is at most $\bar{q}_{k_1k_2}C^{-1}2^{-(1-\delta')j}+1\leqslant C2^{(1-\delta')j}$, as desired.

Assume now that $\bar{q}_{k_1k_2}\leqslant 2^{j+4}$. We divide the interval $[-2^{j+2},2^{j+2}]$ into at most $C2^j/\bar{q}_{k_1k_2}$ intervals J of length $\leqslant\bar{q}_{k_1k_2}/2$. By the same argument as before,

$$\#\{v\in J\cap\mathbf{Z}:\varrho(\bar{a}_{k_1k_2}v/\bar{q}_{k_1k_2}+\mu)\leqslant C^{-1}2^{-(1-\delta')j}\}\leqslant\bar{q}_{k_1k_2}C^{-1}2^{-(1-\delta')j}+1,$$

for any of these intervals J and any $\mu\in\mathbf{R}$. The bound (3.17) follows since $2^{\delta j}\leqslant\bar{q}_{k_1k_2}$, see conditions (3.11). $\square$

LEMMA 3.3. (Major arcs) *Assume that a/q is an irreducible d'-fraction, $\theta\in\mathbf{R}^{d'}$,*

$$q\leqslant 2^{j/4}\quad\textit{and}\quad|\theta-a/q|\leqslant 2^{-7j/4}. \tag{3.18}$$

Then

$$\|\mathcal{U}_j^\theta(g)\|_{L^2(\mathbf{Z}^d)}\leqslant Cq^{-1/2}(1+2^{2j}|\theta-a/q|)^{-1/4}\|g\|_{L^2(\mathbf{Z}^d)}. \tag{3.19}$$

Proof. We may assume that $j\geqslant C$ and let $\theta=a/q+\xi$. Since R_0 is bilinear, we may assume that the functions g and $\mathcal{U}_j^\theta(g)$ are supported in the ball $\{m:|m|\leqslant C2^j\}$. We write

$$m=qm'+\mu\quad\text{and}\quad n=qn'+\nu,$$

with $\mu,\nu\in\mathbf{Z}_q^d$ and $|m'|,|n'|\leqslant C2^j/q$, and identify $\mathbf{Z}^d$ with $\mathbf{Z}^d\times\mathbf{Z}_q^d$ using these maps. Since R_0 is bilinear, it follows from inequalities (3.9) and (3.18) that

$$\begin{aligned}&\Phi_j(m-n)e^{-2\pi i R_0(m-n,n)\cdot\theta}\\&\qquad=[q^d\Phi_j(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}][q^{-d}e^{-2\pi i R_0(\mu-\nu,\nu)\cdot a/q}]+E(m,n),\end{aligned}\tag{3.20}$$

where $|E(m,n)|\leqslant C2^{-j/2}2^{-dj}\mathbf{1}_{[0,2^{j+3}]}(|m-n|)$. The operator defined by this error term is bounded on L^2 with bound $C2^{-j/2}$, which suffices. Let $\widetilde{\mathcal{U}}_j^\theta$ denote the operator defined by the first term in equation (3.20), i.e.

$$\begin{aligned}&\widetilde{\mathcal{U}}_j^\theta(g)(m',\mu)\\&\quad=\sum_{n'\in\mathbf{Z}^d}\sum_{\nu\in[\mathbf{Z}_q]^d}g(n',\nu)[q^d\Phi_j(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}][q^{-d}e^{-2\pi i R_0(\mu-\nu,\nu)\cdot a/q}]\\&\quad=\sum_{n'\in\mathbf{Z}^d}\mathcal{S}^{a/q}(g)(n',\mu)q^d\Phi_j(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}.\end{aligned}\tag{3.21}$$

In view of Lemma 3.1, for the bound (3.19) it suffices to prove that

$$\biggl\|\sum_{n'\in\mathbf{Z}^d}g'(n')q^d\Phi_j(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}\biggr\|_{L^2(\mathbf{Z}^d)}\leqslant C(1+2^{2j}|\xi|)^{-1/4}\|g'\|_{L^2(\mathbf{Z}^d)},$$

for any (compactly supported) function $g'\colon\mathbf{Z}^d\to\mathbf{C}$. Using the restriction (3.18), it suffices to prove that

$$\|\mathcal{U}_j^\xi(g)\|_{L^2(\mathbf{Z}^d)}\leqslant C(1+2^{2j}|\xi|)^{-1/4}\|g\|_{L^2(\mathbf{Z}^d)},\quad\text{if }|\xi|\leqslant 2^{-5j/4}.\tag{3.22}$$

In proving the bound (3.22) we may assume that $|\xi|\geqslant C2^{-2j}$ (and that j is large). Fix $k_1,k_2\in\{1,...,d\}$ with the property that $|\xi_{k_1k_2}|\geqslant C^{-1}|\xi|$. We repeat the $\mathcal{U}_j^\xi(\mathcal{U}_j^\xi)^*$ argument from Lemma 3.2. In view of inequality (3.16), it suffices to prove that

$$2^{-j}\sup_{\mu\in\mathbf{R}}\sum_{v=-2^{j+2}}^{2^{j+2}}[1+2^j\varrho(\xi_{k_1k_2}v+\mu)]^{-1}\leqslant C(2^{2j}|\xi|)^{-1/2},\tag{3.23}$$

provided that $|\xi_{k_1k_2}|\in[2^{-2j},2^{-5j/4}]$ (see inequality (3.22)). The points

$$\{\xi_{k_1k_2}v+\mu : v\in[-2^{j+2},2^{j+2}]\cap\mathbf{Z}\}$$

lie in an interval of length $\frac{1}{2}$. We partition this interval into $C2^j$ subintervals of length 2^{-j}. Each of these subintervals contains at most $C(2^j|\xi_{k1k2}|)^{-1}$ of the points in the set $\{\xi_{k_1k_2}v+\mu:v\in[-2^{j+2},2^{j+2}]\cap\mathbf{Z}\}$. An easy rearrangement argument then shows that the sum in the left-hand side of inequality (3.23) is dominated by

$$C2^{-j}(2^j|\xi_{k1k2}|)^{-1}\sum_{k\in[1,C2^{2j}|\xi_{k1k2}|]\cap\mathbf{Z}}k^{-1},$$

which proves the bound (3.23). □

Our last lemma in this section concerns Calderón–Zygmund kernels. Assume that

$$K_j : \mathbf{R}^d \longrightarrow \mathbf{C}, \quad j \geqslant 1,$$

are kernels as in (6.1) and (6.2). For any finite set $I \subseteq \{1, ...\}$ we define

$$K^I = \sum_{j \in I} K_j. \tag{3.24}$$

For $\theta \in \mathbf{R}^{d'}$ and (compactly supported) functions $g: \mathbf{Z}^d \to \mathbf{C}$ we define

$$\mathcal{V}_I^\theta(g)(m) = \sum_{n \in \mathbf{Z}^d} K^I(m-n) g(n) e^{-2\pi i R_0(m-n,n)\cdot\theta}. \tag{3.25}$$

Lemma 3.4. *Assume that a/q is an irreducible d'-fraction*, $\theta \in \mathbf{R}^{d'}$ *and*

$$I \subseteq \{j : q^8 \leqslant 2^{2j} \leqslant |\theta - a/q|^{-1}\}. \tag{3.26}$$

Then

$$\|\mathcal{V}_I^\theta(g)\|_{L^2(\mathbf{Z}^d)} \leqslant C q^{-1/2} \|g\|_{L^2(\mathbf{Z}^d)}. \tag{3.27}$$

Proof. Let $\theta = a/q + \xi$. Since R_0 is bilinear, we may assume that the functions g and $\mathcal{V}_I^\theta(g)$ are supported in the ball $\{m : |m| \leqslant C|\xi|^{-1/2}\}$. As in Lemma 3.3, we write

$$m = qm' + \mu \quad \text{and} \quad n = qn' + \nu,$$

with $\mu, \nu \in \mathbf{Z}_q^d$ and $|m'|, |n'| \leqslant C|\xi|^{-1/2}/q$, and identify $\mathbf{Z}^d$ with $\mathbf{Z}^d \times \mathbf{Z}_q^d$ using these maps. Since R_0 is bilinear, it follows from inclusion (3.26) that

$$\begin{aligned} &K^I(m-n) e^{-2\pi i R_0(m-n,n)\cdot\theta} \\ &\quad = [q^d K^I(q(m'-n')) e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}][q^{-d} e^{-2\pi i R_0(\mu-\nu,\nu)\cdot a/q}] + E'(m,n), \end{aligned} \tag{3.28}$$

where $|E'(m,n)| \leqslant Cq|m-n|^{-d-1/2} \mathbf{1}_{[q^4/2, 2|\xi|^{-1/2}]}(|m-n|)$. The operator defined by this error term is bounded on L^2 with bound Cq^{-1}, which suffices. Let $\widetilde{\mathcal{V}}_I^\theta$ denote the operator defined by the first term in equation (3.28), i.e.

$$\begin{aligned} &\widetilde{\mathcal{V}}_I^\theta(g)(m', \mu) \\ &\quad = \sum_{n' \in \mathbf{Z}^d} \sum_{\nu \in \mathbf{Z}_q^d} g(n', \nu) [q^d K^I(q(m'-n')) e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}][q^{-d} e^{-2\pi i R_0(\mu-\nu,\nu)\cdot a/q}] \\ &\quad = \sum_{n' \in \mathbf{Z}^d} \mathcal{S}^{a/q}(g)(n', \mu) q^d K^I(q(m'-n')) e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}. \end{aligned}$$

In view of Lemma 3.1, for the bound (3.27) it suffices to prove that

$$\Big\| \sum_{n'\in\mathbf{Z}^d} g'(n')q^d K^I(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}\Big\|_{L^2(\mathbf{Z}^d)} \leqslant C\|g'\|_{L^2(\mathbf{Z}^d)} \tag{3.29}$$

for any (compactly supported) function $g'\colon \mathbf{Z}^d\to\mathbf{C}$.

Since R_0 is bilinear, if $|m'|,|n'|\leqslant C|\xi|^{-1/2}/q$ then

$$\begin{aligned}|q^dK_j(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}-q^dK_j(q(m'-n'))|&\\ \leqslant C(2^j|\xi|^{1/2})(2^j/q)^{-d}\mathbf{1}_{[2^{j-1}/q,2^{j+1}/q]}(|m'-n'|).&\end{aligned}$$

Thus

$$|q^dK^I(q(m'-n'))e^{-2\pi i R_0(m'-n',n')\cdot q^2\xi}-q^dK^I(q(m'-n'))|\leqslant E''(m'-n'),$$

where $\|E''\|_{L^1(\mathbf{Z}^d)}\leqslant C$. The estimate (3.29) follows from the boundedness of standard singular integrals on $\mathbf{Z}^d$. □

4. The maximal Radon transform

In this section we prove Lemma 2.7. The proof is based on three main ingredients: a strong L^2 bound, a restricted (weak) L^p bound, $p\in(1,2]$, and an interpolation argument. We assume throughout this section that $d'=d^2$ and that $\mathbf{G}_0^{\#}$ is the discrete nilpotent group defined in §2.

4.1. L^2 estimates

The main result in this subsection is Lemma 4.1, which is a quantitative L^2 estimate. The proof of Lemma 4.1 is based on a non-commutative variant of the circle method, in which we divide the Fourier space into major arcs and minor arcs. This partition is achieved using cutoff functions like $\Psi_j^{N,\mathcal{R}}$ defined in equation (4.6). The minor arcs estimate (4.12) is based on Plancherel's theorem and Lemmas 3.2 and 3.3. The major arcs estimate (4.13) is based on the change of variables (4.28), the L^2 boundedness of the standard maximal function on the group $\mathbf{G}_0^{\#}$, and Lemma 3.1.

In this section we assume that $\Omega\colon\mathbf{R}^d\to[0,1]$ is a function supported in $\{x:|x|\leqslant 4\}$, and

$$\begin{aligned}|\Omega(x)|+|\nabla\Omega(x)|&\leqslant 10 &&\text{for any } x\in\mathbf{R}^d,\\ \Omega_j(x)&=2^{-dj}\Omega(x/2^j), && j=0,1,\dots.\end{aligned}\tag{4.1}$$

Clearly, if $\Omega(x)=1$ in the set $\{x:|x|\leqslant 1\}$, then

$$\mathcal{M}_0^{\#}(f)(m,u)\leqslant C\sup_{j\geqslant 0}\sum_{n\in\mathbf{Z}^d}\Omega_j(n)f((n,0)^{-1}\cdot(m,u)),$$

for any (compactly supported) function $f:\mathbf{G}_0^{\#}\to[0,\infty)$. For integers $j\geqslant 0$ and (compactly supported) functions $f:\mathbf{G}_0^{\#}\to\mathbf{C}$ let

$$\mathcal{M}_j(f)(m,u)=\sum_{n\in\mathbf{Z}^d}\Omega_j(n)f((n,0)^{-1}\cdot(m,u)). \tag{4.2}$$

To prove Lemma 2.7, it suffices to prove that for any (compactly supported) function $f:\mathbf{G}_0^{\#}\to\mathbf{C}$,

$$\Big\|\sup_{j\geqslant 0}|\mathcal{M}_j(f)|\Big\|_{L^p(\mathbf{G}_0^{\#})}\leqslant C_p\|f\|_{L^p(\mathbf{G}_0^{\#})},\quad p\in(1,2]. \tag{4.3}$$

For any (compactly supported) function $f:\mathbf{G}_0^{\#}\to\mathbf{C}$ let $\hat{f}$ denote its Fourier transform in the central variable, i.e.,

$$\hat{f}(m,\theta)=\sum_{u\in\mathbf{Z}^{d'}}f(m,u)e^{-2\pi iu\cdot\theta},\quad m\in\mathbf{Z}^d,\ \theta\in\mathbf{R}^{d'}. \tag{4.4}$$

Then

$$\widehat{\mathcal{M}}_j(f)(m,\theta)=\sum_{n\in\mathbf{Z}^d}\Omega_j(m-n)\hat{f}(n,\theta)e^{-2\pi iR_0(m-n,n)\cdot\theta}. \tag{4.5}$$

We use formula (4.5) and multipliers in the Fourier variable θ to decompose the operators $\mathcal{M}_j$.

Let $\psi:\mathbf{R}^{d'}\to[0,1]$ denote a smooth function supported in the set $\{\xi:|\xi|\leqslant 2\}$ and equal to 1 in the set $\{\xi:|\xi|\leqslant 1\}$. Assume that $N\in\big[\frac14,\infty\big)$, $j\in[0,\infty)\cap\mathbf{Z}$ and that $\mathcal{R}\subseteq\mathbf{Q}^{d'}$ is a discrete periodic set (i.e. if $r\in\mathcal{R}$ then $r+a\in\mathcal{R}$ for any $a\in\mathbf{Z}^{d'}$, and $\mathcal{R}\cap[0,1)^{d'}$ is finite). We define

$$\Psi_j^{N,\mathcal{R}}(\theta)=\sum_{r\in\mathcal{R}}\psi(2^{2j}N^{-1}(\theta-r)). \tag{4.6}$$

The function $\Psi_j^{N,\mathcal{R}}$ is periodic in θ (i.e. $\Psi_j^{N,\mathcal{R}}(\theta+a)=\Psi_j^{N,\mathcal{R}}(\theta)$ if $a\in\mathbf{Z}^d$), and supported in the union of the $2N2^{-2j}$-neighborhoods of the points in $\mathcal{R}$. We will always assume that j is sufficiently large (depending on N and $\mathcal{R}$) such that these neighborhoods are disjoint, so $\Psi_j^{N,\mathcal{R}}:\mathbf{R}^{d'}\to[0,1]$. By convention, $\Psi_j^{N,\varnothing}=0$. For (compactly supported) functions $f:\mathbf{G}_0^{\#}\to\mathbf{C}$ we define $\mathcal{M}_j^{N,\mathcal{R}}(f)$ by

$$\widehat{\mathcal{M}_j^{N,\mathcal{R}}}(f)(m,\theta)=\widehat{\mathcal{M}}_j(f)(m,\theta)\Psi_j^{N,\mathcal{R}}(\theta). \tag{4.7}$$

Our main lemma in this subsection is the following L^2 estimate.

LEMMA 4.1. (Strong L^2 bound) *Assume that $N\in[\frac{1}{2},\infty)$, that $\mathcal{R}_N\subseteq\mathbf{Q}^{d'}$ is a discrete periodic set and that $J_{N,\mathcal{R}_N}\in[0,\infty)$ is a real number with the properties*

$$\begin{aligned}&\{a/q: q\in[1,N] \text{ and } (a,q)=1\}\subseteq\mathcal{R}_N,\\ &2^{J_{N,\mathcal{R}_N}}\geqslant\Big[100\max_{a/q\in\mathcal{R}_N \text{ and } (a,q)=1} q\Big]^4.\end{aligned}\tag{4.8}$$

Then

$$\Big\|\sup_{j\geqslant J_{N,\mathcal{R}_N}}|\mathcal{M}_j(f)-\mathcal{M}_j^{N,\mathcal{R}_N}(f)|\Big\|_{L^2(\mathbf{G}_0^\#)}\leqslant C(N+1)^{-\bar{c}}\|f\|_{L^2(\mathbf{G}_0^\#)},\tag{4.9}$$

where $\bar{c}=\bar{c}(d)>0$.

Remark. In §5, Lemma 5.5, we need to allow for slightly more general kernels Ω, that is $\Omega:\mathbf{R}^d\to[0,1]$, supported in the set $\{x:|x|\leqslant 4\}$, equal to 1 in the set $\{x:|x|\leqslant 2\}$, and satisfying

$$|\nabla\Omega(x)|\leqslant A \quad \text{for any } x\in\mathbf{R}^d,$$

where $A\gg 1$. In this case the bound (4.9) becomes

$$\Big\|\sup_{j\geqslant J_{N,\mathcal{R}_N}}|\mathcal{M}_j(f)-\mathcal{M}_j^{N,\mathcal{R}_N}(f)|\Big\|_{L^2(\mathbf{G}_0^\#)}\leqslant AC(N+1)^{-\bar{c}}\|f\|_{L^2(\mathbf{G}_0^\#)}.$$

Proof. Estimate (4.3) for $p=2$ corresponds to the case $N=\frac{1}{2}$, $\mathcal{R}_N=\varnothing$ and $J_{N,\mathcal{R}_N}=0$ in Lemma 4.1. The condition (4.8) guarantees that $\Psi_j^{N,\mathcal{R}_N}:\mathbf{R}^{d'}\to[0,1]$ if $j\geqslant J_{N,\mathcal{R}_N}$. We decompose the operator $\mathcal{M}_j-\mathcal{M}_j^{N,\mathcal{R}_N}$ into the main contribution coming from the "major arcs" (in θ) and an error-type contribution coming from the complement of these major arcs. For integers $j,s\geqslant 0$ let

$$\gamma(j,s)=\begin{cases}1, & \text{if } 2^s\leqslant j^{3/2},\\ 0, & \text{if } 2^s>j^{3/2}.\end{cases}$$

For (compactly supported) functions $f:\mathbf{G}_0^\#\to\mathbf{C}$ we define $\mathcal{N}_{j,s}^{N,\mathcal{R}_N}(f)$ by

$$\widehat{\mathcal{N}_{j,s}^{N,\mathcal{R}_N}}(f)(m,\theta)=\gamma(j,s)[\widehat{\mathcal{M}}_j(f)(m,\theta)-\widehat{\mathcal{M}_j^{N,\mathcal{R}_N}}(f)(m,\theta)]\sum_{2^s\leqslant q<2^{s+1}}\psi(2^{2j+2}(\theta-a/q)),\tag{4.10}$$

where the sum is taken over irreducible d'-fractions a/q with $2^s\leqslant q<2^{s+1}$. Then we write

$$\mathcal{M}_j(f)-\mathcal{M}_j^{N,\mathcal{R}_N}(f)=\sum_{s\geqslant 0}\mathcal{N}_{j,s}^{N,\mathcal{R}_N}(f)+\mathcal{E}_j^{N,\mathcal{R}_N}(f).\tag{4.11}$$

This is our basic decomposition. It follows from properties (4.8) that $\mathcal{N}_{j,s}^{N,\mathcal{R}_N}(f)\equiv 0$ if $2^{s+1}\leqslant N$. Thus, for Lemma 4.1, it suffices to prove that

$$\bigg\|\bigg(\sum_{j\geqslant J_{N,\mathcal{R}_N}}|\mathcal{E}_j^{N,\mathcal{R}_N}(f)|^2\bigg)^{1/2}\bigg\|_{L^2(\mathbf{G}_0^\#)}\leqslant C(N+1)^{-\bar{c}}\|f\|_{L^2(\mathbf{G}_0^\#)} \tag{4.12}$$

and

$$\Big\|\sup_{j\geqslant J_{N,\mathcal{R}_N}}|\mathcal{N}_{j,s}^{N,\mathcal{R}_N}(f)|\Big\|_{L^2(\mathbf{G}_0^\#)}\leqslant C2^{-\bar{c}s}\|f\|_{L^2(\mathbf{G}_0^\#)} \tag{4.13}$$

if $2^{s+1}\geqslant N$.

Proof of estimate (4.12). (Minor arcs estimate) Let $s(j)$ denote the largest integer $\geqslant 0$ with the property that $2^{s(j)}\leqslant j^{3/2}$. Notice that

$$\widehat{\mathcal{E}_j^{N,\mathcal{R}_N}}(f)(m,\theta)=m_j^{N,\mathcal{R}_N}(\theta)\sum_{n\in\mathbf{Z}^d}\Omega_j(m-n)\hat{f}(n,\theta)e^{-2\pi iR_0(m-n,n)\cdot\theta},$$

with

$$m_j^{N,\mathcal{R}_N}(\theta)=[1-\Psi_j^{N,\mathcal{R}_N}(\theta)]\bigg[1-\sum_{q\leqslant 2^{s(j)+1}-1}\psi(2^{2j+2}(\theta-a/q))\bigg], \tag{4.14}$$

where the sum in (4.14) is taken over irreducible d'-fractions a/q with $q\leqslant 2^{s(j)+1}-1$. For $\theta\in\mathbf{R}^{d'}$ and (compactly supported) functions $g\colon\mathbf{Z}^d\to\mathbf{C}$, we define

$$\mathcal{U}_j^\theta(g)(m)=\sum_{n\in\mathbf{Z}^d}\Omega_j(m-n)g(n)e^{-2\pi iR_0(m-n,n)\cdot\theta}. \tag{4.15}$$

By Plancherel's theorem,

$$\bigg\|\bigg(\sum_{j\geqslant J_{N,\mathcal{R}_N}}|\mathcal{E}_j^{N,\mathcal{R}_N}(f)|^2\bigg)^{1/2}\bigg\|_{L^2(\mathbf{G}_0^\#)}^2=\int_{[0,1)^{d'}}\sum_{j\geqslant J_{N,\mathcal{R}_N}}|m_j^{N,\mathcal{R}_N}(\theta)|^2\|\mathcal{U}_j^\theta(\hat{f}(\cdot\,,\theta))\|_{L^2(\mathbf{Z}^d)}^2\,d\theta.$$

Using Plancherel's theorem again, for the bound (4.12) it suffices to prove that

$$\sum_{j\geqslant J_{N,\mathcal{R}_N}}|m_j^{N,\mathcal{R}_N}(\theta)|^2\|\mathcal{U}_j^\theta\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}^2\leqslant C(N+1)^{-2\bar{c}} \tag{4.16}$$

for any $\theta\in\mathbf{R}^{d'}$ fixed.

By Diriclet's principle, for any $\Lambda\geqslant 1$ and $\xi\in\mathbf{R}$ there are $q\in\mathbf{Z}_\Lambda=\mathbf{Z}\cap[1,\Lambda]$ and $a\in\mathbf{Z}$, $(a,q)=1$, with the property that $|\xi-a/q|\leqslant 1/\Lambda q$. For $\theta\in\mathbf{R}^{d'}$ we apply this to each component $\theta_{l_1l_2}$; thus there are $q_{l_1l_2}\in\mathbf{Z}_\Lambda$ and $a_{l_1l_2}\in\mathbf{Z}$, $(a_{l_1l_2},q_{l_1l_2})=1$, with the property that

$$|\theta_{l_1l_2}-a_{l_1l_2}/q_{l_1l_2}|\leqslant\frac{C}{\Lambda q_{l_1l_2}}. \tag{4.17}$$

Assume that $\theta\in\mathbf{R}^{d'}$ is fixed. For any $j\geqslant J_{N,\mathcal{R}_N}$ we use the approximation (4.17) with $\Lambda=2^{(2-\delta)j}$, where $\delta=\delta(d)>0$ is sufficiently small ($\delta=1/10d'$ would work). Thus there are irreducible 1-fractions $a^j_{l_1l_2}/q^j_{l_1l_2}$ such that

$$1\leqslant q^j_{l_1l_2}\leqslant 2^{(2-\delta)j} \quad \text{and} \quad |\theta_{l_1l_2}-a^j_{l_1l_2}/q^j_{l_1l_2}|\leqslant \frac{C}{2^{(2-\delta)j}q^j_{l_1l_2}}. \tag{4.18}$$

We fix these irreducible 1-fractions $a^j_{l_1l_2}/q^j_{l_1l_2}$ and partition the set $\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty)$ into two subsets:

$$I_1=\{j\in\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty): \max_{l_1,l_2=1,\dots,d} q^j_{l_1l_2}>2^{j/6d'}\}$$

and

$$I_2=\{j\in\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty): \max_{l_1,l_2=1,\dots,d} q^j_{l_1l_2}\leqslant 2^{j/6d'}\}.$$

For $j\in I_1$ we use Lemma 3.2:

$$\sum_{j\in I_1}|m_j^{N,\mathcal{R}_N}(\theta)|^2\|\mathcal{U}_j^\theta\|^2_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}\leqslant\sum_{j\in I_1}2^{-\delta' j}\leqslant C(N+1)^{-\bar{c}},$$

as desired.

For $j\in I_2$ let a_j/q_j denote the irreducible d'-fraction with the property that

$$a_j/q_j=(a^j_{l_1l_2}/q^j_{l_1l_2})_{l_1,l_2=1,\dots,d}.$$

In view of properties (4.18) and the definition of I_2,

$$1\leqslant q_j\leqslant 2^{j/6} \quad \text{and} \quad |\theta-a_j/q_j|\leqslant\frac{C}{2^{(2-\delta)j}}. \tag{4.19}$$

An easy argument, using properties (4.19), shows that if $j,j'\in I_2$ and $j,j'\geqslant C$ then

$$\text{either} \quad a_j/q_j=a_{j'}/q_{j'} \quad \text{or} \quad |q_j/q_{j'}|\notin\left[\tfrac{1}{2},2\right]. \tag{4.20}$$

We further partition the set I_2:

$$I_2=\bigcup_{a/q}I_2^{a/q}, \quad \text{where } I_2^{a/q}=\{j\in I_2: a_j/q_j=a/q\}. \tag{4.21}$$

For $j\in I_2^{a/q}$ we use Lemma 3.3:

$$\sum_{j\in I_2^{a/q}}|m_j^{N,\mathcal{R}_N}(\theta)|^2\|\mathcal{U}_j^\theta\|^2_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}\leqslant C\sum_{j\in I_2^{a/q}}q^{-1}(1+2^{2j}|\theta-a/q|)^{-1/2}|m_j^{N,\mathcal{R}_N}(\theta)|^2. \tag{4.22}$$

To estimate the right-hand side of (4.22), we consider two cases: $q\leqslant N$ and $q>N$. If $q\leqslant N$, then, using (4.8), (4.6) and (4.14), $|m_j^{N,\mathcal{R}_N}(\theta)|^2\leqslant \mathbf{1}_{[1,\infty)}(2^{2j}N^{-1}|\theta-a/q|)$. Thus the right-hand side of (4.22) is dominated by $Cq^{-1}N^{-1/2}$. If $q>N$, then, using (4.14) and the fact that $j\geqslant 2^{2s(j)/3}$, the right-hand side of (4.22) is dominated by

$$C\sum_{j\in I_2^{a/q}\cap[0,Cq^{2/3}]} q^{-1}+C\sum_{j\in I_2^{a/q}\cap[Cq^{2/3},\infty)} q^{-1}(1+2^{2j}|\theta-a/q|)^{-1/2}\mathbf{1}_{[1/2,\infty)}(2^{2j}|\theta-a/q|) \leqslant Cq^{-1/3}.$$

The bound (4.16) follows since the possible denominators q form a lacunary sequence (see property (4.20)). This completes the proof of estimate (4.12).

Proof of estimate (4.13). (Major arcs estimate) Clearly, if $j\geqslant\max(J_{N,\mathcal{R}_N},2^{2s/3},C)$, then

$$\biggl(\sum_{2^s\leqslant q<2^{s+1}}\psi(2^{2j+2}(\theta-a/q))\biggr)(1-\Psi_j^{N,\mathcal{R}_N}(\theta))=\sum_{r\in\mathcal{R}'}\psi(2^{2j+2}(\theta-r)),$$

where $\mathcal{R}'=\{a/q\in\mathbf{Q}^{d'}\setminus\mathcal{R}_N:(a,q)=1 \text{ and } q\in[2^s,2^{s+1})\}$. We define $\mathcal{M}_j^{1/4,\mathcal{R}'}(f)$ by

$$\widehat{\mathcal{M}_j^{1/4,\mathcal{R}'}}(f)(m,\theta)=\widehat{\mathcal{M}}_j(f)(m,\theta)\sum_{r\in\mathcal{R}'}\psi(2^{2j+2}(\theta-r)) \tag{4.23}$$

(compare with equation (4.7)). Thus, for estimate (4.13), it suffices to prove that if $s\geqslant 0$ and $\mathcal{R}'\subseteq\{a/q:(a,q)=1 \text{ and } q\in[2^s,2^{s+1})\}$, then

$$\Bigl\|\sup_{j\geqslant 2^{2s/3}}|\mathcal{M}_j^{1/4,\mathcal{R}'}(f)|\Bigr\|_{L^2(\mathbf{G}_0^\#)}\leqslant C2^{-\bar{c}s}\|f\|_{L^2(\mathbf{G}_0^\#)}.$$

We partition the set $\mathcal{R}'\subseteq\{a/q:(a,q)=1 \text{ and } q\in[2^s,2^{s+1})\}$ into at most $C2^{2s/5}$ subsets with the property that each of these subsets contains irreducible d'-fractions with at most $2^{3s/5}$ denominators q. Thus, it suffices to prove that if $s\geqslant 0$,

$$\mathcal{R}'\subseteq\{a/q:(a,q)=1 \text{ and } q\in S\},\quad S\subseteq[2^s,2^{s+1})\cap\mathbf{Z}\quad\text{and}\quad |S|\leqslant 2^{3s/5}, \tag{4.24}$$

then

$$\Bigl\|\sup_{j\geqslant 2^{2s/3}}|\mathcal{M}_j^{1/4,\mathcal{R}'}(f)|\Bigr\|_{L^2(\mathbf{G}_0^\#)}\leqslant C2^{-s/2}\|f\|_{L^2(\mathbf{G}_0^\#)}. \tag{4.25}$$

In view of the definitions (4.5) and (4.23), and the Fourier inversion formula,

$$\begin{aligned}
&\mathcal{M}_j^{1/4,\mathcal{R}'}(f)(m,u)\\
&\quad=\sum_{(n,v)\in\mathbf{G}_0^\#}f(n,v)\Omega_j(m-n)\int_{[0,1)^{d'}}\biggl(\sum_{r\in\mathcal{R}'}\psi(2^{2j+2}(\theta-r))\biggr)e^{2\pi i(u-v-R_0(m-n,n))\cdot\theta}\,d\theta\\
&\quad=\sum_{(n,v)\in\mathbf{G}_0^\#}f(n,v)\Omega_j(m-n)\eta_{2^{2j+2}}(u-v-R_0(m-n,n))\sum_{r\in\mathcal{R}'\cap[0,1)^{d'}}e^{2\pi i(u-v-R_0(m-n,n))\cdot r},
\end{aligned} \tag{4.26}$$

where $\eta(s)=\int_{\mathbf{R}^{d'}}\psi(\xi)e^{2\pi is\cdot\xi}\,d\xi$ is the Euclidean inverse Fourier transform of ψ, and $\eta_{2^{2j+2}}(s)=2^{-d'(2j+2)}\eta(s/2^{2j+2})$. We recognize that formula (4.26) is the convolution on $\mathbf{G}_0^\#$ of the function f and the kernel

$$(m,u)\longrightarrow\Omega_j(m)\eta_{2^{2j+2}}(u)\sum_{r\in\mathcal{R}'\cap[0,1)^{d'}}e^{2\pi iu\cdot r}.$$

Let $Q=\prod_{q\in S}q$; see (4.24). Since $|S|\leqslant 2^{3s/5}$,

$$Q\leqslant 2^{(s+1)2^{3s/5}}. \tag{4.27}$$

To continue, we introduce new coordinates on $\mathbf{G}_0^\#$ adapted to the factor Q. For integers $Q\geqslant 1$ we define

$$\begin{gathered}\Phi_Q\colon\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]\longrightarrow\mathbf{G}_0^\#,\\ \Phi_Q((m',u'),(\mu,\alpha))=(Qm'+\mu,Q^2u'+\alpha+QR_0(\mu,m')).\end{gathered} \tag{4.28}$$

Notice that $\Phi_Q((m',u'),(\mu,\alpha))=(\mu,\alpha)\cdot(Qm',Q^2u')$ if we regard (μ,α) and (Qm',Q^2u') as elements of $\mathbf{G}_0^\#$. Clearly, the map Φ_Q establishes a bijection between $\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]$ and $\mathbf{G}_0^\#$. Let $F((n',v'),(\nu,\beta))=f(\Phi_Q((n',v'),(\nu,\beta)))$ and

$$G_j((m',u'),(\mu,\alpha))=\mathcal{M}_j^{1/4,\mathcal{R}'}(f)(\Phi_Q((m',u'),(\mu,\alpha))).$$

Since $Qr\in\mathbf{Z}$ for any $r\in\mathcal{R}'$, formula (4.26) is equivalent to

$$\begin{aligned}&G_j((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F((n',v'),(\nu,\beta))\Omega_j(Q(m'-n')+E_1)\\&\qquad\qquad\times\eta_{2^{2j+2}}(Q^2(u'-v'-R_0(m'-n',n'))+E_2)\sum_{r\in\mathcal{R}'\cap[0,1)^{d'}}e^{2\pi i(\alpha-\beta-R_0(\mu-\nu,\nu))\cdot r},\end{aligned}$$

where $E_1=\mu-\nu$ and

$$E_2=(\alpha-\beta-R_0(\mu-\nu,\nu))+Q(R_0(\mu,m'-n')-R_0(m'-n',\nu)).$$

In view of estimates (4.25) and (4.27), $2^j\geqslant 2^{2^{2s/3}}$ and $Q\leqslant 2^{(s+1)2^{3s/5}}$, thus $C2^j\geqslant Q^{10}$. Clearly, $|E_1|\leqslant CQ$ and $|E_2|\leqslant C2^jQ$ if $|m'-n'|\leqslant C2^j/Q$. Let

$$\begin{aligned}&\widetilde{G}_j((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F((n',v'),(\nu,\beta))\Omega_j(Q(m'-n'))\\&\qquad\qquad\times\eta_{2^{2j+2}}(Q^2(u'-v'-R_0(m'-n',n')))\sum_{r\in\mathcal{R}'\cap[0,1)^{d'}}e^{2\pi i(\alpha-\beta-R_0(\mu-\nu,\nu))\cdot r}.\end{aligned} \tag{4.29}$$

In view of the estimates above on $|E_1|$ and $|E_2|$, and the fact that the sum over $r\in \mathcal{R}'\cap[0,1)^{d'}$ in equation (4.29) has at most $C2^{2s}$ terms, we have

$$\begin{aligned}&|G_j((m',u'),(\mu,\alpha))-\widetilde{G}_j((m',u'),(\mu,\alpha))|\\ &\quad\leqslant C2^{Cs}(Q/2^j)\sum_{(n',v')\in\mathbf{G}_0^\#}\ \sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}|F((n',v'),(\nu,\beta))|Q^{-d}Q^{-2d'}\\ &\qquad\qquad\times(2^j/Q)^{-d}\mathbf{1}_{[0,C2^j/Q]}(|m'-n'|)\phi_{2^{2j}/Q^2}(u'-v'-R_0(m'-n',n')),\end{aligned}$$

where, as in definition (7.7),

$$\phi(s)=(1+|s|^2)^{-(d'+d+1)/2}\quad\text{and}\quad \phi_r(s)=r^{-d'}\phi(s/r),\quad r\geqslant 1.$$

Thus,

$$\sum_{j\geqslant 2^{2s/3}}\|G_j-\widetilde{G}_j\|_{L^2(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C2^{-2^{s/2}}\|F\|_{L^2(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}.$$

For estimate (4.25), it suffices to prove that

$$\Big\|\sup_{2^j\geqslant Q}|\widetilde{G}_j|\Big\|_{L^2(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C2^{-s/2}\|F\|_{L^2(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])},\tag{4.30}$$

where $\widetilde{G}_j$ is defined in equation (4.29). For this, we notice that the function $\widetilde{G}_j$ is obtained as the composition of the operator

$$\mathcal{A}(f)(\mu,\alpha)=Q^{-d}Q^{-2d'}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}f(\nu,\beta)\sum_{r\in\mathcal{R}'\cap[0,1)^{d'}}e^{2\pi i(\alpha-\beta-R_0(\mu-\nu,\nu))\cdot r}\tag{4.31}$$

acting on functions $f\colon\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}\to\mathbf{C}$, followed by an average over a standard ball of radius $\approx 2^j/Q$ in $\mathbf{G}_0^\#$ (with the terminology of §7). In view of estimates (7.11) with $N=1$, for estimate (4.30) it suffices to prove that

$$\|\mathcal{A}(f)\|_{L^2(\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'})}\leqslant C2^{-s/2}\|f\|_{L^2(\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'})}.\tag{4.32}$$

For functions $f\colon\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}\to\mathbf{C}$, we define the Fourier transform in the second variable

$$\tilde{f}(\mu,a/Q^2)=\sum_{\alpha\in\mathbf{Z}_{Q^2}^{d'}}f(\mu,\alpha)e^{-2\pi i\alpha\cdot a/Q^2},\quad a\in\mathbf{Z}^{d'}.$$

It is easy to see that

$$\|f\|_{L^2(\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'})}=Q^{-d'}\bigg(\sum_{\mu\in\mathbf{Z}_Q^d}\ \sum_{a\in\mathbf{Z}_{Q^2}^{d'}}|\tilde{f}(\mu,a/Q^2)|^2\bigg)^{1/2},$$

for any $f\colon \mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}\to\mathbf{C}$ (Plancherel's identity). Since $\mathcal{R}'\subseteq\{a/Q^2:a\in\mathbf{Z}^{d'}\}$ (see (4.24) and the definition of Q), it follows from equation (4.31) that

$$\widetilde{\mathcal{A}(f)}(\mu,a/Q^2)=\mathbf{1}_{\mathcal{R}'}(a/Q^2)Q^{-d}\sum_{\nu\in\mathbf{Z}_Q^d}\tilde{f}(\nu,a/Q^2)e^{-2\pi iR_0(\mu-\nu,\nu)\cdot a/Q^2}.$$

By Plancherel's identity, for estimate (4.32) it suffices to prove that for any $r\in\mathcal{R}'$ and any $g\colon\mathbf{Z}_q^d\to\mathbf{C}$,

$$\left\|Q^{-d}\sum_{\nu\in\mathbf{Z}_Q^d}g(\nu)e^{-2\pi iR_0(\mu-\nu,\nu)\cdot r}\right\|_{L^2_\mu(\mathbf{Z}_Q^d)}\leqslant C2^{-s/2}\|g\|_{L^2_\nu(\mathbf{Z}_Q^d)}.$$

This follows from Lemma 3.1 and the fact that $r=a/q$, $(a,q)=1$, $q\in[2^s,2^{s+1})$ (see (4.24)). This completes the proof of Lemma 4.1. □

4.2. A restricted L^p estimate

Recall that the operators $\mathcal{M}_j$ were defined in equation (4.2). In the rest of this section, in addition to conditions (4.1), we assume that $\Omega(x)=1$ if $|x|\leqslant 2$. In this subsection we prove the following restricted L^p estimate.

LEMMA 4.2. (Restricted L^p estimate) *Assume that $J\geqslant 2$ is an integer. Then*

$$\Big\|\sup_{j\in[J+1,2J]}|\mathcal{M}_j(f)|\Big\|_{L^p(\mathbf{G}_0^\#)}\leqslant C_p(\log J)\|f\|_{L^p(\mathbf{G}_0^\#)},\quad p\in(1,2].\tag{4.33}$$

The idea of using restricted L^p estimates like (4.33) together with L^2 bounds to prove the full L^p estimates (4.3) originates in Bourgain's paper [5]. In proving Lemma 4.2, we exploit the positivity of the operators $\mathcal{M}_j$. Let $\widetilde{\Omega}_j\colon\mathbf{G}_0^\#\to[0,\infty)$ denote the kernel $\widetilde{\Omega}_j(m,u)=\Omega_j(m)\mathbf{1}_{\{0\}}(u)$, so $\mathcal{M}_j(f)=f*\widetilde{\Omega}_j$, and let $\Omega'_j(h)=\widetilde{\Omega}_j(h^{-1})$. To be able to use the same notation as in the previous section, it is more convenient to prove the maximal inequality

$$\Big\|\sup_{j\in[J+1,2J]}|f*\Omega'_j|\Big\|_{L^p(\mathbf{G}_0^\#)}\leqslant C_p(\log J)\|f\|_{L^p(\mathbf{G}_0^\#)},\quad p\in(1,2].\tag{4.34}$$

The bounds (4.33) and (4.34) are equivalent, in view of the duality argument following the statement of Lemma 2.8. By interpolation, we may assume that $p'=p/(p-1)$ is an integer $\geqslant 2$ and it suffices to prove the $L^p\to L^{p,\infty}$ estimate

$$\Big\|\sup_{j\in[J+1,2J]}|f*\Omega'_j|\Big\|_{L^{p,\infty}(\mathbf{G}_0^\#)}\leqslant C_p(\log J)\|f\|_{L^p(\mathbf{G}_0^\#)},\quad p'\in[2,\infty)\cap\mathbf{Z}.\tag{4.35}$$

By duality, the bound (4.35) is equivalent to the inequality

$$\Big\|\sum_{j=J+1}^{2J} f_j*\widetilde{\Omega}_j\Big\|_{L^k(\mathbf{G}_0^\#)} \leqslant C_k(\log J)\Big\|\sum_{j=J+1}^{2J} f_j\Big\|_{L^k(\mathbf{G}_0^\#)},$$

where $k=p/(p-1)$ is an integer $\geqslant 2$ and the f_j's are characteristic functions of disjoint, bounded sets. We may assume that $J\geqslant C_k$ and partition the set $[J+1,2J]\cap\mathbf{Z}$ into at most $C_k(\log J)$ subsets S with the separation property

$$S\subseteq[J+1,2J]\cap\mathbf{Z} \quad\text{and}\quad \text{if } j\neq j'\in S \text{ then } |j-j'|\geqslant A_k(\log J), \tag{4.36}$$

where A_k is a large constant to be fixed later. It suffices to prove that if S is as above and $k\geqslant 2$ is an integer, then

$$\Big\|\sum_{j\in S} f_j*\widetilde{\Omega}_j\Big\|_{L^k(\mathbf{G}_0^\#)} \leqslant C_k\Big\|\sum_{j\in S} f_j\Big\|_{L^k(\mathbf{G}_0^\#)}, \tag{4.37}$$

where the f_j's are characteristic functions of disjoint, bounded sets. Let ϱ denote the smallest constant $C_k\geqslant 1$ for which the bound (4.37) holds. By expanding the left-hand side of inequality (4.37),

$$\begin{aligned}\Big\|\sum_{j\in S} f_j*\widetilde{\Omega}_j\Big\|_{L^k(\mathbf{G}_0^\#)}^k \leqslant C_k &\sum_{j_1<\ldots<j_k}\int_{\mathbf{G}_0^\#}(f_{j_1}*\widetilde{\Omega}_{j_1})\ldots(f_{j_k}*\widetilde{\Omega}_{j_k})\,dg\\ &+C_k\int_{\mathbf{G}_0^\#}\Big(\sum_{j\in S} f_j*\widetilde{\Omega}_j\Big)^{k-1}dg,\end{aligned} \tag{4.38}$$

since the f_j's are characteristic functions. The second term in the right-hand side of inequality (4.38) is dominated by $C_k\varrho^{k-1}\|\sum_{j\in S} f_j\|_{L^k(\mathbf{G}_0^\#)}^k$.

To deal with the first term, we will prove the bound

$$\|[(f_{j_2}*\widetilde{\Omega}_{j_2})\ldots(f_{j_k}*\widetilde{\Omega}_{j_k})]*(\Omega'_{j_1}-\Omega'_J)\|_{L^2(\mathbf{G}_0^\#)} \leqslant C_kJ^{-k}\|f_{j_2}+\ldots+f_{j_k}\|_{L^2(\mathbf{G}_0^\#)}, \tag{4.39}$$

provided $f_{j_2},\ldots,f_{j_k}$ are characteristic functions of disjoint, bounded sets, $j_1<\ldots<j_k\in S$, and the constant A_k in (4.36) is sufficiently large. Assuming the bound (4.39), we would have

$$\begin{aligned}\Big|\int_{\mathbf{G}_0^\#}(f_{j_1}*\widetilde{\Omega}_{j_1})\ldots(f_{j_k}*\widetilde{\Omega}_{j_k})\,dg-\int_{\mathbf{G}_0^\#}(f_{j_1}*\widetilde{\Omega}_J)(f_{j_2}*\widetilde{\Omega}_{j_2})\ldots(f_{j_k}*\widetilde{\Omega}_{j_k})\,dg\Big|\\ \leqslant C_kJ^{-k}\Big\|\sum_{j\in S} f_j\Big\|_{L^2(\mathbf{G}_0^\#)}^2 = C_kJ^{-k}\Big\|\sum_{j\in S} f_j\Big\|_{L^k(\mathbf{G}_0^\#)}^k,\end{aligned}$$

since the f_j's are characteristic functions of disjoint, bounded sets. Thus the first term in the right-hand side of inequality (4.38) can be estimated by

$$C_k\Big\|\sum_{j\in S} f_j\Big\|^k_{L^k(\mathbf{G}_0^\#)} + C_k \sum_{j_2<\ldots<j_k} \int_{\mathbf{G}_0^\#} \Big(\sum_{j\in S} f_j * \widetilde{\Omega}_J\Big)(f_{j_2} * \widetilde{\Omega}_{j_2}) \ldots (f_{j_k} * \widetilde{\Omega}_{j_k})\, dg. \tag{4.40}$$

Since the f_j's are characteristic functions of disjoint, bounded sets, $\sum_{j\in S} f_j * \widetilde{\Omega}_J \leqslant C$. Thus the expression in (4.40) can be estimated by $C_k(1+\varrho^{k-1})\|\sum_{j\in S} f_j\|^k_{L^k(\mathbf{G}_0^\#)}$. It follows from estimate (4.38) that $\varrho^k \leqslant C_k(1+\varrho^{k-1})$, so $\varrho \leqslant C_k$ as desired.

It remains to prove the bound (4.39). Clearly, we may assume that $J \geqslant C_k$. We start with a sequence of appropriate constants $B_2 < \ldots < B_k$, which depend only on the constant $\bar{c}>0$ in Lemma 4.1, and define $N_l = J^{B_l}$ and $\mathcal{R}_{N_l} = \{a/q : q\in[1,N_l] \text{ and } (a,q)=1\}$, $l=2,\ldots,k$. By Lemma 4.1,

$$\|\mathcal{M}_{j_l}(f_{j_l}) - \mathcal{M}_{j_l}^{N_l,\mathcal{R}_{N_l}}(f_{j_l})\|_{L^2(\mathbf{G}_0^\#)} \leqslant CJ^{-\bar{c}B_l}\|f_{j_l}\|_{L^2(\mathbf{G}_0^\#)}, \quad l=2,\ldots,k. \tag{4.41}$$

A computation similar to (4.26) shows that

$$\begin{aligned} &\mathcal{M}_{j_l}^{N_l,\mathcal{R}_{N_l}}(f_{j_l}) = f_{j_l} * L_{j_l}^{N_l,\mathcal{R}_{N_l}},\\ &L_{j_l}^{N_l,\mathcal{R}_{N_l}}(m,u) = \Omega_{j_l}(m)\eta_{2^{2j_l}/N_l}(u) \sum_{r\in\mathcal{R}_{N_l}\cap[0,1)^{d'}} e^{2\pi i u\cdot r}. \end{aligned} \tag{4.42}$$

Since $\mathcal{R}_{N_l}$ has at most $CJ^{(d'+1)B_l}$ elements, $\|L_{j_l}^{N_l}\|_{L^1(\mathbf{G}_0^\#)} \leqslant CJ^{(d'+1)B_l}$. Thus

$$\|\mathcal{M}_{j_l}^{N_l,\mathcal{R}_{N_l}}(f_{j_l})\|_{L^\infty(\mathbf{G}_0)} \leqslant CJ^{(d'+1)B_l}, \quad l=2,\ldots,k, \tag{4.43}$$

since the f_{j_l}'s are characteristic functions of sets. Now, by replacing each $\mathcal{M}_{j_l}(f_{j_l})$ by $\mathcal{M}_{j_l}^{N_l,\mathcal{R}_{N_l}}(f_{j_l})$, for $l=2,\ldots,k$, one at a time, the left-hand side of (4.39) is dominated by

$$\begin{aligned} &\|\mathcal{M}_{j_2}(f_{j_2}) - \mathcal{M}_{j_2}^{N_2,\mathcal{R}_{N_2}}(f_{j_2})\|_{L^2} \ldots \|\mathcal{M}_{j_k}(f_{j_k})\|_{L^\infty}\\ &\qquad + \ldots + \|\mathcal{M}_{j_2}^{N_2,\mathcal{R}_{N_2}}(f_{j_2})\|_{L^\infty} \ldots \|\mathcal{M}_{j_k}(f_{j_k}) - \mathcal{M}_{j_k}^{N_k,\mathcal{R}_{N_k}}(f_{j_k})\|_{L^2}\\ &\qquad + \|[\mathcal{M}_{j_2}^{N_2,\mathcal{R}_{N_2}}(f_{j_2}) \ldots \mathcal{M}_{j_k}^{N_k,\mathcal{R}_{N_k}}(f_{j_k})] * (\Omega'_{j_1} - \Omega'_J)\|_{L^2}. \end{aligned} \tag{4.44}$$

By choosing the constants B_l in geometric progression and using estimates (4.41) and (4.43), for (4.39) it remains to control the last term in expression (4.44). We now examine formula (4.42) and notice that each kernel $L_{j_l}^{N_l,\mathcal{R}_{N_l}}$ is the sum over r of at most $C_kJ^{C_k}$ kernels. For any irreducible d'-fraction a_l/q_l, let

$$L_{j_l}^{N_l,a_l/q_l}(m,u) = \Omega_{j_l}(m)\eta_{2^{2j_l}/N_l}(u)e^{2\pi i u\cdot a_l/q_l}. \tag{4.45}$$

To control the last term in (4.44) it suffices to prove the following lemma.

LEMMA 4.3. *With the notation above, for any constant* $\widetilde{B}_k$,

$$\| [(f_{j_2} * L_{j_2}^{N_2, a_2/q_2}) \dots (f_{j_k} * L_{j_k}^{N_k, a_k/q_k})] * (\Omega'_{j_1} - \Omega'_J) \|_{L^2} \leqslant C_k J^{-\widetilde{B}_k} \| f_{j_2} + \dots + f_{j_k} \|_{L^2}, \quad (4.46)$$

provided $f_{j_2}, \dots, f_{j_k}$ *are characteristic functions of disjoint, bounded sets,* $N_l \leqslant J^{\widetilde{B}_k}$, a_l/q_l *are irreducible* d'*-fractions with* $q_l \leqslant J^{\widetilde{B}_k}$, $l=2, \dots, k$, $J < j_1 < j_2 < \dots < j_k \leqslant 2J$ *and* $j_2 - j_1 \geqslant A_k \log J$, A_k *sufficiently large depending on* $\widetilde{B}_k$.

Proof. From the definitions,

$$\begin{aligned}[](f_{j_2} * L_{j_2}^{N_2, a_2/q_2}) \dots (f_{j_k} * L_{j_k}^{N_k, a_k/q_k})] * (\Omega'_{j_1} - \Omega'_J)(g) \\ = \int_{[\mathbf{G}_0^\#]^{k-1}} f_{j_2}(h_2) \dots f_{j_k}(h_k) H(g \cdot h_2^{-1}, \dots, g \cdot h_k^{-1}) \, dh_2 \dots dh_k,\end{aligned} \quad (4.47)$$

where

$$H(g_2, \dots, g_k) = \sum_{n \in \mathbf{Z}^d} (\Omega_{j_1}(n) - \Omega_J(n)) L_{j_2}^{N_2, a_2/q_2}((n,0) \cdot g_2) \dots L_{j_k}^{N_k, a_k/q_k}((n,0) \cdot g_k). \quad (4.48)$$

Let $g_l = (m_l, u_l)$, $l = 2, \dots, k$. With ϕ as in defition (7.7), we show that

$$|H(g_2, \dots, g_k)| \leqslant C_k J^{\widetilde{B}_k} (2^{j_1 - j_2} + 2^{-J/2}) \prod_{l=2}^k \Omega_{j_l+2}(m_l) \phi_{2^{2j_l}/N_l}(u_l). \quad (4.49)$$

Assuming (4.49), the bound (4.46) follows easily from equation (4.47) and the fact that the f_{j_l}'s are characteristic functions.

To prove the bound (4.49) let

$$Q = q_2 \dots q_k, \quad Q \leqslant J^{(k-1)\widetilde{B}_k}. \quad (4.50)$$

Writing $n = Qn' + \nu$, $n' \in \mathbf{Z}^d$, $\nu \in \mathbf{Z}_Q^d$, equation (4.48) becomes

$$\begin{aligned}|H(g_2, \dots, g_k)| = \bigg| \sum_{n' \in \mathbf{Z}^d} \sum_{\nu \in \mathbf{Z}_Q^d} (\Omega_{j_1}(Qn' + \nu) - \Omega_J(Qn' + \nu)) \\ \times \prod_{l=2}^k \Omega_{j_l}(m_l + Qn' + \nu) \eta_{2^{2j_l}/N_l}(u_l + R_0(Qn' + \nu, m_l)) e^{2\pi i R_0(\nu, m_l) \cdot a_l/q_l} \bigg|.\end{aligned} \quad (4.51)$$

We use (4.50) on Q and the observation that $|n'| \leqslant 100 \cdot 2^{j_1}/Q$ in (4.51). It follows that

$$\begin{aligned}|\Omega_{j_l}(m_l + Qn' + \nu) \eta_{2^{2j_l}/N_l}(u_l + R_0(Qn' + \nu, m_l)) - \Omega_{j_l}(m_l) \eta_{2^{2j_l}/N_l}(u_l)| \\ \leqslant C N_l 2^{j_1 - j_l} \Omega_{j_l+2}(m_l) \phi_{2^{2j_l}/N_l}(u_l).\end{aligned}$$

Thus, using equation (4.51),

$$|H(g_2,...,g_k)| \leqslant C_k \prod_{l=2}^{k} \Omega_{j_l+2}(m_l)\phi_{2^{2j_l}/N_l}(u_l) \\ \times \biggl(J^{\widetilde{B}_k} 2^{j_1-j_2} + \biggl| \sum_{n'\in\mathbf{Z}^d} \sum_{\nu\in\mathbf{Z}_Q^d} (\Omega_{j_1}(Qn'+\nu)-\Omega_J(Qn'+\nu)) \prod_{l=2}^{k} e^{2\pi i R_0(\nu,m_l)\cdot a_l/q_l} \biggr| \biggr). \tag{4.52}$$

We make the simple observations

$$|\Omega_{j_1}(Qn'+\nu)-\Omega_{j_1}(Qn')| \leqslant CQ2^{-j_1}\Omega_{j_1+2}(Qn'),$$
$$|\Omega_J(Qn'+\nu)-\Omega_J(Qn')| \leqslant CQ2^{-J}\Omega_J(Qn'),$$

since $|\nu|\leqslant Q$. In addition, since $\int_{\mathbf{R}^d}(\Omega_{j_1}(x')-\Omega_J(x'))\,dx'=0$, we have

$$Q^d \biggl| \sum_{n'\in\mathbf{Z}^d} (\Omega_{j_1}(Qn')-\Omega_J(Qn')) \biggr| \leqslant CQ2^{-J}.$$

The bound (4.49) follows from inequality (4.52). □

4.3. Proof of Lemma 2.7

In this subsection we prove the bound (4.3) for any $p>1$, thus completing the proof of Lemma 2.7. Our main ingredients are the bound (7.11) in §7, and Lemmas 4.1 and 4.2. The bound (4.3) follows by interpolation (see [8, §7]) from the following more quantitative estimate.

LEMMA 4.4. *Assume that $p\in(1,2]$ is an exponent and $\varepsilon=(p-1)/2$. Then, for any $\lambda\in(0,\infty)$, there are linear operators $\mathcal{A}_j^\lambda=\mathcal{A}_j^{\lambda,\varepsilon}$ and $\mathcal{B}_j^\lambda=\mathcal{B}_j^{\lambda,\varepsilon}$ with $\mathcal{M}_j=\mathcal{A}_j^\lambda+\mathcal{B}_j^\lambda$,*

$$\Bigl\| \sup_{j\geqslant 0} |\mathcal{A}_j^\lambda(f)| \Bigr\|_{L^2(\mathbf{G}_0^\#)} \leqslant \frac{C_\varepsilon}{\lambda} \|f\|_{L^2(\mathbf{G}_0^\#)} \tag{4.53}$$

and

$$\Bigl\| \sup_{j\geqslant 0} |\mathcal{B}_j^\lambda(f)| \Bigr\|_{L^p(\mathbf{G}_0^\#)} \leqslant C_\varepsilon \lambda^\varepsilon \|f\|_{L^p(\mathbf{G}_0^\#)}. \tag{4.54}$$

The rest of this subsection is concerned with the proof of Lemma 4.4. In view of Lemma 4.1 with $N=\frac{1}{2}$ and $\mathcal{R}_N=\varnothing$, in proving Lemma 4.4 we may assume $\lambda\geqslant C_\varepsilon$. With $\bar{c}$ as in Lemma 4.1, we define

$$N_0=\lambda^{1/\bar{c}}, \quad \mathcal{R}_{N_0}=\{a/N_0! : a\in\mathbf{Z}^{d'}\} \quad \text{and} \quad J_{N_0,\mathcal{R}_{N_0}}=N_0^2. \tag{4.55}$$

Property (4.8) is clearly satisfied if λ is sufficiently large. For $j<J_{N_0,\mathcal{R}_{N_0}}$, let $\mathcal{A}_j^\lambda\equiv 0$ and $\mathcal{B}_j^\lambda\equiv\mathcal{M}_j$. By Lemma 4.2,

$$\Big\|\sup_{j\in[0,J_{N_0,\mathcal{R}_{N_0}})\cap\mathbf{Z}}|\mathcal{B}_j^\lambda(f)|\Big\|_{L^p(\mathbf{G}_0^\#)}\leqslant C(\log\lambda)^2\|f\|_{L^p(\mathbf{G}_0^\#)},$$

which is better than estimate (4.54). For $j\geqslant J_{N_0,\mathcal{R}_{N_0}}$, let

$$\mathcal{A}_j^\lambda\equiv\mathcal{M}_j-\mathcal{M}_j^{N_0,\mathcal{R}_{N_0}}\quad\text{and}\quad\mathcal{B}_j^\lambda\equiv\mathcal{M}_j^{N_0,\mathcal{R}_{N_0}}.$$

By Lemma 4.1 and definition (4.55),

$$\Big\|\sup_{j\geqslant J_{N_0,\mathcal{R}_{N_0}}}|\mathcal{A}_j^\lambda(f)|\Big\|_{L^2(\mathbf{G}_0^\#)}\leqslant\frac{C}{\lambda}\|f\|_{L^2(\mathbf{G}_0^\#)},$$

which gives the bound (4.53). To complete the proof of Lemma 4.4, it suffices to show that

$$\Big\|\sup_{j\geqslant J_{N_0,\mathcal{R}_{N_0}}}|\mathcal{M}_j^{N_0,\mathcal{R}_{N_0}}(f)|\Big\|_{L^p(\mathbf{G}_0^\#)}\leqslant C_p(\log N_0)\|f\|_{L^p(\mathbf{G}_0^\#)}.\tag{4.56}$$

To prove the bound (4.56), we use estimates (7.11) and the change of coordinates (4.28). By the Fourier inversion formula, as in equation (4.26),

$$\begin{aligned}\mathcal{M}_j^{N_0,\mathcal{R}_{N_0}}(f)(m,u)=\sum_{(n,v)\in\mathbf{G}_0^\#}&f(n,v)\Omega_j(m-n)\eta_{2^{2j}/N_0}(u-v-R_0(m-n,n))\\&\times\sum_{r\in\mathcal{R}_{N_0}\cap[0,1)^{d'}}e^{2\pi i(u-v-R_0(m-n,n))\cdot r}.\end{aligned}\tag{4.57}$$

Let $Q=N_0!$. Definition (4.55) shows that

$$\sum_{r\in\mathcal{R}_{N_0}\cap[0,1)^{d'}}e^{2\pi i(u-v-R_0(m-n,n))\cdot r}=\delta_Q(u-v-R_0(m-n,n)),$$

where

$$\delta_Q:\mathbf{Z}^{d'}\longrightarrow\mathbf{Z},\quad\delta_Q(u)=\begin{cases}Q^{d'},&\text{if }u/Q\in\mathbf{Z}^{d'},\\0,&\text{if }u/Q\notin\mathbf{Z}^{d'}.\end{cases}\tag{4.58}$$

We use the change of coordinates $\Phi_Q:\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]\to\mathbf{G}_0^\#$ described in (4.28). Let $F((n',v'),(\nu,\beta))=f(\Phi_Q((n',v'),(\nu,\beta)))$ and

$$G_j((m',u'),(\mu,\alpha))=\mathcal{M}_j^{N_0,\mathcal{R}_{N_0}}(f)(\Phi_Q((m',u'),(\mu,\alpha))).$$

Formula (4.57) is equivalent to

$$\begin{aligned}&G_j((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^{\#}}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F((n',v'),(\nu,\beta))\Omega_j(Q(m'-n')+E_1)\\&\qquad\qquad\times\eta_{2^{2j}/N_0}(Q^2(u'-v'-R_0(m'-n',n'))+E_2)\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu)),\end{aligned}$$

where $E_1=\mu-\nu$ and

$$E_2=(\alpha-\beta-R_0(\mu-\nu,\nu))+Q(R_0(\mu,m'-n')-R_0(m'-n',\nu)).$$

Clearly, $2^j\geqslant 2^{N_0^2}$ and $Q\leqslant 2^{N_0^{3/2}}$, and thus $2^j\geqslant Q^{10}$. Also,

$$|E_1|\leqslant CQ\quad\text{and}\quad |E_2|\leqslant C2^jQ\ \text{ if }\ |m'-n'|\leqslant\frac{C2^j}{Q}.$$

Let

$$\begin{aligned}&\widetilde{G}_j((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^{\#}}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F((n',v'),(\nu,\beta))\Omega_j(Q(m'-n'))\\&\qquad\qquad\times\eta_{2^{2j}/N_0}(Q^2(u'-v'-R_0(m'-n',n')))\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu)).\end{aligned}$$

In view of the estimates above on $|E_1|$ and $|E_2|$, we have

$$\begin{aligned}&|G_j((m',u'),(\mu,\alpha))-\widetilde{G}_j((m',u'),(\mu,\alpha))|\\&\quad\leqslant C\frac{N_0Q}{2^j}\sum_{(n',v')\in\mathbf{G}_0^{\#}}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}|F((n',v'),(\nu,\beta))|Q^{-d}Q^{-2d'}\\&\qquad\qquad\times\left(\frac{2^j}{Q}\right)^{-d}\mathbf{1}_{[0,C2^j/Q]}(|m'-n'|)\phi_{2^{2j}/Q^2N_0}(u'-v'-R_0(m'-n',n')),\end{aligned}$$

where ϕ is as in definition (7.7). Thus,

$$\sum_{j\geqslant N_0^2}\|G_j-\widetilde{G}_j\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C\|F\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}.$$

For the bound (4.56), it remains to prove that

$$\Big\|\sup_{j\geqslant N_0^2}|\widetilde{G}_j|\Big\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C_p(\log N_0)\|F\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}. \tag{4.59}$$

For this, we notice that the function $\widetilde{G}_j$ is obtained as the composition of the operator

$$f\longmapsto Q^{-d}Q^{-2d'}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}f(\nu,\beta)\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu))$$

acting on functions $f\colon\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}\to\mathbf{C}$, which is clearly bounded on $L^p(\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'})$, followed by an average dominated by the maximal operator $\mathcal{M}_*^{N_0}$ of Lemma 7.1. The bound (4.59) follows from estimates (7.11).

5. The ergodic theorem

In this section we prove Theorem 1.3. We first reduce matters to proving Theorem 5.1 below (in fact, we only need this theorem for a special group $\mathbf{G}^\#$ and a special polynomial mapping P). Then we use a maximal ergodic theorem (which follows from Theorem 1.1 and a transference argument) and adapt a limiting argument of Bourgain [5].

5.1. Preliminary reductions and a maximal ergodic theorem

Assume that (X,μ) is a finite measure space. A result equivalent to Theorem 1.3 can be formulated in terms of the action of the step 2 discrete nilpotent group $\mathbf{G}^\#$ defined in (2.2) and (2.4), corresponding to a bilinear mapping $R:\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d'}$. Suppose that $\mathbf{G}^\#$ acts on X via measure-preserving transformations, and denote the action $\mathbf{G}^\#\times X\to X$ by $(g,x)\mapsto g\cdot x$. For a polynomial map $P:\mathbf{Z}^d\to\mathbf{Z}^{d'}$ of degree at most 2, and $F\in L^p(X)$, $p\in(1,\infty]$, define the averages

$$M_r(F)(x)=\frac{1}{|B_r\cap\mathbf{Z}^d|}\sum_{n\in B_r\cap\mathbf{Z}^d}F((n,P(n))\cdot x). \tag{5.1}$$

THEOREM 5.1. *For every $F\in L^p(X)$, $p\in(1,\infty)$, there exists $F_*\in L^p(X)$ such that*

$$\lim_{r\to\infty}M_r(F)=F_* \quad \textit{almost everywhere and in } L^p. \tag{5.2}$$

Moreover, if the action of the subgroup $(q\mathbf{Z})^d\times(q\mathbf{Z})^{d'}$ is ergodic on X for every integer $q\geqslant1$, then

$$F_*=\frac{1}{\mu(X)}\int_X F\,d\mu. \tag{5.3}$$

We now prove the equivalence of Theorems 1.3 and 5.1, and reduce matters to proving Theorem 5.1 on a special discrete group $\mathbf{G}^\#$ with special polynomial map P_0. We first show that Theorem 1.3 implies Theorem 5.1. Assume that $\mathbf{G}^\#$ is as in Theorem 5.1 and acts on X via measure-preserving transformations. For $g\in\mathbf{G}^\#$ define the transformation $T_g:X\to X$ by $T_g(x)=g\cdot x$. Let $\{g_j\}_{j=1}^d\cup\{h_k\}_{k=1}^{d'}$ denote the standard basis of $\mathbf{Z}^d\times\mathbf{Z}^{d'}$, and let $T_j=T_{g_j}$ and $S_k=T_{h_k}$. For $n=(n_1,...,n_d)\in\mathbf{Z}^d$, $m=(m_1,...,m_{d'})\in\mathbf{Z}^{d'}$ it follows from the definitions that

$$\prod_{j=1}^d T_j^{n_j}\prod_{k=1}^{d'}S_k^{m_k}=T_{(n,\,m+Q_0(n))}, \tag{5.4}$$

where $Q_0:\mathbf{Z}^d\to\mathbf{Z}^{d'}$ is a polynomial mapping of degree 2. Thus, the averages in (5.1) reduce to those in (1.12) associated with the polynomial map $Q(n)=P(n)-Q_0(n)$. Also, it is clear from equation (5.4) that the family of transformations $\{T_j^q\}_{j=1}^d\cup\{S_k^q\}_{k=1}^{d'}$ generates the subgroup $\mathbf{G}_q^{\#}=(q\mathbf{Z})^d\times(q\mathbf{Z})^{d'}$, hence the ergodicity of the action of the subgroup implies that of the family $\{T_j^q\}_{j=1}^d\cup\{S_k^q\}_{k=1}^{d'}$.

We now start the proof of Theorem 1.3. Notice that the coefficients of the polynomials $Q_l\colon\mathbf{Z}^d\to\mathbf{Z}$ of degree at most 2 must be integers or half integers. Writing $n_j=2n'_j+\varepsilon_j$, $1\leqslant j\leqslant d$, for some fixed residue classes ε_j modulo 2, it follows that the average in (1.12) can be written as a linear combination of 2^d averages, where the exponents are polynomials with integer coefficients. Thus one can assume that the polynomial mapping Q in (1.11) has integer coefficients. Also, one may write

$$\prod_{l=1}^{d'} S_l^{Q_l(n)} = \bar S_0 \prod_{j=1}^{d} \bar S_j^{n_j} \prod_{1\leqslant j\leqslant k\leqslant d} \bar S_{jk}^{n_jn_k}, \tag{5.5}$$

by expanding $S_l^{Q_l(n)}$ into a product of factors with monomial exponents n_j and n_jn_k, and collecting all the resulting factors with a given exponent. If one puts $\bar T_j=T_j\bar S_j$, $1\leqslant j\leqslant d$, then the transformations $\bar T_j$, $1\leqslant j\leqslant d$, and $\bar S_{jk}$, $1\leqslant j\leqslant k\leqslant d$, satisfy the commutator relations (1.10). Moreover, the ergodicity of the family $\{\bar T_j^q\}_{j=1}^d\cup\{\bar S_{jk}^q\}_{1\leqslant j\leqslant k\leqslant d}$ implies that of the family $\{T_j^q\}_{j=1}^d\cup\{S_l^q\}_{l=1}^{d'}$. Thus, it is enough to prove Theorem 1.3 for the special polynomial map

$$Q_0\colon \mathbf{Z}^d \longrightarrow \mathbf{Z}^{d(d+1)/2} \quad \text{with} \quad Q_0^{jk}(n_1,...,n_d)=n_jn_k,\ \ 1\leqslant j\leqslant k\leqslant d. \tag{5.6}$$

We identify the group generated by the transformations $\bar T_j$, $1\leqslant j\leqslant d$, and $\bar S_{jk}$, $1\leqslant j\leqslant k\leqslant d$, as an isomorphic image of a step 2 nilpotent group $\mathbf{G}^{\#}$ on $\mathbf{Z}^d\times\mathbf{Z}^{d^2}$. More precisely, it follows from the relations (1.10) that

$$\prod_{j=1}^{d} T_j^{n_j} \prod_{k=1}^{d} T_k^{n'_k} = \prod_{j=1}^{d} T_j^{n_j+n'_j} \prod_{1\leqslant k<j\leqslant d} [T_j,T_k]^{n_jn'_k}. \tag{5.7}$$

This implies that the group $\overline{\mathbf{G}}_0^{\#}$ defined by the bilinear form $\bar R_0\colon\mathbf{Z}^d\times\mathbf{Z}^d\to\mathbf{Z}^{d^2}$ with components

$$\bar R_0^{jk}(n,n') = \begin{cases} n_jn'_k, & \text{if } 1\leqslant k<j\leqslant d,\\ 0, & \text{if } 1\leqslant j\leqslant k\leqslant d,\end{cases} \tag{5.8}$$

acts on X via

$$(n,m)\cdot x = \prod_{j=1}^{d} \bar T_j^{n_j} \prod_{1\leqslant k<j\leqslant d} [\bar T_j,\bar T_k]^{m_{jk}} \prod_{1\leqslant j\leqslant k\leqslant d} \bar S_{jk}^{m_{jk}}(x), \tag{5.9}$$

where $n=(n_j)_{j=1}^d$ and $m=(m_{jk})_{j,k=1}^d$. In terms of this action, the averages in (1.12) take the form

$$A_r(F)(x)=\frac{1}{|B_r\cap\mathbf{Z}^d|}\sum_{n\in B_r\cap\mathbf{Z}^d}F((n,0,P_0(n))\cdot x). \tag{5.10}$$

Thus, Theorem 1.3 reduces to Theorem 5.1 in the special case $d'=d^2$,

$$P(n)=\sum_{1\leqslant j\leqslant k\leqslant d}n_jn_k\cdot e_{jk} \quad\text{and}\quad R(n,n')=\sum_{1\leqslant k<j\leqslant d}n_jn'_k\cdot e_{jk}, \tag{5.11}$$

where $\{e_{jk}\}_{j,k=1}^d$ denotes a standard orthonormal basis of $\mathbf{R}^{d^2}$.

We conclude this subsection with a maximal ergodic theorem, which follows from Theorem 1.1 and a general transference argument.

Theorem 5.2. (Maximal ergodic theorem) *With the notation as in Theorem* 5.1, *let* $\mathcal{M}(F)(x)=\sup_{r>0}|M_rF(x)|$. *Then*

$$\|\mathcal{M}(F)\|_{L^p(X)}\leqslant C_p\|F\|_{L^p(X)}. \tag{5.12}$$

Using Theorem 5.2 and the Lebesgue dominated convergence theorem, it suffices to prove the almost everywhere convergence in statement (5.2). We can also assume in Theorem 5.1 that F is in a suitable dense subspace of $L^p(X)$, such as $L^\infty(X)$.

5.2. Pointwise convergence

Assume that $F\in L^\infty(X)$ and, for a given $1<\delta\leqslant 2$, define the averages

$$\mathcal{M}_j^\delta(F)(x)=\frac{1}{\|\Omega^\delta\|_{L^1}\delta^j}\sum_{n\in\mathbf{Z}^d}\Omega^\delta(n/\delta^j)F((n,P(n))\cdot x), \tag{5.13}$$

where $\Omega^\delta:\mathbf{R}^d\to[0,1]$ is a smooth function, such that $\Omega^\delta(y)=1$ for $|y|\leqslant 1$ and $\Omega^\delta(y)=0$ for $|y|\geqslant\delta$. For a given $r>1$, let j be such that $\delta^j\leqslant r<\delta^{j+1}$ and compare the averages $M_r(F)$ and $\mathcal{M}_j^\delta(F)$. Since $F\in L^\infty$, it follows easily that for any $x\in X$,

$$|M_r(F)(x)-\mathcal{M}_j^\delta(F)(x)|\leqslant C_d(\delta^{-j}+\delta^d-1)\|F\|_{L^\infty}.$$

Thus, it suffices to show that for each $1<\delta\leqslant 2$ the averages $\mathcal{M}_j^\delta(F)$ converge almost everywhere as $j\to\infty$. For simplicity of notation, we drop the superscript δ and write $\mathcal{M}_j(F)=\mathcal{M}_j^\delta(F)$.

Next, we identify subspaces of $L^2(X)$ on which the convergence of $\mathcal{M}_j(F)$ is immediate. For integers $q\geqslant 1$, let $\mathbf{G}_q^\#=(q\mathbf{Z})^d\times(q\mathbf{Z})^{d'}$, i.e. the subgroup of points with all the coordinates divisible by q. Define the corresponding space of invariant functions by

$$L_q^2(X)=\{F\in L^2(X):T_gF=F \text{ for all } g\in\mathbf{G}_q^\#\}\quad\text{and}\quad L_{\mathrm{inv}}^2(X)=\overline{\bigcup_{q\geqslant 1}L_q^2(X)}, \tag{5.14}$$

where $T_g F(x) = F(g\cdot x)$. Notice that $L^2_{q_1}(X) \subseteq L^2_{q_2}(X)$ if q_1 divides q_2, hence $L^2_{\mathrm{inv}}(X)$ is a closed subspace of $L^2(X)$.

LEMMA 5.3. *Assume that* $q\geqslant 1$ *and let* $F\in L^2_q(X)$. *Then, for every* $x\in X$,

$$\lim_{j\to\infty} \mathcal{M}_j(F)(x) = q^{-d} \sum_{\nu\in(\mathbf{Z}/q\mathbf{Z})^d} F((\nu, P(\nu))\cdot x) \tag{5.15}$$

Proof. If $n\equiv\nu \pmod q$, then $(n, P(n))\equiv(\nu, P(\nu)) \pmod q$ (see (5.11)), hence there is a $g\in\mathbf{G}^{\#}_q$ such that $(n, P(n)) = g\cdot(\nu, P(\nu))$. Thus $F((n, P(n))\cdot x) = F((\nu, P(\nu))\cdot x)$, since $F\in L^2_q(X)$. In view of the definitions, it is enough to show that for every $\nu\in(\mathbf{Z}/q\mathbf{Z})^d$,

$$\lim_{j\to\infty} \frac{1}{\|\Omega^\delta\|_{L^1}\,\delta^j} \sum_{n\equiv\nu \,(\mathrm{mod}\ q)} \Omega^\delta(n/\delta^j) = q^{-d},$$

which is an elementary observation. □

If for each q the action of $\mathbf{G}^{\#}_q$ on X is ergodic, then $L^2_{\mathrm{inv}}(X)$ contains only constant functions. Thus, for statements (5.2) and (5.3), it suffices to prove that for $F\in L^2_{\mathrm{inv}}(X)^{\perp}$,

$$\lim_{j\to\infty} \mathcal{M}_j(F)(x) = 0 \quad \text{for almost every } x\in X. \tag{5.16}$$

We now identify a dense subspace of the orthogonal complement of $L^2_q(X)$.

LEMMA 5.4. *Assume that* $q\geqslant 1$. *Then*

$$L^2_q(X)^{\perp} = \overline{\mathrm{Span}\{T_g H - H : g\in\mathbf{G}^{\#}_q \text{ and } H\in L^\infty(X)\}}, \tag{5.17}$$

where Span S *denotes the subspace spanned by the set* S.

Proof. Let $F\in L^2(X)$ and assume that for all $H\in L^\infty(X)$ and $g\in\mathbf{G}^{\#}_q$,

$$\langle F, T_g H - H\rangle = 0.$$

That is, for every $g\in\mathbf{G}^{\#}_q$,

$$\langle T_{g^{-1}} F - F, H\rangle = 0 \quad \text{for all } H\in L^\infty(X),$$

which means that $T_{g^{-1}}F = F$ for all $g\in\mathbf{G}^{\#}_q$, so $F\in L^2_q(X)$. This proves the lemma. □

Following an idea described in [3], we will show statement (5.16) by proving L^2 bounds for a family of truncated maximal functions. We will use the following construction: let $\mathcal{L}_j$, $j\in\mathbf{N}$, be a family of bounded linear operators on $L^2(X)$, and let $\{j_k\}_{k\in\mathbf{Z}_+}$ be an increasing sequence of natural numbers. Then we define the maximal operators

$$\mathcal{L}^*_k(F)(x) = \max_{j_k\leqslant j< j_{k+1}} |\mathcal{L}_j(F)(x)|.$$

Let $F\in L^2_{\rm inv}(X)^\perp$, and assume indirectly that for a set of positive measure

$$\lim_{j\to\infty}\mathcal{M}_j(F)(x)\neq 0.$$

Then, there exists $\varepsilon>0$ such that

$$\mu\Big\{x\in X:\limsup_{j\to\infty}|\mathcal{M}_j(F)(x)|>\varepsilon\Big\}>\varepsilon.$$

It is now easy to see that there is an increasing sequence $\{j_k\}_{k\in\mathbf{Z}_+}$ such that

$$\|\mathcal{M}_k^*(F)\|^2_{L^2(X)}\geqslant\frac{\varepsilon^3}{2}\tag{5.18}$$

for all $k\in\mathbf{Z}_+$. Moreover, the sequence $\{j_k\}_{k\in\mathbf{Z}_+}$ can be chosen to be rapidly increasing, so we may assume that $j_{k+1}\geqslant 3j_k$.

Let $\widetilde{\chi}:\mathbf{R}\to[0,1]$ denote a smooth function supported in $[-2,2]$ and equal to 1 in $[-1,1]$. For $x\in X$ and $L\gg 1$, we define

$$f_{L,x}(g)=F(g\cdot x)\chi_L(g),\tag{5.19}$$

where $\chi_L:\mathbf{R}^d\times\mathbf{R}^{d^2}\to[0,1]$ is given by

$$\chi_L(m,u)=\widetilde{\chi}\bigg(\frac{|m|}{L}\bigg)\widetilde{\chi}\bigg(\frac{|u|}{L^2}\bigg)$$

(recall that $\mathbf{G}^\#=\mathbf{Z}^d\times\mathbf{Z}^{d^2}$ as sets). Clearly, $\|\chi_L\|_{L^1(\mathbf{G}^\#)}\approx L^{d+2d^2}$. For $f:\mathbf{G}^\#\to\mathbf{C}$, $j\geqslant 0$ and $\delta\in(1,2]$, we define, as in formula (5.13),

$$\widetilde{\mathcal{M}}_j(f)(g)=\frac{1}{\|\Omega^\delta\|_{L^1}\,\delta^j}\sum_{n\in\mathbf{Z}^d}\Omega^\delta(n/\delta^j)\,f((n,P(n))\cdot g).\tag{5.20}$$

Using the definitions, for any $k\in\mathbf{Z}_+$ and $L\geqslant L_k$ large enough,

$$\|\mathcal{M}_k^*(F)\|^2_{L^2(X)}\leqslant\frac{C}{L^{d+2d^2}}\int_X\|\widetilde{\mathcal{M}}_k^*(f_{L,x})\|^2_{L^2(\mathbf{G}^\#)}\,d\mu(x).$$

We assume from now on that the sequence $j_1<j_2<...$ is fixed. To summarize, for the statement (5.16), it suffices to prove Lemma 5.5 below.

LEMMA 5.5. *Assume that $F\in L^2_{\rm inv}(X)^\perp$ and define $f_{L,x}$ as in* (5.19). *Then for every $\varepsilon>0$ and $\delta\in(1,2]$ (see* (5.13)) *there exist $k=k(F,\varepsilon,\delta)$ and $L(j_{k+1},F,\varepsilon,\delta)$ such that*

$$\frac{1}{L^{d+2d^2}}\int_X\Big\|\sup_{j_k\leqslant j<j_{k+1}}|\widetilde{\mathcal{M}}_j(f_{L,x})|\Big\|^2_{L^2(\mathbf{G}^\#)}\,d\mu(x)\leqslant\varepsilon\tag{5.21}$$

for any $L\geqslant L(j_{k+1},F,\varepsilon,\delta)$.

We now show how to reduce Lemma 5.5 to Lemma 5.6 below. We may assume that $\|F\|_{L^2(X)}=1$, so

$$\frac{1}{L^{d+2d^2}}\int_X \|f_{L,x}\|^2_{L^2(\mathbf{G}^\#)}\,d\mu(x)\leqslant C\|F\|_{L^2(X)}\leqslant C \quad \text{for any } L\geqslant 1. \tag{5.22}$$

Also, for $f\in L^2(\mathbf{G}^\#)$, we may redefine

$$\widetilde{\mathcal{M}}_j(f)(g)=2^{-dj}\sum_{n\in\mathbf{Z}^d}\Omega^\delta(n/2^j)f((n,P(n))\cdot g), \tag{5.23}$$

where $\Omega^\delta:\mathbf{R}^d\to[0,1]$ is a smooth function, $\Omega^\delta(y)=1$ for $|y|\leqslant c_0$ and $\Omega^\delta(y)=0$ for $|y|\geqslant c_0\delta$, $1\leqslant c_0\leqslant 2$.

We will use the notation and the results of §4.1, especially the remark following Lemma 4.1. Assume that $\varepsilon>0$ and $\eta\in(1,2]$ are fixed. We now relate the averages $\widetilde{\mathcal{M}}_j(f)$ in equation (5.23) and $\mathcal{M}_j(f)$ in equation (4.2). We identify $\mathbf{G}^\#$ and $\mathbf{G}_0^\#$ with $\mathbf{Z}^d\times\mathbf{Z}^{d^2}$. By taking the Fourier transform in the central variable, for θ in $\mathbf{R}^{d^2}$ we have

$$\widehat{\mathcal{M}}_j(f)(m,\theta)=\sum_{n\in\mathbf{Z}^d}\Omega_j^\delta(m-n)\hat{f}(n,\theta)e^{-2\pi iR_0(m-n,n)\cdot\theta},$$
$$\widehat{\widetilde{\mathcal{M}}}_j(f)(m,\theta)=\sum_{n\in\mathbf{Z}^d}\Omega_j^\delta(m-n)\hat{f}(n,\theta)e^{-2\pi i(-P(n-m)-R(n-m,m))\cdot\theta},$$

where $\Omega_j^\delta(x)=2^{-dj}\Omega^\delta(-x/2^j)$. For $N=N(\varepsilon,\delta)$ sufficiently large, let

$$\mathcal{R}_N=\{a/q\in\mathbf{Q}^{d^2}:q\leqslant N \text{ and } (a,q)=1\}.$$

For $j\geqslant N$ define, as in equations (4.6) and (4.7),

$$\Psi_j^N(\theta)=\sum_{r\in\mathcal{R}_N}\psi(2^{2j}N^{-1}(\theta-r))$$

and

$$\widehat{\widetilde{\mathcal{M}}^N_j}(f)(m,\theta)=\widehat{\widetilde{\mathcal{M}}}_j(f)(m,\theta)\cdot\Psi_j^N(\theta).$$

Simple changes of variables, using (5.11), and the remark following Lemma 4.1 show that

$$\Big\|\sup_{j\geqslant N}|\widetilde{\mathcal{M}}_j(f)-\widetilde{\mathcal{M}}_j^N(f)|\Big\|_{L^2(\mathbf{G}^\#)}\leqslant\frac{\varepsilon}{C}\|f\|_{L^2(\mathbf{G}^\#)}$$

for any $f\in L^2(\mathbf{G}^\#)$, provided $N=N(\varepsilon,\delta)$ is fixed sufficiently large. Thus, using (5.22),

$$\frac{1}{L^{d+2d^2}}\int_X\Big\|\sup_{j\geqslant N}|\widetilde{\mathcal{M}}_j(f_{L,x})-\widetilde{\mathcal{M}}_j^N(f_{L,x})|\Big\|^2_{L^2(\mathbf{G}^\#)}\,d\mu(x)\leqslant\frac{\varepsilon}{2}. \tag{5.24}$$

Assume from now on that N is fixed. We examine the operator $\widetilde{\mathcal{M}}_j^N$ and, for $a/q\in\mathcal{R}_N$, $j\geqslant N$ and $f\in L^2(\mathbf{G}^\#)$, we define

$$\widehat{\widetilde{\mathcal{M}}_{j,a/q}^N}(f)(m,\theta)=\widehat{\widetilde{\mathcal{M}}}_j(f)(m,\theta)\sum_{b\in\mathbf{Z}^{d^2}}\psi(2^{2j}N^{-1}(\theta-a/q-b)). \tag{5.25}$$

Thus, for Lemma 5.5, it suffices to prove Lemma 5.6 below.

LEMMA 5.6. *Assume that* $F\in L^2_{\mathrm{inv}}(X)^\perp$, $N\geqslant 1$, $a/q\in\mathcal{R}_N$ *and* $\delta\in(1,2]$, *and define* $f_{L,x}$ *as in equation* (5.19) *and* $\widetilde{\mathcal{M}}_{j,a/q}^N$ *as in equations* (5.25) *and* (5.23). *Then, for every* $\varepsilon>0$, *there exist* $k=k(F,N,\varepsilon,\delta)$ *and* $L(j_{k+1},F,N,\varepsilon,\delta)$ *such that*

$$\frac{1}{L^{d+2d^2}}\int_X\Big\|\sup_{j_k\leqslant j<j_{k+1}}|\widetilde{\mathcal{M}}_{j,a/q}^N(f_{L,x})|\Big\|^2_{L^2(\mathbf{G}^\#)}\,d\mu(x)\leqslant\varepsilon \tag{5.26}$$

for any $L\geqslant L(j_{k+1},F,N,\varepsilon,\delta)$.

Proof. As in formula (4.26), by the Fourier inversion formula,

$$\widetilde{\mathcal{M}}_{j,a/q}^N(f)(m,u)=\sum_{(n,v)\in\mathbf{Z}^d\times\mathbf{Z}^{d^2}}f(n,v)\Omega_j^\delta(m-n)\eta_{2^{2j}/N}(u-v+P(n-m)+R(n-m,m))\\ \times e^{2\pi i(u-v+P(n-m)+R(n-m,m))\cdot a/q}, \tag{5.27}$$

where $\eta\in\mathcal{S}(\mathbf{R}^{d^2})$ is defined as in formula (4.26) and $\eta_r(s)=r^{-d^2}\eta(s/r)$, $r\geqslant 1$. As in §7, we define $\phi\colon\mathbf{R}^{d^2}\to[0,1]$ by $\phi(s)=(1+|s|^2)^{-(d^2+d+1)}$, and $\phi_r(s)=r^{-d^2}\phi(s/r)$. Then

$$|\widetilde{\mathcal{M}}_{j,a/q}^N(f)(m,u)|\leqslant C_N\sum_{(n,v)\in\mathbf{Z}^d\times\mathbf{Z}^{d^2}}|f(n,v)|\Omega_j^\delta(m-n)\phi_{2^{2j}}(u-v+R(n-m,m)), \tag{5.28}$$

so the maximal function $f\mapsto\sup_{j_k\leqslant j\leqslant j_{k+1}}|\widetilde{\mathcal{M}}_{j,a/q}^N(f)|$ is bounded on $L^2(\mathbf{G}^\#)$ (compare with Lemma 7.1). Thus, using Lemma 5.4 and statement (5.22), in proving Lemma 5.6 we may assume that

$$F(x)=H(g_0\cdot x)-H(x)\quad\text{for some } g_0\in\mathbf{G}_q^\#\ \text{ and }\ H\in L^\infty(X)\ \text{ with }\ \|H\|_{L^\infty}=1. \tag{5.29}$$

We may also replace the function η with a smooth function $\tilde{\eta}$ compactly supported in the set $\{s\in\mathbf{R}^{d^2}:|s|\leqslant N'(\varepsilon,N)\}$; this is due to the fact that the bound (5.28) gains an additional small factor on the right-hand side for the part of the operator corresponding to $\eta-\tilde{\eta}$.

Using equations (5.19) and (5.29),

$$f_{L,x}(g)=\chi_L(g)[H(g_0g\cdot x)-H(g\cdot x)]. \tag{5.30}$$

It suffices to prove that for k and L as in Lemma 5.6, and $f_{L,x}$ as in equation (5.30),

$$\frac{1}{L^{d+2d^2}}\int_X \Big\| \sup_{j_k \leqslant j < j_{k+1}} |\mathcal{M}'_j(f_{L,x})| \Big\|^2_{L^2(\mathbf{G}^\#)} d\mu(x) \leqslant \frac{\varepsilon}{2}, \tag{5.31}$$

where $\mathcal{M}'_j(f)$ is defined as in equation (5.27), with $\tilde{\eta}$ replacing η.

We define the kernels $K_j\colon \mathbf{G}^\# \to \mathbf{C}$ by

$$K_j(n,v) = \Omega_j^\delta(-n)\tilde{\eta}_{2^{2j}/N}(-v+P(n))e^{2\pi i(-v+P(n))\cdot a/q}. \tag{5.32}$$

So, using formula (5.27),

$$\mathcal{M}'_j(f)(m,u) = \sum_{(n,v)\in\mathbf{G}^\#} f(n,v)K_j((n,v)\cdot(m,u)^{-1}).$$

Using equation (5.30) and simple changes of variables, it follows that

$$\mathcal{M}'_j(f_{L,x})(g) = \sum_{h\in\mathbf{G}^\#} H(hg\cdot x)[\chi_L(g_0^{-1}hg)K_j(g_0^{-1}h) - \chi_L(hg)K_j(h)] \tag{5.33}$$

for any $g\in\mathbf{G}^\#$. We now use (5.29), i.e. $\|H\|_{L^\infty}=1$. Since $g_0^{-1}\in\mathbf{G}_q^\#$, the oscillatory parts of $K_j(h)$ and $K_j(g_0^{-1}h)$ agree. Simple estimates then show that (with $h=(n,v)$)

$$|\chi_L(g_0^{-1}hg)K_j(g_0^{-1}h) - \chi_L(hg)K_j(h)| \leqslant C(g_0,N,\varepsilon,\delta) j_k^{-1}\chi_{4L}(g)\Omega_{j+2}^\delta(n)\phi_{2^{2j}}(v),$$

if k is sufficiently large, and then L is sufficiently large compared to j_{k+1}. Thus,

$$|\mathcal{M}'_j(f_{L,x})(g)| \leqslant \frac{C(g_0,N,\varepsilon,\delta)}{j_k}\chi_{4L}(g),$$

and inequality (5.31) follows. □

6. The singular Radon transform

In this section we prove Lemma 2.8. The main ingredients are the L^2 bounds in Lemma 6.1, a super-orthogonality argument of Ionescu and Wainger [8] which reduces matters to square function estimates, and the weighted inequality in Lemma 7.4. We assume throughout this section that $d'=d^2$, and that $\mathbf{G}_0^\#$ is the discrete nilpotent group defined in §2.

6.1. L^2 estimates

Our main result in this subsection is Lemma 6.1, which is a quantitative L^2 estimate. The proof of Lemma 6.1 is based on Plancherel's theorem and Lemmas 3.2–3.4.

Let K denote the Calderón–Zygmund kernel defined in §1. Without loss of generality (compare with [14, p. 624]), we may assume that $K=\sum_{j=0}^{\infty} K_j$, where K_j is supported in the set $\{x:|x|\in[2^{j-1},2^{j+1}]\}$ and satisfies the bound

$$|x|^d|K_j(x)|+|x|^{d+1}|\nabla K_j(x)|\leqslant 1, \quad x\in\mathbf{R}^d,\ j\geqslant 1, \tag{6.1}$$

and the cancellation condition

$$\int_{\mathbf{R}^d} K_j(x)\,dx=0, \quad j\geqslant 1. \tag{6.2}$$

As in §4,

$$\mathcal{T}_0^{\#}(f)=\sum_{j=1}^{\infty}\mathcal{T}_j(f), \text{ where } \widehat{\mathcal{T}}_j(f)(m,\theta)=\sum_{n\in\mathbf{Z}^d}K_j(m-n)\hat{f}(n,\theta)e^{-2\pi iR_0(m-n,n)\cdot\theta}. \tag{6.3}$$

As in §4, let $\psi:\mathbf{R}^{d'}\to[0,1]$ denote a smooth function supported in the set $\{\xi:|\xi|\leqslant 2\}$ and equal to 1 in the set $\{\xi:|\xi|\leqslant 1\}$, $N\in\left[\frac{1}{2},\infty\right)$ a real number, $j\in[0,\infty)\cap\mathbf{Z}$ a non-negative integer and $\mathcal{R}\subseteq\mathbf{Q}^{d'}$ a discrete periodic set. As in equation (4.6), let

$$\Psi_j^{N,\mathcal{R}}(\theta)=\sum_{r\in\mathcal{R}}\psi(2^{2j}N^{-1}(\theta-r)),$$

and, by convention, $\Psi_j^{N,\varnothing}=0$. For (compactly supported) functions $f:\mathbf{G}_0^{\#}\to\mathbf{C}$ we define $\mathcal{T}_j^{N,\mathcal{R}}(f)$ by

$$\widehat{\mathcal{T}_j^{N,\mathcal{R}}}(f)(m,\theta)=\widehat{\mathcal{T}}_j(f)(m,\theta)\Psi_j^{N,\mathcal{R}}(\theta). \tag{6.4}$$

Our main lemma in this section is the following L^2 estimate.

LEMMA 6.1. (Strong L^2 bound) *As in Lemma* 4.1, *assume that* $N\in\left[\frac{1}{2},\infty\right)$, $\mathcal{R}_N\subseteq\mathbf{Q}^{d'}$ *is a discrete periodic set and* $J_{N,\mathcal{R}_N}\in[0,\infty)$ *is a real number with the properties*

$$\begin{aligned}&\{a/q: q\in[1,N] \text{ and } (a,q)=1\}\subseteq\mathcal{R}_N,\\ &2^{J_{N,\mathcal{R}_N}}\geqslant\Big[100\max_{a/q\in\mathcal{R}_N \text{ and } (a,q)=1} q\Big]^4.\end{aligned} \tag{6.5}$$

Then

$$\Big\|\sum_{j\geqslant J_{N,\mathcal{R}_N}}(\mathcal{T}_j-\mathcal{T}_j^{N,\mathcal{R}_N})(f)\Big\|_{L^2(\mathbf{G}_0^{\#})}\leqslant C(N+1)^{-\bar{c}}\|f\|_{L^2(\mathbf{G}_0^{\#})} \tag{6.6}$$

for any $N\geqslant 0$, *where* $\bar{c}=\bar{c}(d)>0$.

Proof. Notice that the case $N=\frac{1}{2}$, $\mathcal{R}_N=\varnothing$ and $J_{N,\mathcal{R}_N}=0$ corresponds to L^2 boundedness of the operator $\mathcal{T}_0^{\#}$. For $\theta\in\mathbf{R}^{d'}$ and (compactly supported) functions $g:\mathbf{Z}^d\to\mathbf{C}$, let

$$\mathcal{U}_j^\theta(g)(m)=\sum_{n\in\mathbf{Z}^d}K_j(m-n)g(n)e^{-2\pi iR_0(m-n,n)\cdot\theta}. \tag{6.7}$$

By Plancherel's theorem,

$$\begin{aligned}\bigg\|\sum_{j\geqslant J_{N,\mathcal{R}_N}}&(\mathcal{T}_j-\mathcal{T}_j^{N,\mathcal{R}_N})(f)\bigg\|_{L^2(\mathbf{G}_0^{\#})}^2\\&=\int_{[0,1)^{d'}}\sum_{m\in\mathbf{Z}^d}\bigg|\sum_{j\geqslant J_{N,\mathcal{R}_N}}(1-\Psi_j^{N,\mathcal{R}_N}(\theta))\mathcal{U}_j^\theta(\hat{f}(\cdot\,,\theta))(m)\bigg|^2d\theta.\end{aligned}$$

Using Plancherel's theorem again, for the bound (6.6) it suffices to prove that

$$\bigg\|\sum_{j\geqslant J_{N,\mathcal{R}_N}}(1-\Psi_j^{N,\mathcal{R}_N}(\theta))\mathcal{U}_j^\theta\bigg\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}\leqslant C(N+1)^{-\bar{c}} \tag{6.8}$$

for any $\theta\in\mathbf{R}^{d'}$ fixed.

Assume that $\theta\in\mathbf{R}^{d'}$ is fixed. As in §4, for any $j\geqslant J_{N,\mathcal{R}_N}$ we use the approximation (4.17) with $\Lambda=2^{(2-\delta)j}$ and $\delta=\delta(d)>0$ sufficiently small. Thus, there are irreducible 1-fractions $a_{l_1l_2}^j/q_{l_1l_2}^j$ such that

$$1\leqslant q_{l_1l_2}^j\leqslant 2^{(2-\delta)j}\quad\text{and}\quad|\theta_{l_1l_2}-a_{l_1l_2}^j/q_{l_1l_2}^j|\leqslant\frac{C}{2^{(2-\delta)j}q_{l_1l_2}^j}. \tag{6.9}$$

We fix these irreducible 1-fractions $a_{l_1l_2}^j/q_{l_1l_2}^j$ and partition the set $\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty)$ into two subsets:

$$I_1=\Big\{j\in\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty):\max_{l_1,l_2=1,\dots,d}q_{l_1l_2}^j>2^{j/6d'}\Big\}$$

and

$$I_2=\Big\{j\in\mathbf{Z}\cap[J_{N,\mathcal{R}_N},\infty):\max_{l_1,l_2=1,\dots,d}q_{l_1l_2}^j\leqslant 2^{j/6d'}\Big\}.$$

For $j\in I_1$, we use Lemma 3.2:

$$\bigg\|\sum_{j\in I_1}(1-\Psi_j^{N,\mathcal{R}_N}(\theta))\mathcal{U}_j^\theta\bigg\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}\leqslant\sum_{j\in I_1}2^{-\delta'j}\leqslant C(N+1)^{-\bar{c}},$$

as desired.

For $j\in I_2$, let a_j/q_j denote the irreducible d'-fraction with the property that $a_j/q_j=(a_{l_1l_2}^j/q_{l_1l_2}^j)_{l_1,l_2=1,\dots,d}$. In view of (6.9) and the definition of I_2,

$$1\leqslant q_j\leqslant 2^{j/6}\quad\text{and}\quad|\theta-a_j/q_j|\leqslant\frac{C}{2^{(2-\delta)j}}. \tag{6.10}$$

We recall (see property (4.20)) that if $j,j'\in I_2$ and $j,j'\geqslant C$ then

$$\text{either}\quad a_j/q_j=a_{j'}/q_{j'}\quad\text{or}\quad |q_j/q_{j'}|\notin\left[\tfrac{1}{2},2\right]. \tag{6.11}$$

As in §4, we further partition the set I_2:

$$I_2=\bigcup_{a/q} I_2^{a/q},\quad \text{where } I_2^{a/q}=\{j\in I_2: a_j/q_j=a/q\}. \tag{6.12}$$

For $j\in I_2^{a/q}$, we show that

$$\bigg\|\sum_{j\in I_2^{a/q}}(1-\Psi_j^{N,\mathcal{R}_N}(\theta))\mathcal{U}_j^\theta\bigg\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}\leqslant C(N+q)^{-\bar{c}}. \tag{6.13}$$

This would suffice to prove (6.8), since the possible denominators q form a lacunary sequence (see (6.11)). To prove (6.13), we have two cases: $q\leqslant N$ and $q>N$. If $q\leqslant N$, we use Lemma 3.3 together with definitions (4.6) and (4.8). It follows that the left-hand side of (6.13) is dominated by

$$C\sum_{j\in\mathbf{Z}}\mathbf{1}_{[1,\infty)}(2^{2j}N^{-1}|\theta-a/q|)q^{-1/2}(1+2^{2j}|\theta-a/q|)^{-1/4}\leqslant Cq^{-1/2}N^{-1/4},$$

as desired. If $q>N$, then the left-hand side of inequality (6.13) is dominated by

$$\bigg\|\sum_{\substack{j\in I_2^{a/q}\\ 2^j\in[q^6,|\theta-a/q|^{-1/2}]}}(1-\Psi_j^{N,\mathcal{R}_N}(\theta))\mathcal{U}_j^\theta\bigg\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}+\sum_{\substack{j\in I_2^{a/q}\\ 2^j\geqslant|\theta-a/q|^{-1/2}}}\|\mathcal{U}_j^\theta\|_{L^2(\mathbf{Z}^d)\to L^2(\mathbf{Z}^d)}. \tag{6.14}$$

For the first term in (6.14), we use Lemma 3.4 for the kernels $(1-\Psi_j^{N,\mathcal{R}_N}(\theta))K_j(m)$. To control the second term in (6.14), we use Lemma 3.3. It follows that the expression (6.14) is dominated by $Cq^{-1/2}$, which suffices to prove the bound (6.13). This completes the proof of Lemma 6.1. □

6.2. An orthogonality lemma

In this subsection we review a partition of integers and a square function estimate from [8]. The point of this construction is to find a suitable decomposition of the singular integral operator and exploit the super-orthogonality (i.e. orthogonality in L^{2r}, $r\in\mathbf{Z}_+$) of the components.

Recall that for any integer $\mu\geqslant 1$, $\mathbf{Z}_\mu=\{1,...,\mu\}$. Assume that $\delta\in\left(0,\frac{1}{10}\right]$ is given and D denotes the smallest integer $\geqslant 2/\delta$. Assume that $N\geqslant 10$ is an integer. Let N' denote the smallest integer $\geqslant N^{\delta/2}$ and $V=\{p_1,p_2,...,p_\nu\}$ the set of prime numbers between $N'+1$ and N. For any $k\in\mathbf{Z}_D$, let

$$W^k(N)=\{p_{i_1}^{\alpha_{i_1}}\ldots p_{i_k}^{\alpha_{i_k}}: p_{i_l}\in V \text{ distinct and } \alpha_{i_l}\in\mathbf{Z}_D,\ l=1,...,k\}$$

and let $W(N)=\bigcup_{k\in\mathbf{Z}_D}W^k(N)$ denote the set of products of up to D factors in V, raised to powers between 1 and D.

We say that a set $W'\subseteq W(N)$ has the *orthogonality property* O if there is $k\in\mathbf{Z}_D$ and k sets $S_1,S_2,...,S_k$, $S_j=\{q_{j,1},...,q_{j,\beta(j)}\}$, $j\in\mathbf{Z}_k$, with the following properties:

(i) $q_{j,s}=p_{j,s}^{\alpha_j}$ for some $p_{j,s}\in V$ and $\alpha_j\in\mathbf{Z}_D$;

(ii) $(q_{j,s},q_{j',s'})=1$ if $(j,s)\neq(j',s')$;

(iii) for any $w'\in W'$ there are (unique) numbers $q_{1,s_1}\in S_1,...,q_{k,s_k}\in S_k$ with

$$w'=q_{1,s_1}\ldots q_{k,s_k}.$$

For simplicity of notation, we say that the set $W'=\{1\}$ has the orthogonality property O with $k=0$. The orthogonality property O is connected to Lemma 6.3 below. Notice that if a set has the orthogonality property O then all its elements have the same number of prime factors. The main result in [8, §3] is the following decomposition.

LEMMA 6.2. (Partition of integers) *With the notation above, the set $W(N)\cup\{1\}$ can be written as a disjoint union of at most $C_D(\log N)^{D-1}$ subsets with the orthogonality property O.*

Let $Q_0=[N'!]^D$ and define

$$Y_N=\{wQ': w\in W(N)\cup\{1\} \text{ and } Q'|Q_0\}. \tag{6.15}$$

Notice that for any $m\in\mathbf{Z}_N$ there is a unique decomposition $m=wQ'$, with $w\in W(N)\cup\{1\}$ and $Q'|Q_0$. In addition, $wQ'\leqslant N^{D^2}[N'!]^D\leqslant e^{N^\delta}$ if $N\geqslant C_\delta$. Thus, for $N\geqslant C_\delta$,

$$\mathbf{Z}_N\subseteq Y_N\subseteq\mathbf{Z}_{e^{N^\delta}}. \tag{6.16}$$

Let

$$W(N)\cup\{1\}=\bigcup_{s\in S}W'_s$$

denote the decomposition (guaranteed by Lemma 6.2) of $W(N)\cup\{1\}$ as a disjoint union of subsets W'_s with the orthogonality property O, where $|S|\leqslant C_D(\log N)^{2/\delta}$. Using this decomposition, we write $Y_N=\bigcup_{s\in S}Y_N^s$ (disjoint union), where

$$Y_N^s=\{wQ': w\in W'_s \text{ and } Q'|Q_0\}. \tag{6.17}$$

This is the partition of integers we will use in §6.3.

For any integer $q\geqslant 1$ let

$$P_q=\{a\in\mathbf{Z}^{d'}:(a,q)=1\}\quad\text{and}\quad \widetilde{P}_q=P_q\cap[0,q)^{d'}.$$

Let $S_1,S_2,...,S_k$ denote sets of integers $S_j=\{q_{j,1},...,q_{j,\beta(j)}\}$, $j\in\mathbf{Z}_k$. Assume that for some $\widetilde{Q}$,

$$q_{j,s}\in[2,\widetilde{Q}]\quad\text{for any } j\in\mathbf{Z}_k \text{ and } s\in\mathbf{Z}_{\beta(j)}, \tag{6.18}$$

and

$$(q_{j,s},q_{j',s'})=1\quad\text{if } (j,s)\neq(j',s'). \tag{6.19}$$

For any $j\in\mathbf{Z}_k$, let

$$T_{\{j\}}=\{a_{j,s}/q_{j,s}:s\in\mathbf{Z}_{\beta(j)}\text{ and } a_{j,s}\in P_{q_{j,s}}\}\subseteq\mathbf{Q}^{d'}$$

denote the set of irreducible fractions with denominators in S_j. Furthermore, for any set $A=\{j_1,...,j_{k'}\}\subseteq\mathbf{Z}_k$, let

$$T_A=\{r_{j_1}+...+r_{j_{k'}}:r_{j_l}\in T_{\{j_l\}}\text{ for } l\in\mathbf{Z}_{k'}\}\subseteq\mathbf{Q}^{d'}.$$

Finally, for $A=\{j_1,...,j_{k'}\}\subseteq\mathbf{Z}_k$ and any $(s_{j_1},...,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times...\times\mathbf{Z}_{\beta(j_{k'})}$, let

$$U_{A,s_{j_1},...,s_{j_{k'}}}=\{a_{j_1,s_{j_1}}/q_{j_1,s_{j_1}}+...+a_{j_{k'},s_{j_{k'}}}/q_{j_{k'},s_{j_{k'}}}:a_{j_l,s_{j_l}}\in P_{q_{j_l,s_{j_l}}}\text{ for } l\in\mathbf{Z}_{k'}\},$$

that is the subset of elements of T_A with fixed denominators $q_{j_1,s_{j_1}},...,q_{j_{k'},s_{j_{k'}}}$. If $A=\varnothing$ then, by definition, $T_A=U_A=\mathbf{Z}^{d'}$. Notice that the sets T_A and $U_{A,s_{j_1},...,s_{j_{k'}}}$ are discrete periodic subsets of $\mathbf{Q}^{d'}$. Let $\widetilde{T}_A=T_A\cap[0,1)^{d'}$ and $\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}=U_{A,s_{j_1},...,s_{j_{k'}}}\cap[0,1)^{d'}$.

Assume that $Q\geqslant 1$ is an integer with the property that

$$(Q,q_{j,s})=1\quad\text{for any } j\in\mathbf{Z}_k\text{ and } s\in\mathbf{Z}_{\beta(j)}. \tag{6.20}$$

Assume that $p\geqslant 1$ is an integer and fix

$$\gamma=(8pQ^{2p}\widetilde{Q}^{2pk})^{-1}, \tag{6.21}$$

where $\widetilde{Q}$ is such that condition (6.18) holds.

For any $r\in T_{\mathbf{Z}_k}$, let $f_r\in L^2(\mathbf{Z}^{d'})$ denote a function whose Fourier transform is supported in a γ-neighborhood of the set $\{r+a/Q:a\in\mathbf{Z}^{d'}\}$, i.e. in the set

$$\bigcup_{a\in\mathbf{Z}} r+a/Q+B(\gamma),$$

where $B(\gamma)=\{\xi:|\xi|\leqslant\gamma\}$. We assume that $f_r=f_{r+a}$ for any $a\in\mathbf{Z}^{d'}$. Let $(\mathbf{Z}^{d'},dn)$ denote the set of lattice points in $\mathbf{R}^{d'}$ with the counting measure. The main estimate in this subsection is the following lemma.

LEMMA 6.3. (Square function estimate) *With the notation above we have*

$$\begin{aligned}\int_{\mathbf{Z}^{d'}}\Big|\sum_{r\in\widetilde{T}_{\mathbf{Z}_k}} f_r(u)\Big|^{2p}\,du \\ \leqslant C_{k,p}\sum_{A=\{j_1,\dots,j_{k'}\}}\ \sum_{s_{j_1},\dots,s_{j_{k'}}}\int_{\mathbf{Z}^{d'}}\bigg(\sum_{r'\in\widetilde{T}_{c_A}}\Big|\sum_{\mu\in\widetilde{U}_{A,s_{j_1},\dots,s_{j_{k'}}}} f_{\mu+r'}(u)\Big|^2\bigg)^p du,\end{aligned} \tag{6.22}$$

where the sum in the right-hand side is taken over all sets $A=\{j_1,\dots,j_{k'}\}\subseteq\mathbf{Z}_k$ *and all* $(s_{j_1},\dots,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times\dots\times\mathbf{Z}_{\beta(j_{k'})}$. *The constant* $C_{k,p}$ *may depend only on* k *and* p.

See [8, §2] for a proof.

6.3. Proof of Lemma 2.8

In this subsection we complete the proof of Lemma 2.8. The main ingredients are the L^2 estimate in Lemma 6.1, the partition of the integers in Lemma 6.2, the square function estimate in Lemma 6.3 and the weighted estimate in Lemma 7.4. The kernels K_j satisfy conditions (6.1) and (6.2), and the operators $\mathcal{T}_j$ are as in equation (6.3). Lemma 2.8 follows by interpolation (see [8, §7]) from the following more quantitative lemma.

LEMMA 6.4. *Assume that* $2p\geqslant 4$ *is an even integer and* $\varepsilon=1/(2p-2)$. *Then, for any* $\lambda\in(0,\infty)$, *there are two linear operators* $\mathcal{A}_j^\lambda=\mathcal{A}_j^{\lambda,\varepsilon}$ *and* $\mathcal{B}_j^\lambda=\mathcal{B}_j^{\lambda,\varepsilon}$ *with* $\mathcal{T}_j=\mathcal{A}_j^\lambda+\mathcal{B}_j^\lambda$,

$$\Big\|\sum_{j\geqslant 1}\mathcal{A}_j^\lambda(f)\Big\|_{L^2(\mathbf{G}_0^\#)}\leqslant\frac{C_\varepsilon}{\lambda}\|f\|_{L^2(\mathbf{G}_0^\#)} \tag{6.23}$$

and

$$\Big\|\sum_{j\geqslant 1}\mathcal{B}_j^\lambda(f)\Big\|_{L^{2p}(\mathbf{G}_0^\#)}\leqslant C_\varepsilon\lambda^\varepsilon\|f\|_{L^{2p,1}(\mathbf{G}_0^\#)}. \tag{6.24}$$

In estimate (6.24), $L^{2p,1}(\mathbf{G}_0^\#)$ denotes the standard Lorentz space on $\mathbf{G}_0^\#$.

Proof. In view of Lemma 6.1, we may assume that $\lambda\geqslant C_\varepsilon$. With $\bar{c}$ as in Lemma 6.1, let

$$\delta=\frac{\bar{c}\varepsilon}{100}. \tag{6.25}$$

Let

$$\begin{gathered}N_0\text{ denote the smallest integer }\geqslant\lambda^{1/\bar{c}},\\ \mathcal{R}_{N_0}=\{a/q:(a,q)=1\text{ and }q\in Y_{N_0}\},\\ J_{N_0,\mathcal{R}_{N_0}}=\lambda^\varepsilon,\end{gathered} \tag{6.26}$$

where Y_{N_0} denotes the set defined in (6.15) with $\delta=\bar{c}\varepsilon/100$ as in (6.25). Property (6.5) is satisfied for $\lambda\geqslant C_\varepsilon$, using (6.16). For $j<J_{N_0,\mathcal{R}_{N_0}}$, let $\mathcal{A}_j^\lambda\equiv 0$ and $\mathcal{B}_j^\lambda\equiv\mathcal{T}_j$. Clearly,

$$\biggl\|\sum_{j\in[1,J_{N_0,\mathcal{R}_{N_0}})\cap\mathbf{Z}}\mathcal{B}_j^\lambda(f)\biggr\|_{L^{2p}(\mathbf{G}_0^\#)}\leqslant C\lambda^\varepsilon\|f\|_{L^{2p}(\mathbf{G}_0^\#)},$$

which gives the bound (6.24). For $j\geqslant J_{N_0,\mathcal{R}_{N_0}}$ let $\mathcal{A}_j^\lambda\equiv\mathcal{T}_j-\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}}$ and $\mathcal{B}_j^\lambda\equiv\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}}$, with $\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}}$ defined as in (6.4). By Lemma 6.1,

$$\biggl\|\sum_{j\geqslant J_{N_0,\mathcal{R}_{N_0}}}\mathcal{A}_j^\lambda(f)\biggr\|_{L^{2}(\mathbf{G}_0^\#)}\leqslant\frac{C_\varepsilon}{\lambda}\|f\|_{L^{2}(\mathbf{G}_0^\#)},$$

which gives the bound (6.23). To complete the proof of Lemma 6.4 it suffices to show that

$$\biggl\|\sum_{j\geqslant J_{N_0,\mathcal{R}_{N_0}}}\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}}(f)\biggr\|_{L^{2p}(\mathbf{G}_0^\#)}\leqslant C_\varepsilon\lambda^\varepsilon\|f\|_{L^{2p}(\mathbf{G}_0^\#)}\tag{6.27}$$

for any characteristic function of a bounded set f.

For simplicity of notation, let

$$J_0=J_{N_0,\mathcal{R}_{N_0}}=\lambda^\varepsilon.$$

We use the notation in §6.2, with $\delta=\bar{c}\varepsilon/100$, D the smallest integer $\geqslant 2/\delta$, $N=N_0$, $N'=N_0'$ and

$$Q_0=[N_0'!]^D\leqslant e^{\lambda^{\varepsilon/10}}.\tag{6.28}$$

Then $Y_{N_0}=\bigcup_{s\in S}Y_{N_0}^s$ and $\mathcal{R}_{N_0}=\bigcup_{s\in S}\mathcal{R}_{N_0}^{W_s'}$ (disjoint unions), where $Y_{N_0}^s$ is defined in equation (6.17) and

$$\mathcal{R}_{N_0}^{W_s'}=\{a'/w'+b/Q_0:a',b\in\mathbf{Z}^{d'},(a',w')=1\text{ and }w'\in W_s'\}.\tag{6.29}$$

Clearly, for $j\geqslant J_0$,

$$\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}}(f)=\sum_{s\in S}\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}^{W_s'}}(f).$$

Since $|S|\leqslant C_\varepsilon(\log\lambda)^{C_\varepsilon}$ (see Lemma 6.2), for the bound (6.27), it suffices to prove that for any set $W'\subseteq W(N_0)\cup\{1\}$ with the orthogonality property O,

$$\biggl\|\sum_{j\geqslant J_0}\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}^{W'}}(f)\biggr\|_{L^{2p}(\mathbf{G}_0^\#)}\leqslant C_\varepsilon\lambda^{\varepsilon/2}\|f\|_{L^{2p}(\mathbf{G}_0^\#)}\tag{6.30}$$

for any characteristic function of a bounded set f.

We fix the set W' in inequality (6.30) and assume that $W'\neq\{1\}$ (the case $W'=\{1\}$ is significantly easier). Let $S_1, ..., S_k$, $S_j=\{q_{j,1}, ..., q_{j,\beta(j)}\}$, denote the sets in the definition of the orthogonality property O. Clearly $k\leqslant C_\varepsilon$ and

$$q_{j,s}\in[2,\lambda^{C_\varepsilon}]. \tag{6.31}$$

For any $s_1\in\mathbf{Z}_{\beta(1)}, ..., s_k\in\mathbf{Z}_{\beta(k)}$, let

$$\gamma(q_{1,s_1}\dots q_{k,s_k})=\begin{cases}1, & \text{if } q_{1,s_1}\dots q_{k,s_k}\in W',\\ 0, & \text{if } q_{1,s_1}\dots q_{k,s_k}\notin W'.\end{cases}$$

Any irreducible d'-fraction a'/w', $w'\in W'$, can be written uniquely in the form

$$\frac{a_{1,s_1}}{q_{1,s_1}}+\dots+\frac{a_{k,s_k}}{q_{k,s_k}} \pmod{\mathbf{Z}^{d'}},$$

with $q_{l,s_l}\in S_l$ and $a_{l,s_l}\in\widetilde{P}_{q_{l,s_l}}$, $l=1, ..., k$. Conversely, if $\gamma(q_{1,s_1}\dots q_{k,s_k})=1$, then any sum of the form $a_{1,s_1}/q_{1,s_1}+\dots+a_{k,s_k}/q_{k,s_k}$, with $q_{l,s_l}\in S_l$ and $a_{l,s_l}\in\widetilde{P}_{q_{l,s_l}}$, $l=1, ..., k$, belongs to the set $\{a'/w':(a',w')=1 \text{ and } w'\in W'\}$. Thus

$$\begin{aligned}\Psi_j^{N_0,\mathcal{R}_{N_0}^{W'}}(\theta)=\sum_{s_1,a_{1,s_1},...,s_k,a_{k,s_k}}\sum_{b\in\mathbf{Z}^{d'}}&\gamma(q_{1,s_1}\dots q_{k,s_k})\\ &\times\psi(2^{2j}(\theta-a_{1,s_1}/q_{1,s_1}-\dots-a_{k,s_k}/q_{k,s_k}-b/Q_0)/N_0),\end{aligned} \tag{6.32}$$

where the sum is taken over all $s_l\in\mathbf{Z}_{\beta(l)}$ and $a_{l,s_l}\in\widetilde{P}_{q_{l,s_l}}$. For any

$$r=a_{1,s_1}/q_{1,s_1}+\dots+a_{k,s_k}/q_{k,s_k},\quad s_l\in\mathbf{Z}_{\beta(l)} \text{ and } a_{l,s_l}\in P_{q_{l,s_l}}$$

(so $r\in T_{\mathbf{Z}_k}$ with the notation in §6.2), we define $G_r\in L^2(\mathbf{G}_0^\#)$ by the formula

$$\widehat{G}_r(m,\theta)=\gamma(q_{1,s_1}\dots q_{k,s_k})\sum_{j\geqslant J_0}\widehat{\mathcal{T}}_j(f)(m,\theta)\sum_{b\in\mathbf{Z}^{d'}}\psi(2^{2j}(\theta-r-b/Q_0)/N_0). \tag{6.33}$$

In view of equation (6.32),

$$\sum_{j\geqslant J_0}\mathcal{T}_j^{N_0,\mathcal{R}_{N_0}^{W'}}(f)=\sum_{r\in\widetilde{T}_{\mathbf{Z}_k}}G_r,$$

with $\widetilde{T}_{\mathbf{Z}_k}$ defined as in §6.2. Clearly, $\widehat{G}_r(m,\cdot\,)$ is supported in a $2N_0 2^{-2J_0}$-neighborhood of the set $\{r+b/Q_0:b\in\mathbf{Z}\}$. Condition (6.21) with $Q=Q_0$ and $\widetilde{Q}=\lambda^{C_\varepsilon}$ is satisfied if $\lambda\geqslant C_\varepsilon$

(see formulas (6.26), (6.28) and (6.31)). We apply Lemma 6.3 to the functions $G_r(m,u)$, for any $m\in\mathbf{Z}^d$. With the notation in Lemma 6.3, it follows that

$$\Big\|\sum_{r\in\widetilde{T}_{\mathbf{Z}_k}} G_r\Big\|^{2p}_{L^{2p}(\mathbf{G}_0^\#)} \leqslant C_\varepsilon \sum_{A=\{j_1,...,j_{k'}\}}\sum_{s_{j_1},...,s_{j_{k'}}}\int_{\mathbf{G}_0^\#}\bigg(\sum_{r'\in\widetilde{T}_{cA}}\bigg|\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}} G_{r'+\mu}(m,u)\bigg|^2\bigg)^p\,dm\,du.$$

The sum over the sets $A\subseteq\mathbf{Z}_k$ above has $2^k=C_\varepsilon$ terms. To summarize, for estimate (6.30), it suffices to prove that for any set $A=\{j_1,...,j_{k'}\}\subseteq\mathbf{Z}_k$,

$$\sum_{s_{j_1},...,s_{j_{k'}}}\int_{\mathbf{G}_0^\#}\bigg(\sum_{r'\in\widetilde{T}_{cA}}\bigg|\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}} G_{r'+\mu}(m,u)\bigg|^2\bigg)^p\,dm\,du\leqslant C_\varepsilon\lambda^{p\varepsilon}\|f\|^{2p}_{L^{2p}(\mathbf{G}_0^\#)} \tag{6.34}$$

for any characteristic function of a bounded set f.

For $(s_{j_1},...,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times...\times\mathbf{Z}_{\beta(j_{k'})}$ and $r'\in\widetilde{T}_{cA}$, let

$$\widetilde{G}_{r',s_{j_1}...,s_{j_{k'}}}=\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}} G_{r'+\mu}.$$

We also define the function $f_{r',s_{j_1}...,s_{j_{k'}}}\in L^2(\mathbf{G}_0^\#)$ by the formula

$$\mathcal{F}(f_{r',s_{j_1}...,s_{j_{k'}}})(m,\theta)=\hat{f}(m,\theta)\sum_{b\in\mathbf{Z}^{d'}}\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}}\psi(2^{2J_0-1}(\theta-r'-\mu-b/Q_0)/N_0). \tag{6.35}$$

For the bound (6.34) it suffices to prove that

$$\sum_{s_{j_1},...,s_{j_{k'}}}\int_{\mathbf{G}_0^\#}\bigg(\sum_{r'\in\widetilde{T}_{cA}}|f_{r',s_{j_1}...,s_{j_{k'}}}(m,u)|^2\bigg)^p\,dm\,du\leqslant C_\varepsilon\|f\|^{2p}_{L^{2p}(\mathbf{G}_0^\#)} \tag{6.36}$$

and

$$\int_{\mathbf{G}_0^\#}\bigg(\sum_{r'\in\widetilde{T}_{cA}}|\widetilde{G}_{r',s_{j_1}...,s_{j_{k'}}}(m,u)|^2\bigg)^p\,dm\,du \leqslant C_\varepsilon\lambda^{p\varepsilon}\int_{\mathbf{G}_0^\#}\bigg(\sum_{r'\in\widetilde{T}_{cA}}|f_{r',s_{j_1}...,s_{j_{k'}}}(m,u)|^2\bigg)^p\,dm\,du \tag{6.37}$$

for any $(s_{j_1},...,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times...\times\mathbf{Z}_{\beta(j_{k'})}$ fixed. The bound (6.36) follows from Lemma 6.5 below. The bound (6.37) follows from Lemma 6.6 below and the identity

$$\widetilde{G}_{r',s_{j_1}...,s_{j_{k'}}}=\gamma(q(r')q_{j_1,s_{j_1}}...\,q_{j_{k'},s_{j_{k'}}})\sum_{j\geqslant J_0}\mathcal{T}_j^{N_0,\mathcal{R}_{r',Q'}}(f_{r',s_{j_1}...,s_{j_{k'}}}), \tag{6.38}$$

where $Q'=Q_0 q_{j_1,s_{j_1}} \dots q_{j_{k'},s_{j_{k'}}}$ and $q(r')$ is the denominator of the irreducible d'-fraction r' (see the notation in Lemma 6.6). The identity (6.38) follows from the definitions and the observation

$$\sum_{b\in\mathbf{Z}^{d'}} \sum_{\mu\in\widetilde{U}_{A,s_{j_1},\dots,s_{j_{k'}}}} \psi(2^{2j}(\theta-r'-\mu-b/Q_0)/N_0)$$
$$=\sum_{b'\in\mathbf{Z}^{d'}} \psi(2^{2j}(\theta-r'-b'/Q')/N_0) \sum_{b\in\mathbf{Z}^{d'}} \sum_{\mu\in\widetilde{U}_{A,s_{j_1},\dots,s_{j_{k'}}}} \psi(2^{2J_0-1}(\theta-r'-\mu-b/Q_0)/N_0).$$

□

LEMMA 6.5. *With the notation above*,

$$\sum_{s_{j_1},\dots,s_{j_{k'}}} \int_{\mathbf{G}_0^\#} \biggl(\sum_{r'\in\widetilde{T}_{cA}} |f_{r',s_{j_1}\dots,s_{j_{k'}}}(m,u)|^2 \biggr)^p dm\,du \leqslant C_\varepsilon \|f\|_{L^{2p}(\mathbf{G}_0^\#)}^{2p}$$

for any characteristic function of a bounded set f.

Proof. This is similar to the proof of [8, Lemma 4.3], and is inspired by the Littlewood–Paley inequality in [13]. Clearly, since $f\colon \mathbf{G}_0^\# \to \{0,1\}$, $\|f\|_{L^{2p}(\mathbf{G}_0^\#)}^{2p} = \|f\|_{L^2(\mathbf{G}_0^\#)}^2$. In addition, by Plancherel's theorem,

$$\sum_{s_{j_1},\dots,s_{j_{k'}}} \int_{\mathbf{G}_0^\#} \sum_{r'\in\widetilde{T}_{cA}} |f_{r',s_{j_1}\dots,s_{j_{k'}}}(m,u)|^2\, dm\,du \leqslant C\|f\|_{L^2(\mathbf{G}_0^\#)}^2,$$

since the function $\mathcal{F}[f_{r',s_{j_1}\dots,s_{j_{k'}}}](m,\cdot)$ is supported in a $4N_0 2^{-2J_0}$-neighborhood of the set

$$r'+\sum_{b\in\mathbf{Z}^{d'}} \sum_{\mu\in\widetilde{U}_{A,s_{j_1},\dots,s_{j_{k'}}}} b/Q_0+\mu.$$

These neighborhoods are disjoint, as $r'\in\widetilde{T}_{cA}$ and $(s_{j_1},\dots,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times\dots\times\mathbf{Z}_{\beta(j_{k'})}$; see formulas (6.26), (6.28), and (6.31). Thus it suffices to prove that for any $(s_{j_1},\dots,s_{j_{k'}})\in \mathbf{Z}_{\beta(j_1)}\times\dots\times\mathbf{Z}_{\beta(j_{k'})}$ and $(m,u)\in\mathbf{Z}^d\times\mathbf{Z}^{d'}$,

$$\sum_{r'\in\widetilde{T}_{cA}} |f_{r',s_{j_1}\dots,s_{j_{k'}}}(m,u)|^2 \leqslant C_\varepsilon.$$

Thus, it suffices to prove that for $(s_{j_1},\dots,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times\dots\times\mathbf{Z}_{\beta(j_{k'})}$ fixed,

$$\biggl| \sum_{r'\in\widetilde{T}_{cA}} \nu(r') f_{r',s_{j_1}\dots,s_{j_{k'}}}(m,u) \biggr| \leqslant C_\varepsilon$$

for any $(m,u)\in\mathbf{Z}^d\times\mathbf{Z}^{d'}$ and any complex numbers $\nu(r')$ with

$$\sum_{r'\in\widetilde{T}_{c_A}}|\nu(r')|^2=1. \tag{6.39}$$

Since $\|f\|_{L^\infty}\leqslant 1$, it suffices to prove that for $(s_{j_1},...,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times...\times\mathbf{Z}_{\beta(j_{k'})}$ fixed,

$$\left\|\mathcal{F}^{-1}\left(\theta\mapsto\sum_{r'\in\widetilde{T}_{c_A}}\nu(r')\sum_{b\in\mathbf{Z}^{d'}}\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}}\psi(2^{2J_0-1}(\theta-r'-\mu-b/Q_0)/N_0)\right)\right\|_{L^1(\mathbf{Z}^{d'})}\leqslant C_\varepsilon. \tag{6.40}$$

As before, let $\eta(x)=\int_{\mathbf{R}^{d'}}\psi(\xi)e^{2\pi ix\cdot\xi}\,d\xi$ denote the Euclidean inverse Fourier transform of the function ψ. An easy calculation shows that

$$\begin{aligned}&\mathcal{F}^{-1}\left(\theta\mapsto\sum_{r'\in\widetilde{T}_{c_A}}\nu(r')\sum_{b\in\mathbf{Z}^{d'}}\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}}\psi(2^{2J_0-1}(\theta-r'-\mu-b/Q_0)/N_0)\right)(u)\\&\quad=\left(\sum_{r'\in\widetilde{T}_{c_A}}\nu(r')e^{2\pi iu\cdot r'}\right)\eta_{2^{2J_0-1}/N_0}(u)\left(\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}}\sum_{b\in\mathbf{Z}^{d'}_{Q_0}}e^{2\pi iu\cdot(b/Q_0+\mu)}\right).\end{aligned} \tag{6.41}$$

We first consider the sum over b and μ in equation (6.41). For any integer $Q'\geqslant 1$, define the function $\delta_{Q'}:\mathbf{Z}^{d'}\to\mathbf{Z}$ as in formula (4.58). Clearly, $\sum_{b\in\mathbf{Z}^{d'}_{Q'}}e^{2\pi iu\cdot b/Q'}=\delta_{Q'}(u)$. Recall from §6.2 that $q_{j,s_j}=p_{j,s_j}^{\alpha_j}$ for some primes $p_{j,s_j}\in V$ and $\alpha_j\in[1,C_\varepsilon]\cap\mathbf{Z}$. In addition, it is easy to see that if $q=p^\alpha$ and $(Q,p)=1$, then

$$\{a/q+b/Q: b\in\mathbf{Z}^{d'}\text{ and }a\in\widetilde{P}_q\}=\{b'/Qp^\alpha: b'\in\mathbf{Z}^{d'}\}\setminus\{b'/Qp^{\alpha-1}: b'\in\mathbf{Z}^{d'}\}.$$

Thus, for $(s_{j_1},...,s_{j_{k'}})\in\mathbf{Z}_{\beta(j_1)}\times...\times\mathbf{Z}_{\beta(j_{k'})}$ fixed,

$$\sum_{\mu\in\widetilde{U}_{A,s_{j_1},...,s_{j_{k'}}}}\sum_{b\in\mathbf{Z}^{d'}_{Q_0}}m(\mu+b/Q_0)=\sum_{\varepsilon_{j_1},...,\varepsilon_{j_{k'}}\in\{0,1\}}(-1)^{\varepsilon_{j_1}+...+\varepsilon_{j_{k'}}}\sum_{b\in\mathbf{Z}^{d'}_{Q'}}m(b/Q') \tag{6.42}$$

for any periodic function $m:\mathbf{R}^{d'}\to\mathbf{C}$, where

$$Q'=Q_0p_{j_1,s_{j_1}}^{\alpha_{j_1}-\varepsilon_{j_1}}\,...\,p_{j_{k'},s_{j_{k'}}}^{\alpha_{j_{k'}}-\varepsilon_{j_{k'}}}.$$

The possible values of Q' are products of Q_0 and $p_{j_l,s_{j_l}}^{\alpha_l}$ or $p_{j_l,s_{j_l}}^{\alpha_l-1}$, $l=1,...,k'$, and the sum over $\varepsilon_{j_1},...,\varepsilon_{j_{k'}}\in\{0,1\}$ contains $2^{k'}=C_\varepsilon$ terms. Thus, for estimate (6.40), it suffices to prove that

$$\left\|\left(\sum_{r'\in\widetilde{T}_{c_A}}\nu(r')e^{2\pi iu\cdot r'}\right)\eta_{2^{2J_0-1}/N_0}(u)\delta_{Q'}(u)\right\|_{L^1_u(\mathbf{Z}^{d'})}\leqslant C_\varepsilon$$

for any Q' with (see formulas (6.28) and (6.31))

$$Q' \in [1, e^{\lambda^{\varepsilon/5}}] \cap \mathbf{Z} \quad \text{and} \quad (Q', q_{j,s}) = 1 \text{ for any } j \in {}^{c}A,\ s \in \mathbf{Z}_{\beta(j)}. \tag{6.43}$$

This is equivalent to proving that

$$\left\| \left(\sum_{r' \in \widetilde{T}_{cA}} \nu(r') e^{2\pi i Q' u \cdot r'} \right) \eta_{2^{2J_0}/2N_0 Q'}(u) \right\|_{L^1_u(\mathbf{Z}^{d'})} \leqslant C_\varepsilon, \tag{6.44}$$

provided conditions (6.39) and (6.43) hold.

Let $\gamma_0 = 2^{-2J_0} 2N_0 Q' \ll 1$. The function η is a Schwartz function on $\mathbf{R}$; by Hölder's inequality, for estimate (6.44) it suffices to prove that

$$\gamma_0^{d'/2} \left\| \left(\sum_{r' \in \widetilde{T}_{cA}} \nu(r') e^{2\pi i Q' u \cdot r'} \right) (1 + \gamma_0^2 |u|^2)^{-d'} \right\|_{L^2_u(\mathbf{Z}^{d'})} \leqslant C_\varepsilon. \tag{6.45}$$

The left-hand side of inequality (6.45) is equal to

$$\gamma_0^{d'/2} \left(\sum_{r'_1, r'_2 \in \widetilde{T}_{cA}} \nu(r'_1) \overline{\nu(r'_2)} \int_{\mathbf{Z}^{d'}} (1 + \gamma_0^2 |u|^2)^{-2d'} e^{2\pi i u \cdot Q'(r'_1 - r'_2)}\, du \right)^{1/2}. \tag{6.46}$$

It remains to estimate the integrals over $\mathbf{Z}^{d'}$ in expression (6.46). If $r'_1 = r'_2$ then

$$\left| \int_{\mathbf{Z}^{d'}} (1 + \gamma_0^2 |u|^2)^{-2d'} e^{2\pi i u \cdot Q'(r'_1 - r'_2)}\, du \right| \leqslant C \gamma_0^{-d'}. \tag{6.47}$$

Recall that $d' = d^2$. If $r'_1 \neq r'_2$ then, by (6.43), $Q'(r'_1 - r'_2) \notin \mathbf{Z}^{d'}$. Let $\zeta = (\zeta_{l_1 l_2})_{l_1, l_2 = 1, \ldots, d}$ denote the fractional part of $Q'(r'_1 - r'_2)$. Since the denominators of r'_1 and r'_2 are bounded by λ^{C_ε}, there are $l_1, l_2 \in \{1, \ldots, d\}$ with the property that $\zeta_{l_1 l_2} \in [\lambda^{-C_\varepsilon}, 1 - \lambda^{-C_\varepsilon}]$. By summation by parts in the variable $u_{l_1 l_2}$ corresponding to this $\zeta_{l_1 l_2}$,

$$\left| \int_{\mathbf{Z}^{d'}} (1 + \gamma_0^2 |u|^2)^{-2d'} e^{2\pi i u \cdot Q'(r'_1 - r'_2)}\, du \right| \leqslant C \gamma_0^{-d'+1} \lambda^{C_\varepsilon} \tag{6.48}$$

if $r'_1 \neq r'_2$. We substitute inequalities (6.47) and (6.48) in expression (6.46). It follows that the left-hand side of inequality (6.45) is dominated by

$$C \left(\sum_{r' \in \widetilde{T}_{cA}} |\nu(r')|^2 + \gamma_0 \lambda^{C_\varepsilon} \left(\sum_{r' \in \widetilde{T}_{cA}} |\nu(r')| \right)^2 \right)^{1/2}.$$

Since $|\widetilde{T}_{cA}| \leqslant \lambda^{C_\varepsilon}$ and $\gamma_0 \leqslant e^{-\lambda^{\varepsilon/2}}$, the bound (6.45) follows from condition (6.39) and Hölder's inequality. This completes the proof of Lemma 6.5. □

LEMMA 6.6. *Assume, as before, that* $Q\in[1,e^{\lambda^{\varepsilon/5}}]\cap\mathbf{Z}$, $J_0=\lambda^{\varepsilon}$ *and* $N_0\leqslant\lambda^C$. *For any irreducible* d'*-fraction* $r=a/q$ *with* $q\in[1,\lambda^{C_\varepsilon}]\cap\mathbf{Z}$ *and* $(Q,q)=1$ *let*

$$\mathcal{R}_{r,Q}=\{r+b/Q:b\in\mathbf{Z}^{d'}\},$$

and, as in equation (6.4),

$$\mathcal{F}[\mathcal{T}_j^{N_0,\mathcal{R}_{r,Q}}(f)](m,\theta)=\widehat{\mathcal{T}}_j(f)\sum_{b\in\mathbf{Z}^{d'}}\psi(2^{2j}(\theta-r-b/Q)/N_0).$$

Then,

$$\biggl\|\biggl(\sum_r\biggl|\sum_{j\geqslant J_0}\mathcal{T}_j^{N_0,\mathcal{R}_{r,Q}}(f_r)\biggr|^2\biggr)^{1/2}\biggr\|_{L^{2p}(\mathbf{G}_0^\#)}\leqslant C_\varepsilon(\log\lambda)^C\biggl\|\biggl(\sum_r|f_r|^2\biggr)^{1/2}\biggr\|_{L^{2p}(\mathbf{G}_0^\#)}, \qquad (6.49)$$

for any (*compactly supported*) *function* $f_r:\mathbf{G}_0^\#\to\mathbf{C}$, *where the sums are taken over irreducible* d'*-fractions* $r=a/q$ *with* $q\in[1,\lambda^{C_\varepsilon}]\cap\mathbf{Z}$ *and* $(Q,q)=1$.

Proof. As in formula (4.26), in view of the definitions and the Fourier inversion formula,

$$\begin{aligned}\mathcal{T}_j^{N_0,\mathcal{R}_{r,Q}}(f_r)(m,u)=\sum_{(n,v)\in\mathbf{G}_0^\#}&f_r(n,v)K_j(m-n)\eta_{2^{2j}/N_0}(u-v-R_0(m-n,n))\\&\times e^{2\pi i(u-v-R_0(m-n,n))\cdot r}\delta_Q(u-v-R_0(m-n,n)),\end{aligned} \qquad (6.50)$$

where δ_Q is defined in (4.58). We use the change of variable

$$\Phi_Q:\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]\longrightarrow\mathbf{G}_0^\#$$

defined in (4.28). Let $F_r((n',v'),(\nu,\beta))=f_r(\Phi_Q((n',v'),(\nu,\beta)))$ and

$$G_r((m',u'),(\mu,\alpha))=\sum_{j\geqslant J_0}\mathcal{T}_j^{N_0,\mathcal{R}_{r,Q}}(f_r)(\Phi_Q((m',u'),(\mu,\alpha))).$$

Then, by formula (6.50),

$$\begin{aligned}&G_r((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^\#}\;\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F_r((n',v'),(\nu,\beta))\sum_{j\geqslant J_0}K_j(Q(m'-n')+E_1)\\&\qquad\times\eta_{2^{2j}/N_0}(Q^2(u'-v'-R_0(m'-n',n'))+E_2)\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu))e^{2\pi iE_3\cdot r},\end{aligned}$$

where $E_1=\mu-\nu$,

$$E_2=(\alpha-\beta-R_0(\mu-\nu,\nu))+Q(R_0(\mu,m'-n')-R_0(m'-n',\nu))$$

and

$$E_3=Q^2(u'-v'-R_0(m'-n',n'))+E_2.$$

Clearly, $|E_1|\leqslant CQ$ and $|E_2|\leqslant C2^jQ$ if $|m'-n'|\leqslant C2^j/Q$. Let

$$\begin{aligned}\widetilde{G}_r&((m',u'),(\mu,\alpha))\\ &=\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}F_r((n',v'),(\nu,\beta))\sum_{j\geqslant J_0}K_j(Q(m'-n'))\\ &\qquad\times\eta_{2^{2j}/N_0}(Q^2(u'-v'-R_0(m'-n',n')))\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu))e^{2\pi iE_3\cdot r}.\end{aligned}\tag{6.51}$$

In view of the estimates above on $|E_1|$ and $|E_2|$, and the relative sizes of Q, J_0 and N_0 (see the statement of Lemma 6.6),

$$\begin{aligned}&|G_r((m',u'),(\mu,\alpha))-\widetilde{G}_r((m',u'),(\mu,\alpha))|\\ &\quad\leqslant C\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}}|F_r((n',v'),(\nu,\beta))|Q^{-d}Q^{-2d'}\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu))\\ &\qquad\times\sum_{j\geqslant J_0}\frac{N_0Q}{2^j}\left(\frac{2^j}{Q}\right)^{-d}\mathbf{1}_{[0,C2^j/Q]}(|m'-n'|)\phi_{2^{2j}/Q^2N_0}(u'-v'-R_0(m'-n',n')),\end{aligned}$$

where ϕ is as in definition (7.7). The kernel in the formula defining

$$|G_r((m',u'),(\mu,\alpha))-\widetilde{G}_r((m',u'),(\mu,\alpha))|$$

has L^1 norm dominated by $CN_0Q/2^{J_0}\leqslant C$. In view of the Marcinkiewicz–Zygmund theorem,

$$\left\|\left(\sum_r|G_r-\widetilde{G}_r|^2\right)^{1/2}\right\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C\left\|\left(\sum_r|F_r|^2\right)^{1/2}\right\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}.$$

Thus, for estimate (6.49), it remains to prove that

$$\left\|\left(\sum_r|\widetilde{G}_r|^2\right)^{1/2}\right\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}\leqslant C_\varepsilon(\log\lambda)^C\left\|\left(\sum_r|F_r|^2\right)^{1/2}\right\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])},\tag{6.52}$$

where $\widetilde{G}_r$ is defined in (6.51).

Assume that $r=a/q$, $(a,q)=1$, and for any (ν,β) fixed define

$$\begin{aligned} &H_r((m',u'),(\nu,\beta)) \\ &\quad = \sup_{\substack{a_1\in\mathbf{Z}^d\\ a_2\in\mathbf{Z}^{d'}}} \bigg| \sum_{(n',v')\in\mathbf{G}_0^\#} F_r((n',v'),(\nu,\beta)) \sum_{j\geqslant J_0} K_j(Q(m'-n')) \\ &\qquad\qquad \times Q^d \eta_{2^{2j}/Q^2N_0}(u'-v'-R_0(m'-n',n'))e^{2\pi i[a_1\cdot(m'-n')+a_2\cdot(u'-v'-R_0(m'-n',n'))]/q}\bigg|. \end{aligned}$$

Clearly,

$$|\widetilde{G}_r((m',u'),(\mu,\alpha))| \leqslant \sum_{(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}} H_r((m',u'),(\nu,\beta))Q^{-d}Q^{-2d'}\delta_Q(\alpha-\beta-R_0(\mu-\nu,\nu)),$$

so, using the Marcinkiewicz–Zygmund theorem again,

$$\bigg\| \bigg(\sum_r |\widetilde{G}_r|^2\bigg)^{1/2}\bigg\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])} \leqslant C \bigg\| \bigg(\sum_r |H_r|^2\bigg)^{1/2}\bigg\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}. \tag{6.53}$$

Thus, for estimate (6.52), it suffices to prove that

$$\bigg\| \bigg(\sum_r |H_r|^2\bigg)^{1/2}\bigg\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])} \leqslant C_\varepsilon(\log\lambda)^C \bigg\| \bigg(\sum_r |F_r|^2\bigg)^{1/2}\bigg\|_{L^{2p}(\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}])}. \tag{6.54}$$

To prove estimate (6.54), we use Lemma 7.4. The connection between weighted estimates and vector-valued inequalities is well-known (see, for example, [7, Chapter V, Theorem 6.1]). In our case, let $p'\in(1,\infty]$ denote the exponent dual to p. The left-hand side of inequality (6.54) is dominated by

$$\sup_{\substack{w:\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]\to[0,\infty)\\ \|w\|_{L^{p'}}=1}} \bigg(\int_{\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]} \sum_r |H_r|^2 w\bigg)^{1/2}. \tag{6.55}$$

We examine the definition of the functions H_r above and notice that for fixed $(\nu,\beta)\in \mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}$,

$$H_r(h,(\nu,\beta)) \leqslant \widetilde{\mathcal{T}}_*^{\widetilde{N}_0,q}[F_r(\cdot,(\nu,\beta))](h), \quad h\in\mathbf{G}_0^\#,$$

with the notation in Lemma 7.3. The operators $\widetilde{\mathcal{T}}_*^{\widetilde{N}_0,q}$ are as in the statement of Lemma 7.3, using the kernels $\widetilde{K}_j(x)=Q^dK_{j+j_1}(Qx)$, $j\geqslant\lambda^\varepsilon/2$, where j_1 is the smallest integer such that $2^{j_1}\geqslant Q$, and $\widetilde{N}_0=Q^2N_0/2^{2j_1}$. These kernels $\widetilde{K}_j$ clearly satisfy the basic properties (6.1), (6.2) and condition (7.23). For fixed $(\nu,\beta)\in\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}$ we define the function $w_*^{\widetilde{N}_0}(\cdot,(\nu,\beta))$ as in equation (7.21), and use the bounds (7.22) and Lemma 7.4 with $\varrho=C_\varepsilon\log(\widetilde{N}_0+1)$. The expression (6.55) is dominated by

$$\sup_{\substack{w:\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]\to[0,\infty)\\ \|w\|_{L^{p'}}=1}} C_\varepsilon(\log\lambda)^C\biggl(\int_{\mathbf{G}_0^\#\times[\mathbf{Z}_Q^d\times\mathbf{Z}_{Q^2}^{d'}]}\sum_r |F_r|^2 w_*^{\widetilde{N}_0}\biggr)^{1/2},$$

which easily leads to estimate (6.54) (using again the bounds (7.22)). □

7. Real-variable theory on the group $\mathbf{G}_0^\#$

In this section, which is self-contained, we discuss some features of the real-variable theory on the group $\mathbf{G}_0^\#$. Our basic reference is [14, Chapters I, II and V]. The main results in this section are the bound (7.11), which is used in §4.3, and Lemma 7.4, which is used in §6.3. We assume throughout this section that $d'=d^2$ and $\mathbf{G}_0^\#$ is the discrete nilpotent group defined in §2.

7.1. Weighted maximal functions

We define the "distance" function $d:\mathbf{G}_0^\#\times\mathbf{G}_0^\#\to[0,\infty)$,

$$d(0,(m,u))=\max(|m|,|u|^{1/2}),\quad d(h,h')=d(0,h'\cdot h^{-1})\ \text{ if }\ h,h'\in\mathbf{G}_0^\#. \tag{7.1}$$

It is easy to see that $d(h,h')\approx d(h',h)$ and

$$d(h,h'')\leqslant C(d(h,h')+d(h',h''))\quad\text{for any }\ h,h',h''\in\mathbf{G}_0^\#.$$

We define the family of non-isotropic balls on $\mathbf{G}_0^\#$:

$$\mathbf{B}=\bigl\{B(h,r)=\{g\cdot h:d(0,g)\leqslant r\}:h\in\mathbf{G}_0^\#\text{ and }r\geqslant\tfrac12\bigr\}, \tag{7.2}$$

and notice that we have the basic properties

$$\begin{gathered}\text{if }\ B(h,r)\cap B(h',r)\neq\varnothing\ \text{ then }\ B(h',r)\subseteq B(h,C_1r),\\ |B(h,C_1r)|\leqslant C_2|B(h,r)|,\end{gathered} \tag{7.3}$$

for any $h,h'\in\mathbf{G}_0^\#$ and $r\geqslant\frac12$. As a consequence, we have the Whitney decomposition (see [14, p. 15]): if $O\subseteq\mathbf{G}_0^\#$ is a finite set, then there are balls $B_k\in\mathbf{B}$, $k=1,...,K$, with the properties

$$B_k\cap B_{k'}=\varnothing\ \text{ for any }\ k\neq k',\quad O=\bigcup_{k=1}^K B_k^*\quad\text{and}\quad B_k^{**}\cap{}^cO\neq\varnothing, \tag{7.4}$$

where, if $B=B(h,r)$, then $B^*=B(h,c^*r)$ and $B^{**}=B(h,(c^*)^2r)$ for a sufficiently large constant c^*. In addition, there are pairwise disjoint Whitney "cubes" Q_k with the properties $\bigcup_{k=1}^K Q_k=O$ and $B_k\subseteq Q_k\subseteq B_k^*$.

For any set $E\subseteq\mathbf{G}_0^\#$ and any function $w:\mathbf{G}_0^\#\to[0,\infty)$ let $w(E)=\int_E w(h)\,dh$. If $w:\mathbf{G}_0^\#\to[0,\infty)$ is a non-negative function, we define $L^p(w)$, $p\in[1,\infty]$, and $L^{1,\infty}(w)$ as the corresponding weighted spaces on $\mathbf{G}_0^\#$. It follows from properties (7.3) that the standard non-centered maximal function

$$\widetilde{\mathcal{M}}(f)(h)=\sup_{h\in B\in\mathbf{B}}\frac{1}{|B|}\int_B|f(g)|\,dg \tag{7.5}$$

extends to a bounded operator from $L^1(w)$ to $L^{1,\infty}(\widetilde{\mathcal{M}}(w))$:

$$\alpha w(\{h:\widetilde{\mathcal{M}}(f)(h)>\alpha\})\leqslant C\int_{\mathbf{G}_0^\#}|f(h)|\widetilde{\mathcal{M}}(w)(h)\,dh \tag{7.6}$$

for any $f:\mathbf{G}_0^\#\to\mathbf{C}$ and $\alpha\in(0,\infty)$ (see [14, p. 53]).

Let Ω and Ω_j be defined as in formula (4.1). In this section we assume, in addition, that $\Omega(x)=1$ if $|x|\leqslant 2$. Let $\phi,\phi_r:\mathbf{R}^{d'}\to[0,1]$ denote the functions

$$\phi(s)=(1+|s|^2)^{-(d'+d+1)/2}\quad\text{and}\quad\phi_r(s)=r^{-d'}\phi(s/r),\ \ r\geqslant 1. \tag{7.7}$$

Assume that $N\geqslant 1$ is a real number. For integers $j\geqslant\log_2 N$ we define the kernels

$$A_j^N,A_j'^N:\mathbf{G}_0^\#\longrightarrow[0,\infty),$$

by

$$A_j^N(m,u)=\Omega_j(m)\phi_{2^{2j}/N}(u)\quad\text{and}\quad A_j'^N(g)=A_j^N(g^{-1}),\ \ g\in\mathbf{G}_0^\#.$$

For $N\geqslant 1$ and $f:\mathbf{G}_0^\#\to\mathbf{C}$, let

$$\mathcal{M}_*^N(f)(h)=\sup_{j\geqslant\log_2 N}|f*(A_j^N+A_j'^N)(h)|+\sup_{j\geqslant 0}|f*(A_j^1+A_j'^1)(h)|. \tag{7.8}$$

We start with a weighted maximal inequality.

Lemma 7.1. *Assume that $N,\varrho\in[1,\infty)$ and that $w:\mathbf{G}_0^\#\to(0,\infty)$ is a function with the property that*

$$\mathcal{M}_*^N(w)(h)\leqslant\varrho w(h)\quad\textit{for any }h\in\mathbf{G}_0^\#. \tag{7.9}$$

Then, for any compactly supported function $f:\mathbf{G}_0^\#\to\mathbf{C}$,

$$\begin{gathered}\|\mathcal{M}_*^N(f)\|_{L^{1,\infty}(w)}\leqslant C\varrho^2\log(N+1)\|f\|_{L^1(w)},\\ \|\mathcal{M}_*^N(f)\|_{L^p(w)}\leqslant C_p\varrho^2\log(N+1)\|f\|_{L^p(w)},\quad p\in(1,\infty].\end{gathered} \tag{7.10}$$

In particular, if $w\equiv 1$, then

$$\begin{gathered}\|\mathcal{M}_*^N(f)\|_{L^{1,\infty}(\mathbf{G}_0^\#)}\leqslant C\log(N+1)\|f\|_{L^1(\mathbf{G}_0^\#)},\\ \|\mathcal{M}_*^N(f)\|_{L^p(\mathbf{G}_0^\#)}\leqslant C_p\log(N+1)\|f\|_{L^p(\mathbf{G}_0^\#)},\quad p\in(1,\infty].\end{gathered} \tag{7.11}$$

Proof. The main issue is to prove that there is only a logarithmic loss in N in estimates (7.10) and (7.11). Since the non-centered maximal operator $\widetilde{\mathcal{M}}$ in equation (7.5) is dominated by $C\mathcal{M}^1_*$, it follows from property (7.9) that

$$\frac{w(B)}{|B|} \leqslant C\varrho \min_{h\in B} w(h) \quad \text{for any ball } B\in \mathbf{B}. \tag{7.12}$$

We recall the Calderón–Zygmund decomposition of functions on $\mathbf{G}_0^\#$: if $f\in L^1(\mathbf{G}_0^\#)$ and $\alpha\in(0,\infty)$ is a given "height", let $E_\alpha=\{h:\widetilde{\mathcal{M}}(f)(h)>\alpha\}$ and $E_\alpha=\bigcup_{k=1}^K B_k^*=\bigcup_{k=1}^K Q_k$ be the Whitney decomposition of the set E_α (see properties (7.4)). Let

$$f_0(h)=\mathbf{1}_{^cE_\alpha}(h)f(h)+\sum_{k=1}^K \mathbf{1}_{Q_k}(h)\frac{1}{|Q_k|}\int_{Q_k} f(h')\,dh',$$
$$b_k(h)=\mathbf{1}_{Q_k}(h)\biggl(f(h)-\frac{1}{|Q_k|}\int_{Q_k} f(h')\,dh'\biggr).$$

Clearly, $f=f_0+\sum_{k=1}^K b_k$; in addition, directly from the definitions,

$$\begin{gathered} |f_0(h)|\leqslant C\alpha \quad \text{ for any } h\in \mathbf{G}_0^\#, \\ b_k \text{ is supported in } Q_k \quad\text{and}\quad \int_{\mathbf{G}_0^\#} b_k(h)\,dh=0. \end{gathered} \tag{7.13}$$

Also, using property (7.12) for the balls B_k^* and the definition of b_k,

$$\int_{\mathbf{G}_0^\#} |b_k(h)|w(h)\,dh \leqslant C\varrho \|f\mathbf{1}_{Q_k}\|_{L^1(w)}. \tag{7.14}$$

By interpolation, we only need to prove the $L^1(w)\to L^{1,\infty}(w)$ bound in (7.10). Assume that $f\colon \mathbf{G}_0^\#\to\mathbf{C}$ is a compactly supported function and fix $\alpha\in(0,\infty)$. It suffices to prove that

$$\alpha w(\{h:\mathcal{M}_*^N(f)(h)>\alpha\})\leqslant C\varrho^2\log(N+1)\|f\|_{L^1(w)}.$$

We use the Calderón–Zygmund decomposition $f=f_0+\sum_{k=1}^K b_k=f_0+b$ at height α/C, for C sufficiently large. It suffices to prove that

$$\alpha w\bigl(\bigl\{h:\mathcal{M}_*^N(b)(h)>\tfrac{1}{2}\alpha\bigr\}\bigr)\leqslant C\varrho^2\log(N+1)\|f\|_{L^1(w)}. \tag{7.15}$$

For estimate (7.15), it suffices to prove that

$$\alpha\sum_{k=1}^K w(B_k^{**})\leqslant C\varrho^2\|f\|_{L^1(w)} \tag{7.16}$$

and

$$\sum_{k=1}^{K}\int_{^cB_k^{**}}\mathcal{M}_*^N(b_k)(h)w(h)\,dh\leqslant C\varrho^2\log(N+1)\|f\|_{L^1(w)},\tag{7.17}$$

where the B_k^{**}'s are sufficiently large dilations of the balls B_k that appear in the Whitney decomposition of the set $E_{\alpha/C}$.

To prove the bound (7.16), we use estimates (7.12) and (7.6):

$$\begin{aligned}\alpha\sum_{k=1}^{K}w(B_k^{**})&\leqslant C\varrho\alpha\sum_{k=1}^{K}|B_k^{**}|\min_{h\in B_k^{**}}w(h)\leqslant C\varrho\alpha w(\{h:\widetilde{\mathcal{M}}(f)(h)>\alpha/C\})\\&\leqslant C\varrho\int_{\mathbf{G}_0^\#}|f(h)|\widetilde{\mathcal{M}}(w)(h)\,dh\leqslant C\varrho^2\|f\|_{L^1(w)},\end{aligned}$$

as desired.

To prove the bound (7.17), we use estimate (7.14) and the fact that the cubes Q_k are pairwise disjoint. By translation invariance, it suffices to prove that if $B=B(0,r)$ is a ball centered at 0 and $f:\mathbf{G}_0^\#\to\mathbf{C}$ is a function supported in the ball B with the property that $\int_{\mathbf{G}_0^\#}f(g)\,dg=0$, then

$$\begin{aligned}\sum_{j\geqslant\log_2 N}\int_{^cB^*}|f*(A_j^N+A_j'^N)(h)|w(h)\,dh+\sum_{j\geqslant 0}\int_{^cB^*}|f*(A_j^1+A_j'^1)(h)|w(h)\,dh&\\ \leqslant C\varrho\log(N+1)\|f\|_{L^1(w)},&\end{aligned}\tag{7.18}$$

where, as before, $B^*=B(0,c^*r)$, for c^* sufficiently large. To prove estimate (7.18), it suffices to control the first sum in the left-hand side (the second sum corresponds to the particular case $N=1$). Since $r\geqslant\frac12$, fix $k_0\in\mathbf{Z}\cap[-1,\infty)$ such that $2^{k_0}\leqslant r<2^{k_0+1}$. We divide the sum over j into three parts: $j\leqslant k_0$, $j\in[k_0,k_0+2\log(N+1)]$ and $j\geqslant k_0+2\log(N+1)$.

For $\log_2 N\leqslant j\leqslant k_0$, ignoring the condition $\int_{\mathbf{G}_0^\#}f(g)\,dg=0$, we notice that if $h\in{}^cB^*$, $g\in B$ and c^* is sufficiently large, then $\min(d(0,h\cdot g^{-1}),d(0,g\cdot h^{-1}))\geqslant\frac12c^*2^{k_0}$. From the definitions,

$$(A_j^N+A_j'^N)(hg^{-1})\leqslant C2^{j-k_0}(A_{k_0+2}^N+A_{k_0+2}'^N)(gh^{-1}).$$

Thus, using property (7.9),

$$\begin{aligned}\int_{^cB^*}|f*(A_j^N+A_j'^N)(h)|w(h)\,dh&\leqslant C2^{j-k_0}\int_B|f(g)|(w*(A_{k_0+2}^N+A_{k_0+2}'^N)(g))\,dg\\&\leqslant C\varrho2^{j-k_0}\|f\|_{L^1(w)},\end{aligned}\tag{7.19}$$

which suffices to prove estimate (7.18) for this part of the sum.

For $j\geqslant\log_2 N$ and $j\in[k_0,k_0+2\log(N+1)]$, we use property (7.9) as before, and notice that the sum contains at most $C\log(N+1)$ terms.

For $j\geqslant k_0+2\log(N+1)$, we use the condition $\int_{\mathbf{G}_0^\#} f(g)\,dg=0$ and write

$$|f*(A_j^N+A_j'^N)(h)|\leqslant\int_B|f(g)|\,|(A_j^N+A_j'^N)(hg^{-1})-(A_j^N+A_j'^N)(h)|\,dg.$$

Assume that $h=(n,v)\in{}^cB^*$ and $g=(m,u)\in B$. Then $hg^{-1}=(n-m,v-u-R_0(n-m,m))$ and

$$\begin{aligned}|A_j^N(hg^{-1})-A_j^N(h)|&\leqslant|\Omega_j(n-m)-\Omega_j(n)|\phi_{2^{2j}/N}(v)\\&\qquad+\Omega_j(n-m)|\phi_{2^{2j}/N}(v-u-R_0(n-m,m))-\phi_{2^{2j}/N}(v)|\\&\leqslant C(N+1)2^{k_0-j}(2^{-dj}\mathbf{1}_{[0,2^{j+3}]}(n))\phi_{2^{2j}/N}(v)\\&\leqslant C(N+1)2^{k_0-j}A_{j+3}^N(hg^{-1}).\end{aligned}\tag{7.20}$$

Similar estimates show that

$$|A_j'^N(hg^{-1})-A_j'^N(h)|\leqslant C(N+1)2^{k_0-j}A_{j+3}'^N(hg^{-1}).$$

Estimate (7.18) for this part of the sum follows using property (7.9), as in formula (7.19). This completes the proof of Lemma 7.1. □

We now explain how to construct weights with property (7.9). Assume that $p\in(1,\infty]$, $w\colon\mathbf{G}_0^\#\to[0,\infty)$ and $w\in L^p(\mathbf{G}_0^*)$. For $N\geqslant1$, let

$$w_*^N=\sum_{k=0}^{\infty}(C_p\log(N+1))^{-k}(\mathcal{M}_*^N)^k(w),\tag{7.21}$$

where C_p is a sufficiently large constant. Then, using estimates (7.11),

$$\begin{aligned}w(h)&\leqslant w_*^N(h)&&\text{for any }h\in\mathbf{G}_0^\#,\\\|w_*^N\|_{L^p(\mathbf{G}_0^\#)}&\leqslant C\|w\|_{L^p(\mathbf{G}_0^\#)},\\\mathcal{M}_*^N(w_*^N)(h)&\leqslant C_p\log(N+1)w_*^N(h)&&\text{for any }h\in\mathbf{G}_0^\#.\end{aligned}\tag{7.22}$$

In particular, property (7.9) holds for the function w_*^N with $\varrho=C_p\log(N+1)$. We used this construction in the proof of Lemma 6.6 in §6.3.

7.2. Maximal oscillatory singular integrals

We now consider singular integrals on the group $\mathbf{G}_0^\#$. The main result in this subsection is Lemma 7.4. Let $K_j\colon\mathbf{R}^d\to\mathbf{C}$, $j=0,1,\dots$, denote a family of kernels on $\mathbf{R}^d$ with the properties (6.1) and (6.2). In this section it is convenient to make a slightly less restrictive assumption on the supports of the K_j's, namely

$$K_j \text{ is supported in the set } \{x : |x| \in [c_0 2^{j-1}, c_0 2^{j+1}]\} \quad \text{for some } c_0 \in \left[\tfrac{1}{2}, 2\right]. \tag{7.23}$$

Assume that $\eta\in\mathcal{S}(\mathbf{R}^{d'})$ is a fixed Schwartz function and let

$$\eta_r(s) = r^{-d'}\eta(s/r), \quad s \in \mathbf{R}^{d'},\ r \geqslant 1.$$

Let $N\geqslant 1$ be a real number. For integers $j\geqslant \log_2 N$, we define the kernels $L_j^N: \mathbf{G}_0^\# \to \mathbf{C}$,

$$L_j^N(m,u) = K_j(m)\eta_{2^{2j}/N}(u).$$

For (compactly supported) functions $f: \mathbf{G}_0^\# \to \mathbf{C}$ let

$$\mathcal{T}_j^N(f) = f * L_j^N \quad \text{and} \quad \mathcal{T}_{\geqslant j}^N(f) = \sum_{j'=j}^{\infty} \mathcal{T}_{j'}^N(f).$$

LEMMA 7.2. (Maximal singular integrals) *Assume that* $N\in[1,\infty)$. *The maximal singular integral operator*

$$\mathcal{T}_*^N(f)(h) = \sup_{j\geqslant \log N} |\mathcal{T}_{\geqslant j}^N(f)(h)|$$

extends to a bounded (subadditive) operator on $L^p(\mathbf{G}_0^\#)$, $p\in(1,\infty)$, *with*

$$\|\mathcal{T}_*^N\|_{L^p\to L^p} \leqslant C_p(\log(N+1))^2. \tag{7.24}$$

Proof. As in Lemma 7.1, the main issue is to prove that there is only a logarithmic loss in N in inequality (7.24). We first show that

$$\Bigl\| \sum_{j\geqslant \log N} \mathcal{T}_j^N \Bigr\|_{L^2\to L^2} \leqslant C\log(N+1). \tag{7.25}$$

In the proof of estimate (7.25) we assume that the kernels K_j satisfy the slightly different cancellation condition $\sum_{m\in\mathbf{Z}^d} K_j(m)=0$ instead of condition (6.2). The two cancellation conditions are equivalent (at least in the proof of the bound (7.25)) by replacing K_j with $K_j - C_j 2^{-j}\varphi_j$ for suitable constants $|C_j|\leqslant C$, where $\varphi: \mathbf{R}^d\to[0,1]$ is a smooth function supported in $\left\{x:|x|\in\left[\frac{1}{2},2\right]\right\}$ and $\varphi_j(x)=(c_0 2^j)^{-d}\varphi(x/c_0 2^j)$. By abuse of notation, in the proof of estimate (7.25) we continue to denote by $\mathcal{T}_j^N$, L_j^N, etc. the operators and the kernels corresponding to these modified kernels K_j. Clearly, $\|\mathcal{T}_j^N\|_{L^2\to L^2}\leqslant C$ for any $j\geqslant \log_2 N$. By the Cotlar–Stein lemma, it suffices to prove that

$$\|\mathcal{T}_j^N[\mathcal{T}_k^N]^*\|_{L^2\to L^2} + \|\,[\mathcal{T}_j^N]^*\mathcal{T}_k^N\|_{L^2\to L^2} \leqslant C(N+1)2^{-|j-k|} \tag{7.26}$$

for any $j,k\geqslant\log_2 N$ with $|j-k|\geqslant 2\log(N+1)$. Assume that $j\geqslant k$. The kernel of the operator $\mathcal{T}_j^N[\mathcal{T}_k^N]^*$ is

$$L_{j,k}^N(g)=\int_{\mathbf{G}_0^\#}\bar{L}_k^N(h)L_j^N(gh)\,dh.$$

Using the cancellation condition (6.2), with $h=(n,v)$,

$$\begin{aligned}|L_{j,k}^N(g)| &\leqslant \int_{|v|\leqslant 2^{j+k}}|L_k^N(h)|\,|L_j^N(gh)-L_j^N(g)|\,dh+\int_{|v|\geqslant 2^{j+k}}|L_k^N(h)|\,|L_j^N(gh)|\,dh\\ &= I_1(g)+I_2(g).\end{aligned}$$

An estimate similar to (7.20) shows that

$$I_1(m,u)\leqslant C(N+1)2^{-|j-k|}(2^{-dj}\mathbf{1}_{[0,2^{j+3}]}(m))\phi_{2^{2j}/N}(u).$$

Also, by integrating the variable g first, it is easy to see that $\|I_2\|_{L^1(\mathbf{G}_0^\#)}\leqslant C(N+1)2^{-|j-k|}$. The bound for the first term in the left-hand side of (7.26) follows. The bound for the second term is similar, which completes the proof of estimate (7.25).

The proof of estimate (7.20) shows that

$$\sum_{j\geqslant\log_2 N}\int_{{}^cB(0,c^*r)}|L_j^N(hg^{-1})-L_j^N(h)|\,dh\leqslant C\log(N+1)$$

for any $r>0$ and $g\in B(0,r)$. Let $\mathcal{T}^N(f)=\sum_{j\geqslant\log N}\mathcal{T}_j^N$. It follows from estimate (7.25) and standard Calderón–Zygmund theory that

$$\|\mathcal{T}^N\|_{L^1\to L^{1,\infty}}\leqslant C\log(N+1)\quad\text{and}\quad\|\mathcal{T}^N\|_{L^p\to L^p}\leqslant C_p\log(N+1),\;\; p\in(1,\infty).\tag{7.27}$$

We turn now to the proof of the bound (7.24). In view of estimates (7.11) and (7.27), it suffices to prove the pointwise bound

$$\mathcal{T}_*^N(f)(h)\leqslant C\log(N+1)[\widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|))(h)+\widetilde{\mathcal{M}}(|\mathcal{T}^N(f)|)(h)]\tag{7.28}$$

for any $h\in\mathbf{G}_0^\#$, where $\widetilde{\mathcal{M}}$ is the non-centered maximal operator defined in (7.5). By translation invariance, it suffices to prove this bound for $h=0$. Thus, it suffices to prove that for any $k_0\geqslant\log_2 N$,

$$\left|\sum_{j\geqslant k_0}\mathcal{T}_j^N(f)(0)\right|\leqslant C\log(N+1)[\widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|))(0)+\widetilde{\mathcal{M}}(|\mathcal{T}^N(f)|)(0)].\tag{7.29}$$

Assume k_0 fixed and let $f_1=f\mathbf{1}_{B(0,2^{k_0-2})}$ and $f_2=f-f_1$. It follows from the definitions that $\sum_{j\geqslant k_0}\mathcal{T}_j^N(f)(0)=\sum_{j\geqslant k_0}\mathcal{T}_j^N(f_2)(0)$.

We first show that for any $h\in B(0,c2^{k_0})$, for c sufficiently small,

$$\Big|\sum_{j\geqslant k_0}\mathcal{T}_j^N(f_2)(0)-\mathcal{T}^N(f_2)(h)\Big|\leqslant C\log(N+1)[\mathcal{M}_*^N(|f|)(0)+\mathcal{M}_*^N(|f|)(h)]. \tag{7.30}$$

To prove the bound (7.30), we first notice that

$$\Big|\sum_{j\in[k_0,k_0+2\log(N+1))}\mathcal{T}_j^N(f_2)(0)-\sum_{j\in[k_0,k_0+2\log(N+1))}\mathcal{T}_j^N(f_2)(h)\Big|$$

is clearly controlled by the right-hand side of inequality (7.30). In addition,

$$\begin{aligned}\Big|\sum_{j\geqslant k_0+2\log(N+1)}\mathcal{T}_j^N(f_2)(0)-&\sum_{j\geqslant k_0+2\log(N+1)}\mathcal{T}_j^N(f_2)(h)\Big|\\ &\leqslant\sum_{j\geqslant k_0+2\log(N+1)}\int_{\mathbf{G}_0^\#}|f_2(g^{-1})|\,|L_j^N(g)-L_j^N(hg)|\,dg\\ &\leqslant C\sum_{j\geqslant k_0+2\log(N+1)}(N+1)2^{k_0-j}\mathcal{M}_*^N(|f_2|)(0),\end{aligned}$$

using an estimate on the difference $|L_j^N(g)-L_j^N(hg)|$ similar to (7.20). Finally,

$$\begin{aligned}\Big|\sum_{j\in[\log_2 N,k_0)}\mathcal{T}_j^N(f_2)(h)\Big|&\leqslant\int_{\mathbf{G}_0^\#}|f_2(g^{-1})|\bigg(\sum_{j\in[\log_2 N,k_0]}|L_j^N(hg)|\bigg)\,dg\\ &\leqslant C\int_{\mathbf{G}_0^\#}|f_2(g^{-1})|A_{k_0}^N(hg)\,dg\\ &\leqslant C\mathcal{M}_*^N(|f_2|)(h).\end{aligned}$$

The bound (7.30) follows. Thus, for any $h\in B(0,c2^{k_0})$,

$$\Big|\sum_{j\geqslant k_0}\mathcal{T}_j^N(f)(0)\Big|\leqslant C\log(N+1)[\mathcal{M}_*^N(|f|)(0)+\mathcal{M}_*^N(|f|)(h)]+|\mathcal{T}^N(f)(h)|+|\mathcal{T}^N(f_1)(h)|.$$

The proof of the bound (7.29) now follows easily as in [14, Chapter I, §7.3], using estimates (7.11) and (7.27). This completes the proof of the lemma. □

In the proof of Lemma 6.6 we need bounds on more general oscillatory singular integral operators. Assume that $q\geqslant1$ is an integer, $N\geqslant1$ is a real number as before, $a_1\in\mathbf{Z}^d$ and $a_2\in\mathbf{Z}^{d'}$. For integers $j\geqslant\log_2(2Nq)$ and K_j satisfying properties (6.1), (6.2) and (7.23), we define the kernels $L_{j,a_1,a_2}^{N,q}:\mathbf{G}_0^\#\to\mathbf{C}$ by

$$L_{j,a_1,a_2}^{N,q}(m,u)=K_j(m)\eta_{2^{2j}/N}(u)e^{2\pi i(a_1\cdot m+a_2\cdot u)/q}.$$

For (compactly supported) functions $f:\mathbf{G}_0^\#\to\mathbf{C}$ let

$$\mathcal{T}_{j,a_1,a_2}^{N,q}(f)=f*L_{j,a_1,a_2}^{N,q}\quad\text{and}\quad\mathcal{T}_{\geqslant j,a_1,a_2}^{N,q}(f)=\sum_{j'=j}^{\infty}\mathcal{T}_{j',a_1,a_2}^{N,q}(f).$$

LEMMA 7.3. (Maximal oscillatory singular integrals) *Assume that* $N\in[1,\infty)$. *The maximal oscillatory singular integral operator*

$$\mathcal{T}_*^{N,q}(f)(h)=\sup_{\substack{a_1\in\mathbf{Z}^d\\ a_2\in\mathbf{Z}^{d'}}}\sup_{j\geqslant\log_2(2Nq)}|\mathcal{T}^{N,q}_{\geqslant j,a_1,a_2}(f)(h)|$$

extends to a bounded (subadditive) operator on $L^p(\mathbf{G}_0^\#)$, $p\in(1,\infty)$, *with*

$$\|\mathcal{T}_*^{N,q}\|_{L^p\to L^p}\leqslant C_p(\log(N+1))^2. \tag{7.31}$$

Proof. Notice first that the case $q=1$ follows from Lemma 7.2, since $L^{N,1}_{j,a_1,a_2}=L^N_j$. To deal with the case $q\geqslant2$, we use the coordinates (4.28) on $\mathbf{G}_0^\#$ adapted to the factor q:

$$\Phi_q\colon \mathbf{G}_0^\#\times[\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}]\longrightarrow\mathbf{G}_0^\#,$$
$$\Phi_q((m',u'),(\mu,\alpha))=(qm'+\mu,q^2u'+\alpha+qR_0(\mu,m')).$$

Let $F((n',v'),(\nu,\beta))=f(\Phi_q((n',v'),(\nu,\beta)))$ and

$$G_{j,a_1,a_2}((m',u'),(\mu,\alpha))=\mathcal{T}^{N,q}_{\geqslant j,a_1,a_2}(f)(\Phi_q((m',u'),(\mu,\alpha))).$$

The definitions show that

$$\begin{aligned}&G_{j,a_1,a_2}((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}}F((n',v'),(\nu,\beta))\sum_{j'=j}^\infty K_{j'}(q(m'-n')+E_1)\\&\qquad\times\eta_{2^{2j'}/N}(q^2(u'-v'-R_0(m'-n',n'))+E_2)e^{2\pi i[a_1\cdot(\mu-\nu)+a_2\cdot(\alpha-\beta-R_0(\mu-\nu,\nu))]/q},\end{aligned}$$

where $E_1=\mu-\nu$ and

$$E_2=(\alpha-\beta-R_0(\mu-\nu,\nu))+q(R_0(\mu,m'-n')-R_0(m'-n',\nu)).$$

Clearly, $|E_1|\leqslant Cq$ and $|E_2|\leqslant C2^{j'}q$ if $|m'-n'|\leqslant C2^{j'}/q$. Let

$$\begin{aligned}&\widetilde{G}_{j,a_1,a_2}((m',u'),(\mu,\alpha))\\&\quad=\sum_{(n',v')\in\mathbf{G}_0^\#}\sum_{(\nu,\beta)\in\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}}F((n',v'),(\nu,\beta))\sum_{j'=j}^\infty q^dK_{j'}(q(m'-n'))q^{2d'}\\&\qquad\times\eta_{2^{2j'}/N}(q^2(u'-v'-R_0(m'-n',n')))q^{-d}q^{-2d'}e^{2\pi i[a_1\cdot(\mu-\nu)+a_2\cdot(\alpha-\beta-R_0(\mu-\nu,\nu))]/q}.\end{aligned} \tag{7.32}$$

In view of the estimates above on $|E_1|$ and $|E_2|$, we have

$$\begin{aligned}&|G_{j,a_1,a_2}((m',u'),(\mu,\alpha))-\widetilde{G}_{j,a_1,a_2}((m',u'),(\mu,\alpha))|\\&\quad\leqslant C\sum_{(n',v')\in\mathbf{G}_0^{\#}}\;\sum_{(\nu,\beta)\in\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}}|F((n',v'),(\nu,\beta))|q^{-d}q^{-2d'}\\&\qquad\times\sum_{j'=j}^{\infty}\frac{qN}{2^{j'}}\left(\frac{2^{j'}}{q}\right)^{-d}\mathbf{1}_{[0,C2^{j'}/q]}(|m'-n'|)\phi_{2^{2j'}/Nq^2}(u'-v'-R_0(m'-n',n')),\end{aligned}$$

where ϕ is as in definition (7.7). Thus,

$$\Big\|\sup_{a_1,a_2,j\geqslant\log_2(2Nq)}|G_j-\widetilde{G}_j|\Big\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}])}\leqslant C\|F\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}])}.$$

For estimate (7.31), it suffices to prove that

$$\Big\|\sup_{a_1,a_2,j\geqslant\log_2(2Nq)}|\widetilde{G}_j|\Big\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}])}\leqslant C_p(\log(N+1))^2\|F\|_{L^p(\mathbf{G}_0^{\#}\times[\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}])},\qquad(7.33)$$

where $\widetilde{G}_j$ is defined in formula (7.32). We examine definition (7.32) and notice first that

$$q^{2d'}\eta_{2^{2j'}/N}(q^2(u'-v'-R_0(m'-n',n')))=\eta_{(2^{j'}/q)^2/N}(u'-v'-R_0(m'-n',n')).$$

Fix j_0 as the smallest integer with the property that $c_0 2^{j_0}/q=\tilde{c}_0\in\left[\frac{1}{2},2\right]$. The kernels $\widetilde{K}_j(x)=q^dK_{j+j_0}(qx)$, $j\geqslant\log_2 N$, have the properties (6.1), (6.2) and (7.23). Let

$$\widetilde{N}=\frac{q^2N}{2^{2j_0}}$$

and define $\tilde{L}_j^{\widetilde{N}}$ and $\widetilde{T}_*^{\widetilde{N}}$ as before, using the kernels $\widetilde{K}_j$. Then, from definition (7.32),

$$|G_{j,a_1,a_2}((m',u'),(\mu,\alpha))|\leqslant\sum_{(\nu,\beta)\in\mathbf{Z}_q^d\times\mathbf{Z}_{q^2}^{d'}}q^{-d}q^{-2d'}\widetilde{T}_*^{\widetilde{N}}(F((\cdot,\cdot),(\nu,\beta))(m',u').$$

The bound (7.33) follows from Lemma 7.2. □

Finally, we prove a weighted version of Lemma 7.3.

LEMMA 7.4. (Weighted maximal oscillatory singular integrals) *Assume that* $w\in L^\infty(\mathbf{G}_0^{\#})$, $w:\mathbf{G}_0^{\#}\to(0,\infty)$, *satisfies condition* (7.9), *i.e.*

$$\mathcal{M}_*^N(w)(h)\leqslant\varrho w(h)\quad\textit{for any }h\in\mathbf{G}_0^{\#}.$$

Then, for any compactly supported function $f:\mathbf{G}_0\to\mathbf{C}$,

$$\|\mathcal{T}_*^{N,q}(f)\|_{L^p(w)}\leqslant C_p\varrho^6(\log(N+1))^3\|f\|_{L^p(w)},\quad p\in(1,\infty),\qquad(7.34)$$

where $\mathcal{T}_*^{N,q}$ *is the maximal operator defined in Lemma* 7.3.

Proof. We use the method of distributional inequalities, as in [14, Chapter V]. Fix $p_1=\frac{1}{2}(p+1)\in(1,p)$, and assume that we could prove the distributional inequality

$$\begin{aligned} w(\{h:\mathcal{T}_*^{N,q}(f)(h)>\alpha \text{ and } \widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|)^{p_1})^{1/p_1}(h)\leqslant\gamma_1\alpha\}) \\ \leqslant(1-\gamma_3)w(\{h:\mathcal{T}_*^{N,q}(f)(h)>(1-\gamma_2)\alpha\}) \end{aligned} \tag{7.35}$$

for any $\alpha\in(0,\infty)$, for some small constants $\gamma_1,\gamma_2,\gamma_3>0$ depending on p, ϱ and $\log(N+1)$ with the property

$$1-\frac{\gamma_3}{2}<(1-\gamma_2)^p. \tag{7.36}$$

By integrating and using the assumptions that f is compactly supported and $w\in L^\infty(\mathbf{G}_0^\#)$ (so $\mathcal{T}_*^{N,q}(f)\in L^p(w)$, $p\in(1,\infty]$), it would follow that

$$\begin{aligned} \|\mathcal{T}_*^{N,q}(f)\|_{L^p(w)} &\leqslant \frac{1}{\gamma_1[1-(1-\gamma_3)(1-\gamma_2)^{-p}]^{1/p}}\|\widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|)^{p_1})^{1/p_1}\|_{L^p(w)} \\ &\leqslant C_p(\gamma_1\gamma_3)^{-1}\varrho^4\log(N+1)\|f\|_{L^p(w)}, \end{aligned} \tag{7.37}$$

using Lemma 7.1. Thus, for estimate (7.34), it suffices to prove the distributional inequality (7.35) with property (7.36) satisfied and control over $(\gamma_1\gamma_3)^{-1}$.

The bound (7.12) shows easily that if Q is a "cube" (i.e. $B\subseteq Q\subseteq B^*$ for some ball $B\in\mathbf{B}$) and $F\subseteq Q$, then

$$\frac{w(F)}{w(Q)}\leqslant 1-\frac{1}{C\varrho}\left(1-\frac{|F|}{|Q|}\right). \tag{7.38}$$

Indeed, the bound (7.38) is equivalent to $|G|/|Q|\leqslant C\varrho w(G)/w(Q)$ for any $G\subseteq Q$, which follows from (7.12). To prove inequality (7.35), we fix $\gamma_3=(C\varrho)^{-1}$ and $\gamma_2=(C_p\varrho)^{-1}$ such that property (7.36) holds. Let E denote the bounded set

$$E=\{h:\mathcal{T}_*^{N,q}(f)(h)>(1-\gamma_2)\alpha\},$$

and let $E=\bigcup_{k=1}^K Q_k$ be its Whitney decomposition in disjoint cubes (see properties (7.4)). For inequality (7.35) it suffices to prove that

$$w(\{h\in Q_k:\mathcal{T}_*^{N,q}(f)(h)>\alpha \text{ and } \widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|)^{p_1})^{1/p_1}(h)\leqslant\gamma_1\alpha\})\leqslant(1-\gamma_3)w(Q_k)$$

for $k=1,...,K$. In view of inequality (7.38), it suffices to prove that

$$|\{h\in Q_k:\mathcal{T}_*^{N,q}(f)(h)>\alpha \text{ and } \widetilde{\mathcal{M}}(\mathcal{M}_*^N(|f|)^{p_1})^{1/p_1}(h)\leqslant\gamma_1\alpha\}|\leqslant\frac{|Q_k|}{2} \tag{7.39}$$

for $k=1,...,K$ and some constant $\gamma_1>0$.

Since Q_k is a Whitney cube,

$$\mathcal{T}_*^{N,q}(f)(h_0)\leqslant(1-\gamma_2)\alpha \quad \text{for some } h_0\in B_k^{**}. \tag{7.40}$$

In addition, either the inequality (7.39) is trivial or

$$\widetilde{\mathcal{M}}(|f|^{p_1})^{1/p_1}(h_1) \leqslant \gamma_1\alpha \quad \text{for some } h_1 \in B_k^*, \tag{7.41}$$

since $|f(h)| \leqslant \mathcal{M}_*^N(|f|)(h)$ for any $h \in \mathbf{G}_0^\#$. Let $f_1 = f\mathbf{1}_{B_k^{**}}$, $f_2 = f\mathbf{1}_{^cB_k^{**}}$ and $f = f_1 + f_2$. The left-hand side of inequality (7.39) is dominated by

$$\begin{aligned} &\big|\{h : \mathcal{T}_*^{N,q}(f_1)(h) > \tfrac{1}{2}\gamma_2\alpha\}\big| \\ &\qquad + \big|\{h \in B_k^* : \mathcal{T}_*^{N,q}(f_2)(h) > (1-\tfrac{1}{2}\gamma_2)\alpha \text{ and } \mathcal{M}_*^N(|f_2|)(h) \leqslant \gamma_1\alpha\}\big|. \end{aligned} \tag{7.42}$$

However, using Lemma 7.3, the definition $\gamma_2 = (C_p\varrho)^{-1}$ and property (7.41),

$$\begin{aligned} \big|\{h : \mathcal{T}_*^{N,q}(f_1)(h) > \tfrac{1}{2}\gamma_2\alpha\}\big| &\leqslant \frac{C_p}{(\gamma_2\alpha)^{p_1}} \|\mathcal{T}_*^{N,q}(f_1)\|_{L^{p_1}}^{p_1} \\ &\leqslant C_p \alpha^{-p_1} \varrho^{p_1} (\log(N+1))^{2p_1} \int_{B_k^{**}} |f(h)|^{p_1}\, dh \\ &\leqslant C_p (\gamma_1 \varrho (\log(N+1))^2)^{p_1} |Q_k|. \end{aligned} \tag{7.43}$$

We now fix $\gamma_1 = (C_p\varrho(\log(N+1))^2)^{-1}$, for C_p sufficiently large, and show that the set in the second line of expression (7.42) is empty. Assuming this, the bound (7.39) follows and Lemma 7.4 follows from estimate (7.37).

It remains to show that the set in the second line of expression (7.42) is empty. We will use property (7.40) and the definition of the operators $\mathcal{T}_*^{N,q}$. Assume that the ball B_k^* has radius $r \in [2^{k_0}, 2^{k_0+1})$, $k_0 \in [-1, \infty) \cap \mathbf{Z}$. We notice that if $h \in B_k^*$ and $g \in {}^cB_k^{**}$ then $d(0, hg^{-1}) \geqslant \frac{1}{2}c^* 2^{k_0}$. If, in addition, $\log_2 N \leqslant j \leqslant k_0$, then

$$|L_{j,a_1,a_2}^{N,q}(hg^{-1})| \leqslant CA_j^N(hg^{-1}) \leqslant C2^{j-k_0} A_{k_0+2}^N(hg^{-1}),$$

thus, for any $j \in [\log_2 N, k_0] \cap \mathbf{Z}$, $a_1 \in \mathbf{Z}^d$ and $a_2 \in \mathbf{Z}^{d'}$,

$$|\mathcal{T}_{j,a_1,a_2}^{N,q}(f_2)(h)| \leqslant C2^{j-k_0} \mathcal{M}_*^N(|f_2|)(h).$$

Since $|\mathcal{T}_{j,a_1,a_2}^{N,q}(f_2)(h)| \leqslant C\mathcal{M}_*^N(|f_2|)(h)$ for any $j \geqslant \log_2 N$, $j \in [k_0, k_0 + \log N + C]$, we have for any $h \in B_k^*$,

$$\sup_{\substack{a_1 \in \mathbf{Z}^d \\ a_2 \in \mathbf{Z}^{d'}}} \sum_{j \in [\log_2 N, k_0 + \log N + C]} |\mathcal{T}_{j,a_1,a_2}^{N,q}(f_2)(h)| \leqslant C \log(N+1) \mathcal{M}_*^N(|f_2|)(h). \tag{7.44}$$

Now let $j \geqslant \min(\log_2(2Nq), k_0 + \log N + C)$, $a_1 \in \mathbf{Z}^d$, $a_2 \in \mathbf{Z}^{d'}$ and $h = (n, v) \in B_k^*$. With $h_0 = (n_0, v_0)$ as in property (7.40), let $a_{1,0} \in \mathbf{Z}^d$ be such that

$$a_{1,0} \cdot m = a_1 \cdot m + a_2 \cdot R_0(n - n_0, m) \quad \text{for any } m \in \mathbf{Z}^d.$$

Then, from the definitions,

$$\begin{aligned}
&\bigg|\sum_{j'=j}^{\infty}\mathcal{T}^{N,q}_{j',a_1,a_2}(f_2)(h)\bigg|-\bigg|\sum_{j'=j}^{\infty}\mathcal{T}^{N,q}_{j',a_{1,0},a_2}(f_2)(h_0)\bigg| \\
&\qquad\leqslant\bigg|\int_{{}^cB_k^{**}} f_2(m,u)e^{-2\pi i(a_1\cdot m+a_2\cdot u+a_2\cdot R_0(n-m,m))/q} \\
&\qquad\qquad\times\sum_{j'=j}^{\infty}[L^N_{j'}((n,v)\cdot(m,u)^{-1})-L^N_{j'}((n_0,v_0)\cdot(m,u)^{-1})]\,dm\,du\bigg| \\
&\qquad\leqslant\int_{{}^cB_k^{**}}|f_2(g)|\sum_{j=k_0+\log N+C}^{\infty}|L^N_j(hg^{-1})-L^N_j(h_0g^{-1})|\,dg.
\end{aligned}\tag{7.45}$$

An estimate similar to (7.20) shows that

$$|L^N_j(hg^{-1})-L^N_j(h_0g^{-1})|\leqslant C(N+1)2^{k_0-j}A^N_{j+3}(hg^{-1}),$$

since $h,h_0\in B_k^{**}$ and $j\geqslant k_0+\log N+C$. In addition, for $j'\geqslant j\geqslant k_0+\log N+C$,

$$\mathcal{T}^{N,q}_{j',a_{1,0},a_2}(f_2)(h_0)=\mathcal{T}^{N,q}_{j',a_{1,0},a_2}(f)(h_0)-\mathcal{T}^{N,q}_{j',a_{1,0},a_2}(f_1)(h_0)=\mathcal{T}^{N,q}_{j',a_{1,0},a_2}(f)(h_0).$$

Thus, from inequalities (7.44) and (7.45), for any $h\in B_k^*$,

$$\mathcal{T}^{N,q}_*(f_2)(h)\leqslant\mathcal{T}^{N,q}_*(f_2)(h_0)+C\log(N+1)\mathcal{M}^N_*(|f_2|)(h),$$

so the set in the second line of expression (7.42) is empty, as desired. This completes the proof of Lemma 7.4. □

References

[1] Arkhipov, G. I. & Oskolkov, K. I., A special trigonometric series and its applications. *Mat. Sb.*, 134 (176) (1987), 147–157, 287 (Russian); English translation in *Math. USSR–Sb.*, 62 (1989), 145–155.

[2] Bergelson, V. & Leibman, A., A nilpotent Roth theorem. *Invent. Math.*, 147 (2002), 429–470.

[3] Bourgain, J., On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.*, 61 (1988), 39–72.

[4] — On the pointwise ergodic theorem on L^p for arithmetic sets. *Israel J. Math.*, 61 (1988), 73–84.

[5] — Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, 69 (1989), 5–45.

[6] Christ, M., Nagel, A., Stein, E. M. & Wainger, S., Singular and maximal Radon transforms: analysis and geometry. *Ann. of Math.*, 150 (1999), 489–577.

[7] García-Cuerva, J. & Rubio de Francia, J. L., *Weighted Norm Inequalities and Related Topics.* North-Holland Mathematics Studies, 116. North-Holland, Amsterdam, 1985.

[8] IONESCU, A. D. & WAINGER, S., L^p boundedness of discrete singular Radon transforms. *J. Amer. Math. Soc.*, 19 (2006), 357–383.
[9] LEIBMAN, A., Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.*, 146 (2005), 303–315.
[10] MAGYAR, A., STEIN, E. M. & WAINGER, S., Discrete analogues of integral operators: maximal functions on the discrete Heisenberg group with applications to ergodic theory. To appear in *J. Anal. Math.*
[11] RICCI, F. & STEIN, E. M., Harmonic analysis on nilpotent groups and singular integrals. I. Oscillatory integrals. *J. Funct. Anal.*, 73 (1987), 179–194.
[12] — Harmonic analysis on nilpotent groups and singular integrals. II. Singular kernels supported on submanifolds. *J. Funct. Anal.*, 78 (1988), 56–84.
[13] RUBIO DE FRANCIA, J. L., A Littlewood–Paley inequality for arbitrary intervals. *Rev. Mat. Iberoamericana*, 1 (1985), 1–14.
[14] STEIN, E. M., *Harmonic Analysis*: *Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton Mathematical Series, 43. Princeton University Press, Princeton, NJ, 1993.
[15] STEIN, E. M. & WAINGER, S., Discrete analogues of singular Radon transforms. *Bull. Amer. Math. Soc.*, 23 (1990), 537–544.
[16] — Discrete analogues in harmonic analysis. I. l^2 estimates for singular Radon transforms. *Amer. J. Math.*, 121 (1999), 1291–1336.
[17] — Two discrete fractional integral operators revisited. *J. Anal. Math.*, 87 (2002), 451–479.

ALEXANDRU D. IONESCU
University of Wisconsin, Madison
Madison, WI
U.S.A.
ionescu@math.wisc.edu

AKOS MAGYAR
University of Georgia, Athens
Athens, GA
U.S.A.
magyar@math.uga.edu

ELIAS M. STEIN
Princeton University
Princeton, NJ
U.S.A.
stein@math.princeton.edu

STEPHEN WAINGER
University of Wisconsin, Madison
Madison, WI
U.S.A.
wainger@math.wisc.edu

Received March 21, 2006