Annals of Mathematics, **155** (2002), 189–208

Discrete analogues in harmonic analysis: Spherical averages

By A. Magyar, E. M. Stein, and S. Wainger*

Abstract

In this paper we prove an analogue in the discrete setting of $\mathbb{Z}^d$, of the spherical maximal theorem for $\mathbb{R}^d$. The methods used are two-fold: the application of certain "sampling" techniques, and ideas arising in the study of the number of representations of an integer as a sum of d squares, in particular, the "circle method". The results we obtained are by necessity limited to $d \geq 5$, and moreover the range of p for the L^p estimates differs from its analogue in $\mathbb{R}^d$.

1. Introduction

Geometric considerations, in particular curvature, play an important role in harmonic analysis in $\mathbb{R}^d$. Emblematic of this are the properties of the spherical maximal function. Given the significance of this operator, it is an interesting and natural question to ask what happens when we consider its discrete analogue; that is, what can be said of the corresponding version of the spherical maximal theorem taken over $\mathbb{Z}^d$? It is the purpose of this paper to answer this question by proving optimal ℓ^p estimates in this setting.

We shall now describe these results, turning first to $\mathbb{R}^d$. The spherical averages are defined by the operators $\mathcal{A}_\lambda$,where

$$\mathcal{A}_\lambda(f) = f \star d\sigma_\lambda$$

with $d\sigma_\lambda$ the normalized invariant measure on the sphere $|x| = \lambda$. With the definition of the maximal function, $\mathcal{A}_\star(f)(x) = \sup_{0<\lambda<\infty} |\mathcal{A}_\lambda(f)(x)|$, we recall the main estimate for it,

$$\|\mathcal{A}_\star(f)\|_{L^p(\mathbb{R}^d)} \leq A\, \|f\|_{L^p(\mathbb{R}^d)}\,, \quad \text{if} \quad p > \frac{d}{d-1} \quad \text{and} \quad d \geq 2\,. \tag{1.1}$$

(See [S], [SW1], [B1].)

*This work was supported by NSF Grants: DMS-9970899, DMS-9706889, and DMS-9731647.

The discrete analogue of $\mathcal{A}_\lambda$ is the operator

$$A_\lambda(f)(n) = \frac{1}{N(\lambda)} \sum_{|m|=\lambda} f(n-m)\,. \tag{1.2}$$

Here n and m are restricted to range over $\mathbb{Z}^d$; also $N(\lambda) =$ the number of $m \in \mathbb{Z}^d$, so that $|m| = \lambda$. Notice that only those λ for which λ^2 is an integer are relevant; also observe that $N(\lambda) = r_d(\lambda^2)$, where $r_d(k)$ is the standard counting function giving the number of ways of representing k as a sum of d squares.

Now, up to this point, formulating a discrete analogue of the spherical maximal function, i.e. $A_\star(f)(n) = \sup\limits_{0<\lambda<\infty} |A_\lambda(f)(n)|$, and asking the question of its ℓ^p boundedness, seem quite straightforward.* However, this is misleading since quite different ideas must come into play in the discrete analogue, and anyway, the range of exponents is not the same as the version in $\mathbb{R}^d$. The theorem we prove is the following optimal result.

THEOREM. *The maximal operator $A_\star$ is bounded in $\ell^p(\mathbb{Z}^d)$ to itself for $p > \frac{d}{d-2}$, when $d \geq 5$.*

Alex Ionescu has pointed out to us that simple examples show that this result cannot be improved: in fact, when $d \geq 5$, $A_\star$ is not bounded on ℓ^p for $p \leq \frac{d}{d-2}$; moreover when $d < 5$, the ℓ^p boundedness fails for every $p < \infty$ (the case $p = \infty$ is of course trivial). The relevant examples can be found in Section 8. Here the facts that the number of representations $r_d(k)$ is an irregular function of k when $d \leq 4$, while $r_d(k) \approx k^{\frac{d-2}{2}}$ when $d \geq 5$, play a role. (For these assertions about $r_d(k)$, consult [W].)

Our attack on the discrete spherical maximal function proceeds in three stages. To begin with (motivated by the ideas of the circle method) we approximate A_λ by an infinite sum of simpler operators

$$M_\lambda = c_d \sum e^{-2\pi i \lambda^2 a/q}\, M_\lambda^{a/q}\,, \tag{1.3}$$

with each $M_\lambda^{a/q}$ associated to a reduced fraction a/q, with $0 < a/q \leq 1$. Now since each $M_\lambda^{a/q}$ is a convolution operator on $\mathbb{Z}^d$, it corresponds to a Fourier multiplier $m_\lambda^{p/q}(\xi)$, which is given by

$$m_\lambda^{a/q}(\xi) = \sum_{\ell \in \mathbb{Z}^d} G(a/q, \ell)\, \Phi_q(\xi - \ell/q)\, d\hat{\sigma}_\lambda(\xi - \ell/q)\,.$$

*Here we use the notation that for f defined on $\mathbb{Z}^d$, it belongs to $\ell^p(\mathbb{Z}^d)$ if $\sum_{n\in\mathbb{Z}^d} |f(n)|^p$ is finite. The ℓ^p norm is of course the p^{th} root of the sum.

Here G is a normalized Gauss sum, Φ_q is a suitable cut-off function, and $d\hat{\sigma}_\lambda$ is the Fourier transform of the unit measure $d\sigma_\lambda$ on the sphere $|x| = \lambda$.

Notice that the first term of M_λ, corresponding to $a/q = 1 \equiv 0 \bmod 1$, can be viewed as the vestige of the continuous analogue on $\mathbb{R}^d$. All the other terms are approximations corresponding to the other rationals.

The second stage is to study each $M_\lambda^{a/q}$ as a sort of discrete analogue of an operator on $\mathbb{R}^d$. The main tool is a general abstract theorem which allows one to pass from certain convolution operators on $\mathbb{R}^d$ to analogous operators on $\mathbb{Z}^d$. While ideas about special cases of this principle have been implicit in the past, our general approach seems both new interesting in its own right. It is presented in Section 2. It is based in part on variants of "sampling" ideas which go back to Plancherel and Pólya [PP] and which were taken up again later by Shannon [ShW]. Using arguments of a different kind, Bourgain obtained certain results of this form; see [B2, (3.5)].

The final stage of the argument is to show that M_λ is an adequate approximation of A_λ. This is begun in Sections 4 and 5, and is concluded in Sections 6 and 7.

The analysis of our theorem has as its starting point a partial result obtained previously by one of us [M] (see Proposition (4.2) below). The interested reader may also want to compare the related ways the sums $\displaystyle\sum_{|n|=\lambda} e^{2\pi i n\cdot\xi}$ are treated in our paper (see Section 5), and in a previous work of Bleher and Bourgain [BB, §6]. The context of that paper is however quite different from ours.

2. Discrete analogues of convolution operators

Suppose $T(f) = f \star K$ is a convolution operator in $L^p(\mathbb{R}^d)$ to itself with a suitable distribution kernel K. Then, as is known, its Fourier transform $\widehat{K} = m(\xi) = \displaystyle\int_{\mathbb{R}^d} K(x) e^{-2\pi i x\xi} dx$ is a bounded function, and we can think of T as a Fourier multiplier operator given by $(Tf)^\wedge(\xi) = m(\xi)\hat{f}(\xi)$.

To be precise, in what follows we shall assume in this section that in addition to $m(\xi)$ being bounded, it is supported in the fundamental cube $Q = \{\xi = (\xi_j) : -1/2 < \xi_j \leq 1/2,\ j = 1, \dots \ell\}$. In this case $K(x) = \displaystyle\int_{\mathbb{R}^d} e^{2\pi i x\xi}\, m(\xi) d\xi$ is an L^2 function on $\mathbb{R}^d$, which is continuous (in fact, C^∞). Thus $K_{\mathrm{dis}} = K\Big|_{\mathbb{Z}^d}$ is well-defined, as is the convolution operator acting on functions on $\mathbb{Z}^d$ given by

$$T_{\mathrm{dis}}(f) \;=\; f \star K_{\mathrm{dis}}\,, \qquad T_{\mathrm{dis}}(f)(n) \;=\; \sum_{m\in\mathbb{Z}^d} K(m)\, f\,(n-m)\,.$$

Note that the condition that the multiplier be supported in Q is natural. Because then not only does T determine T_{dis}, but conversely T_{dis} determines T, i.e. $K\Big|_{\mathbb{Z}^d}$ determines K. This follows since $K\Big|_{\mathbb{Z}^d}$ determines the Fourier coefficients of the function $m(\xi)$, when expanded as a function on Q.

Let m_{per} be the periodic extension of m, i.e. $m_{\text{per}}(\xi) = \sum_{\ell \in \mathbb{Z}^d} m(\xi - \ell)$. Then $m_{\text{per}}(\xi)$ is the Fourier multiplier corresponding to T_{dis} in the sense

$$\sum_{n \in \mathbb{Z}^d} T_{\text{dis}}(f)(n)\, e^{-2\pi i n \xi} = m_{\text{per}}(\xi) \sum_{n \in \mathbb{Z}^d} f(n) e^{-2\pi i n \xi},$$

for suitable functions f on $\mathbb{Z}^d$.

Let us note that

$$m_{\text{per}}(\xi) = \sum_{n \in \mathbb{Z}^d} K(n)\, e^{-2\pi i n \xi} \tag{2.0}$$

in the sense of L^2 convergence of the series on any compact subset of $\mathbb{R}^d$.

In fact, $m(\xi) = \sum_{n \in \mathbb{Z}^d} K(n)\, e^{-2\pi i n \xi}$ on Q represents the Fourier inversion of the identity $K(n) = \int m(\xi) e^{2\pi i n \xi}\, d\xi$ ($m(\xi)$ is supported in Q); and, moreover, $m_{\text{per}}(\xi)$ is the periodic function which agrees with $m(\xi)$ on Q. This establishes (2.0).

The question we will be concerned with is how the norm of T_{dis} as an operator on $\ell^p(\mathbb{Z}^d)$ is controlled by the norm of the operator T acting on $L^p(\mathbb{R}^d)$. For our applications it will be important to be able to deal with the more general case where the L^p and ℓ^p spaces of complex-valued functions are replaced by the spaces $L^p_B(\mathbb{R}^d)$ and $\ell^p_B(\mathbb{Z}^d)$ of functions taking their values in the Banach space B. In order to avoid technical problems involving definability, measurability, etc., we shall restrict our attention to the case when the Banach spaces in question are finite-dimensional. However, all our estimates will be independent of the Banach spaces in question, so that a limiting argument will encompass the results in the generality needed. In particular, this argument will apply to the case when B is an L^∞ space, which is what is needed for the maximal theorems below.

We shall suppose that B_1 and B_2 are a pair of finite-dimensional Banach spaces, and assume that $m(\xi)$ is a bounded measurable function, taking its values in $\mathcal{L}(B_1, B_2)$; and as we have said we suppose m is supported in Q. Then T, described above, is a bounded mapping from $L^2_{B_1}(\mathbb{R}^d)$ to $L^2_{B_2}(\mathbb{R}^d)$, and similarly T_{dis} is bounded from $\ell^2_{B_1}(\mathbb{Z}^d)$ to $\ell^2_{B_2}(\mathbb{Z}^d)$.

Proposition 2.1. *Fix p, $1 \le p \le \infty$. If T is bounded from $L^p_{B_1}(\mathbb{R}^d)$ to $L^p_{B_2}(\mathbb{R}^d)$, then T_{dis} is bounded from $\ell^p_{B_1}(\mathbb{Z}^d)$ to $\ell^p_{B_2}(\mathbb{Z}^d)$. For these operators we have the norm inequality*

$$\left\| T_{\text{dis}} \right\|_{\ell^p_{B_1} \to \ell^p_{B_2}} \le C \left\| T \right\|_{L^p_{B_1} \to L^p_{B_2}}, \tag{2.1}$$

with a bound C that depends only on the dimension d, but not on p or the Banach spaces B_1 and B_2.

Remarks. (1) It would be interesting to know if C can be taken to be independent of the dimension d, or for that matter if $C = 1$.

(2) There is a converse to (2.1); i.e., a reverse inequality also holds. Since that fact will not be used below we will omit its proof.

The proof of the proposition requires the following "sampling" and extension lemma. We fix the function Ψ on $\mathbb{R}^d$ by

$$\Psi(x) = \left(\frac{\sin \pi x_1}{\pi x_1}\right)^2 \left(\frac{\sin \pi x_2}{\pi x_2}\right)^2 \cdots \left(\frac{\sin \pi x_d}{\pi x_d}\right)^2, \quad x = (x_1, x_2, \dots x_d).$$

For any suitable function f on $\mathbb{Z}^d$ we consider its extension $f_{\text{ext}} = F$ on $\mathbb{R}^d$ given by

$$F(x) = f_{\text{ext}}(x) = \sum_{n \in \mathbb{Z}^d} f(n)\, \Psi(x - n). \tag{2.2}$$

(Note that if $f \in \ell^p$ for some p, the series above converges for every $x \in \mathbb{R}^d$.) We observe that in fact $F\Big|_{\mathbb{Z}^d} = f$, since $\Psi(0) = 1$, and $\Psi(n) = 0$, if $n \in \mathbb{Z}^d$, $n \neq 0$; thus f_{ext} is a genuine extension of f. The following estimate holds for any (finite-dimensional) Banach space B.

Lemma 2.1. *If $f \in \ell^p(\mathbb{Z}^d, B)$, then $F \in L^p(\mathbb{R}^d, B)$, and*

$$(1/A)\, \|f\|_{\ell^p_B} \le \|F\|_{L^p_B} \le A\, \|f\|_{\ell^p_B}. \tag{2.3}$$

Here A is a constant that depends only on d, but not p or the space B.

Ideas of this kind go back to Plancherel and Pólya [PP]. In that work (when e.g. $d = 1$), the function $\frac{\sin \pi x}{\pi x}$ was used in effect in place of $\left(\frac{\sin \pi x}{\pi x}\right)^2$. The resulting version of (2.3) is then more delicate and holds only in the range $1 < p < \infty$, since it involves the Hilbert transform; it also does not cover the case of Banach space-valued functions.

To prove the lemma we observe two easily established estimates,

$$\int_{\mathbb{R}^d} \Psi(x)dx \le A_1 \quad \text{and} \quad \sup_x \sum_{n \in \mathbb{Z}^d} \Psi(x - n) \le A_1.$$

Then for any $p < \infty$, by Hölder' inequality

$$|f_{\text{ext}}(x)|^p \le \left(\sum_n |f(n)|^p\, \Psi(x - n)\right) \left(\sum_n \Psi(x - n)\right)^{p-1}$$

and integration in x then gives

$$\| f_{\text{ext}} \|^p_{L^p_B} \le \| f \|^p_{\ell^p_B} \ A_1^p .$$

The proof of the corresponding result for $p = \infty$ is similar but simpler.

To prove the converse inequality, choose $\widehat{\Phi}$ to be a C^∞ function with compact support so that $\widehat{\Phi}(\xi) = 1$, when $\xi \in 2Q$. Since $\left[\left(\frac{\sin \pi x_1}{\pi x_1}\right)^2\right]^\wedge = (1 - |\xi_1|)_+$, it follows that

$$\widehat{\Psi} \cdot \widehat{\Phi} = \widehat{\Psi}, \text{ and hence } \Psi \star \Phi = \Psi .$$

As a result $f_{\text{ext}} \star \Phi = f_{\text{ext}}$ and since $f_{\text{ext}}(n) = f(n)$, we have

$$f(n) = \int_{\mathbb{R}^d} f_{\text{ext}}(y)\, \Phi(n-y)\, dy .$$

Thus as before,

$$\left|f(n)\right|^p \le \left(\int_{\mathbb{R}^d} \left|f_{\text{ext}}(y)\right|^p \left|\Phi(n-y)\right| dy\right) \left(\int_{\mathbb{R}^d} \left|\Phi(n-y)\right| dy\right)^{p-1} ,$$

and

$$\|f\|^p_{\ell^p_B} \le \|f_{\text{ext}}\|_{L^p_B} \ A_2^p ,$$

if

$$\int_{\mathbb{R}^d} |\Phi(y)|\, dy \le A_2 , \quad \text{and} \quad \sup_y \sum_{n \in \mathbb{Z}^d} |\Phi(n-y)| \le A_2 .$$

The argument also gives the case $p = \infty$. The lemma, inequality (2.3), is thus established with $A = \max(A_1, A_2)$.

To prove the proposition, we consider the cube $3Q$ which can be covered by 3^d disjoint translates of Q. In fact, it is easily verified that $3Q = \underset{\substack{\ell \in \mathbb{Z}^d, \ell = (\ell_1, \dots \ell_d) \\ \sup_j |\ell_j| \le 1}}{U} (Q + \ell)$. Now let $m(\xi)$ be continued periodically to $3Q$, i.e. define $\widetilde{m}(\xi)$ by $\widetilde{m}(\xi) = \sum_{\sup_j |\ell_j| \le 1} m(\xi + \ell)$. We let $\widetilde{T}$ denote Fourier multiplier operator, whose multiplier is $\widetilde{m}(\xi)$. Then clearly

$$\| \widetilde{T} \|_{L^p_{B_1} \to L^p_{B_2}} \le 3^d \ \| T \|_{L^p_{B_1} \to L^p_{B_2}} . \tag{2.4}$$

On the other hand, we claim that

$$\widetilde{T}(f_{\text{ext}}) = (T_{\text{dis}}(f))_{\text{ext}} . \tag{2.5}$$

To verify (2.5) it suffices to do it for $f = \delta_m$, for every fixed m, where

$$\delta_m(n) = \begin{cases} 1 & \text{if} \quad n = m \\ \\ 0 & \text{if} \quad n \ne m . \end{cases}$$

We will check this by taking the Fourier transform of both sides of (2.5). Indeed

$$T_{\mathrm{dis}}(f)(n) \,=\, K(n-m)$$

and

$$(T_{\mathrm{dis}}(f))_{\mathrm{ext}} \,=\, \sum_{n\in\mathbb{Z}^d} K(n-m)\,\Psi(x-n)\,.$$

Hence,

$$\begin{aligned}(T_{\mathrm{dis}}(f)_{\mathrm{ext}})^{\wedge} &= \sum_n K(n-m)\,\widehat{\Psi}\,(\xi)\,e^{-2\pi i n\xi} \\ &= \Big(\sum K(n-m)\,e^{-2\pi i n\xi}\Big)\;\widehat{\Psi}(\xi) \\ &= \left(\sum_n K(n)\,e^{-2\pi i n\xi}\right)\,e^{-2\pi i m\xi}\;\widehat{\Psi}(\xi) \\ &= m_{\mathrm{per}}(\xi)\,\widehat{\Psi}(\xi)\,e^{-2\pi i m\xi} \qquad\qquad \text{(by (2.0))}.\end{aligned}$$

On the other hand

$$\widetilde{T}(f_{\mathrm{ext}})^{\wedge} \,=\, \tilde{m}(\xi)\,(f_{\mathrm{ext}})^{\wedge} \,=\, \tilde{m}(\xi)\,(\Psi(x-m))^{\wedge} \,=\, \tilde{m}(\xi)\widehat{\Psi}(\xi)\,e^{-2\pi i m\xi}.$$

Now we have the desired identity, since $\tilde{m}(\xi) = m_{\mathrm{per}}(\xi)$ on the support of $\widehat{\Psi}$ (note that $2Q \subset 3Q$).

Once (2.5) is established we have

$$\begin{aligned}\Big\|T_{\mathrm{dis}}(f)\Big\|_{\ell^p_{B_2}} &\le A\Big\|\,(T_{\mathrm{dis}}(f))_{\mathrm{ext}}\,\Big\|_{L^p_{B_2}} && \text{(by the lemma)} \\ &= A\Big\|\widetilde{T}(f_{\mathrm{ext}})\Big\|_{L^p_{B_2}} \,\le\, 3^d\,A\Big\|T(f_{\mathrm{ext}})\Big\|_{L^p_{B_2}} \\ &\le 3^d\,A\Big\|T\Big\|_{L^p_{B_1}\to L^p_{B_2}}\;\Big\|f_{\mathrm{ext}}\Big\|_{L^p_{B_1}} && \text{(by (2.4)))} \\ &\le 3^d\,A^2\Big\|T\Big\|_{L^p_{B_1}\to L^p_{B_2}}\;\Big\|f\Big\|_{\ell^p_{B_1}} && \text{(by the lemma)}.\end{aligned}$$

Thus $\Big\|T_{\mathrm{dis}}\Big\|_{\ell^p_{B_1}\to\ell^p_{B_2}} \,\le\, 3^d\,A^2\,\Big\|T\Big\|_{L^p_{B_1}\to L^p_{B_2}}$, and the proposition is proved with $C = 3^d\,A^2$.

We now fix an integer $q \ge 1$. We shall also make the stronger assumption that $m(\xi)$ is supported in Q/q, and consider m^q_{per} defined by

$$m^q_{\mathrm{per}}(\xi) \,=\, \sum_{\ell\in\mathbb{Z}^d} m(\xi - \ell/q)\,. \tag{2.6}$$

Notice that m^q_{per} is periodic with respect to elements in $(1/q)\,\mathbb{Z}^d$, and hence, in particular, periodic with respect to $\mathbb{Z}^d$.

We consider the operator T^q_{dis}, a convolution operator on $\mathbb{Z}^d$, having m^q_{per} as its Fourier multiplier; i.e.,

$$\sum_{m\in\mathbb{Z}^d} T^q_{\text{dis}}\,(f)(n)\,e^{-2\pi i n\xi} \;=\; m^q_{\text{per}}(\xi)\sum_{n\in\mathbb{Z}^d} f(n)\,e^{-2\pi i n\cdot\xi}\,,$$

for suitable f.

Corollary 2.1.

$$\left\|T^q_{\text{dis}}\right\|_{\ell^p_{B_1}\to\ell^p_{B_2}} \;\le\; C\left\|T\right\|_{L^p_{B_1}\to L^p_{B_2}}\,. \tag{2.7}$$

Again the bound C does not depend on p, B_1 and B_2; it is also independent of q.

Proof of the corollary. Let T^q be the operator on $L^p_{B_1}(\mathbb{R}^d)$ to $L^p_{B_2}(\mathbb{R}^d)$ whose multiplier is $m\,(\xi/q)$. Notice that $m(\xi/q)$ is supported for $\xi\in Q$.

Now a simple scaling argument shows

$$\left\|T^q\right\|_{L^p_{B_1}\to L^p_{B_2}} \;=\; \left\|T\right\|_{L^p_{B_1}\to L^p_{B_2}}\,,$$

and so if $(T^q)_{\text{dis}}$ is the discrete analogue in the sense of Proposition 1,

$$\left\|(T^q)_{\text{dis}}\right\|_{\ell^p_{B_1}\to\ell^p_{B_2}} \;\le\; C\left\|T\right\|_{L^p_{B_1}\to L^p_{B_2}}\,. \tag{2.8}$$

However, we must emphasize that $(T^q)_{\text{dis}}\neq T^q_{\text{dis}}$. In fact, the convolution kernel of $(T^q)_{\text{dis}}$, which comes from the multiplier $m(\xi/q)$, is $K^q(n)=q^dK(qn)$, $n\in\mathbb{Z}^d$.

Next we observe the convolution kernel, $K^\#(n)$, of T^q_{dis} is given by

$$K^\#(m) \;=\; \begin{cases} q^dK(m), & \text{if } m\in q\mathbb{Z}^d \\[2ex] =0 & \text{if } m\in\mathbb{Z}^d\,\text{, but } m\notin q\mathbb{Z}^d\,, \end{cases}$$

because

$$\int_Q\left(\sum_{\ell\in\mathbb{Z}^d} m(\xi-\ell/q)\right)\,e^{2\pi i\xi m}\,d\xi$$
$$=\int_Q\left(\sum_{\ell'\in\mathbb{Z}^d} m(\xi-\ell')\right)\,e^{2\pi i\xi m}\,d\xi\;\times\begin{cases} q^d & \text{if} \quad m\in q\mathbb{Z}^d \\[2ex] 0 & \text{if} \quad m\notin q\mathbb{Z}^d\,. \end{cases}$$

Now finally let $T_\#$ denote the operator mapping functions of $q\mathbb{Z}^d$ to itself, given by the kernel $K^\#$, i.e.

$$T_\#(f')(nq) \;=\; \sum_{m\in\mathbb{Z}^d} f'((n-m)q)\, K^\#(mq)\,.$$

Then clearly

$$\Big\|T_\#\Big\|_{\ell^p_{B_1}(q\mathbb{Z}^d)\to\ell^p_{B_2}(q\mathbb{Z}^d)} \;=\; \Big\|(T^q)_{\mathrm{dis}}\Big\|_{\ell^p_{B_1}(\mathbb{Z}^d)\to\ell^p_{B_2}(\mathbb{Z}^d)}$$

which is an immediate consequence of the isomorphism $\mathbb{Z}^d \leftrightarrow q\mathbb{Z}^d$, given by $n \leftrightarrow qn$, $(n\in\mathbb{Z}^d)$.

Finally note that T^q_{dis} can be written as $T_\# \otimes I$, if we write $\ell^p_B(\mathbb{Z}^d)$ as $\ell^p_B(q\mathbb{Z}^d)\otimes\ell^p(\mathbb{Z}^d\diagup q\mathbb{Z}^d)$, with $T_\#$ acting on the first factor, and the identity acting on the second factor.

As a result

$$\Big\|T^q_{\mathrm{dis}}\Big\|_{\ell^p_{B_1}\to\ell^p_{B_2}} \;\le\; \Big\|T_\#\Big\|_{\ell^p_{B_1}(q\mathbb{Z}^d)\to\ell^p_{B_2}(q\mathbb{Z}^d)}\,.$$

Combining this with (2.8) proves Corollary 2.1.

We next consider a version of a convolution operator, whose multiplier is somewhat akin to (2.4). Here we shall consider

$$m(\xi) \;=\; \sum_{\ell\in\mathbb{Z}^d} \gamma_\ell\, \Phi(\xi-\ell/q) \tag{2.9}$$

under the following assumptions:

(a) Φ is a C^∞ function supported on Q/q. As a function on Q it has the Fourier expansion

$$\Phi(\xi) \;=\; \sum_{m\in\mathbb{Z}^d} \varphi_m\, e^{-2\pi i m\xi}$$

with

$$\sum_{m\in\mathbb{Z}^d} |\varphi_m| \;\le\; A\,.$$

(b) $\{\gamma_\ell\}$ is a $q\mathbb{Z}^d$ periodic sequence; i.e., $\gamma_\ell=\gamma_{\ell'}$ if $\ell-\ell'\in q\mathbb{Z}^d$.

Now let $\{\hat\gamma_s\}$ be the Fourier transform of $\{\gamma_\ell\}$; i.e., $\hat\gamma_s \;=\; \sum_{\ell\in\mathbb{Z}^d\diagup q\mathbb{Z}^d} e^{2\pi i s\ell/q}\,\gamma_\ell\,.$

We shall also restrict our attention to scalar-related functions on $\ell^p(\mathbb{Z}^d)$, as opposed to the Banach-space case treated in the previous proposition, because of the specific use of Plancherel's identity. Our result is as follows.

Proposition 2.2. *Let T be the operator on functions on $\mathbb{Z}^d$ whose Fourier multiplier is given by* (2.8), *satisfying the conditions above. Then with* $1 \le p \le 2$,

$$\Big\|T\Big\|_{\ell^p(\mathbb{Z}^d)\to\ell^p(\mathbb{Z}^d)} \le A\left(\sup_\ell |\gamma_\ell|\right)^{2-2/p}\left(\sup_s |\hat{\gamma}_s|\right)^{2/p-1}. \tag{2.10}$$

For the case $p=2$ we use Parseval-Plancherel's theorem for $\mathbb{Z}^d$, together with the disjointedness of the supports of the $\Phi(\xi-\ell/q)$ and the fact that $\sup_\xi|\Phi(\xi)| \le A$. This implies that $|m(\xi)| \le A\sup_\ell|\gamma_\ell|$, yielding the case $p=2$.

For the case $p=1$, we calculate the $\ell^1(\mathbb{Z}^d)$ norm of the kernel $K(n)$, corresponding to the multiplier $m(\xi)$. It is given by

$$\begin{aligned} K(n) &= \int_Q \left(\sum \gamma_\ell \Phi(\xi-\ell/q)\right) e^{2\pi i n\xi}\, d\xi \\ &= \varphi(n)\left(\sum_{\ell\in\mathbb{Z}^d/q\mathbb{Z}^d} \gamma_\ell\, e^{2\pi i n\ell/q}\right) = \varphi(n)\hat{\gamma}_n\,. \end{aligned}$$

Hence by property (a), $\sum_{n\in\mathbb{Z}^d}|K(n)| \le A\sup_n|\hat{\gamma}_n|$, and as a result the case $p=1$ of (2.10) is proved. The general result for $1\le p\le 2$ then follows by Riesz' convexity theorem.

3. The main term

The averages we are interested in,

$$A_\lambda(f)(n) = \frac{1}{N(\lambda)}\sum_{|m|=\lambda} f(n-m)\,,$$

will be replaced by the equivalent averages

$$\frac{1}{\lambda^{d-2}}\sum_{|m|=\lambda} f(n-m)\,,$$

when $d\ge 5$. This equivalence comes about because, as we have pointed out with $N(\lambda)$ = number of $n\in\mathbb{Z}^d$, so that $|n|=\lambda$, we have $N(\lambda)\approx\lambda^{d-2}$, whenever λ^2 is an integer, and $d\ge 5$. In order not to introduce new notation, we shall designate these averages also by A_λ and now write

$$A_\lambda(f)(n) = \frac{1}{\lambda^{d-2}}\sum_{|m|=\lambda} f(n-m)\,, \tag{3.1}$$

and in what follows we shall always assume that λ is restricted so that λ^2 is an integer.

Here we shall deal with the main term in the approximation of A_λ. It is a convolution operator M_λ acting on functions on $\mathbb{Z}^d$, which can be written as

$$M_\lambda = c_d \sum_{q=1}^{\infty} \sum_{\substack{1\le a\le q \\ (a,q)=1}} e^{-2\pi i\lambda^2 a/q}\, M_\lambda^{a/q}, \tag{3.2}$$

where the sum is taken over all reduced fractions a/q, with $0 < a/q \le 1$. Hence c_d is the constant $= \frac{\pi^{d/2}}{\Gamma(d/2)}$. Also each $M_\lambda^{a/q}$ is the convolution operator whose multiplier is

$$\sum_{\ell\in\mathbb{Z}^d} G(a/q,\ell)\, \Psi_q\,(\xi-\ell/q)\, d\widehat{\sigma}_\lambda\,(\xi-\ell/q)\,. \tag{3.3}$$

In the above, $\Psi_q(\xi) = \Psi(q\xi)$, where Ψ is a C^∞ cut-off function supported in the cube $Q/2$, with $\Psi(\xi) = 1$, for $\xi \in Q/4$. Also $G(a/q,\ell)$ is the normalized Gauss sum

$$G(a/q,\ell) = q^{-d} \sum_{n\in\mathbb{Z}^d/q\mathbb{Z}^d} e^{2\pi i(|n|^2 a/q + n\cdot\ell/q)}, \tag{3.4}$$

and $d\hat{\sigma}_\lambda(\xi)$ is the Fourier transform of the normalized invariant measure $d\sigma_\lambda$ supported on the sphere $|x| = \lambda$. Note that (3.3) is periodic on ξ with periods in $\mathbb{Z}^d$, since $G(a/q,\ell) = G(a/q,\ell')$, if $\ell \equiv \ell' \bmod q\mathbb{Z}^d$; also, for each ξ only one term in (3.3) is nonzero.

We define the corresponding maximal operators,

$$M_\star(f)(n) = \sup_{0<\lambda<\infty} |M_\lambda(f)(n)|\,,$$

and

$$M_\star^{a/q}(f)(n) = \sup_{0<\lambda<\infty} \left| M_\lambda^{a/q}(f)(n) \right|.$$

The basic estimates for these are as follows:

Proposition 3.1. (a) *$\| M_\star^{a/q} \|_{\ell^p\to\ell^p} = O\left(q^{-d(1-1/p)}\right)$ if $d \ge 3$, and $\frac{d}{d-1} < p \le 2$,*

(b) $\| M_\star \|_{\ell^p\to\ell^p} \le A$ if $d \ge 5$, and $\frac{d}{d-2} < p \le 2$.

To prove part (a) we write $\Psi = \Psi\cdot\Psi'$, where Ψ' is another C^∞ function, supported in Q, with $\Psi'(\xi) = 1$ for $\xi \in Q/2$. Then the operator corresponding to the multiplier (3.3) can be written as a product of two operators, with multipliers respectively:

$$\sum_{\ell\in\mathbb{Z}^d} G(a/q,\ell)\, \Psi'_q\,(\xi-\ell/q)$$

and

$$\sum_{\ell\in\mathbb{Z}^d} \Psi_q(\xi-\ell/q)\, d\hat{\sigma}_\lambda\,(\xi-\ell/q)\,,$$

where $\Psi'_q(\xi) = \Psi'(q\xi)$, if we recall that for each ξ only one term in each of the above sums is nonvanishing.

To the first multiplier we apply Proposition 2.2 (in §2) with $\gamma_\ell = G(a/q,\ell)$, and $\Phi(\xi) = \Psi'_q(\xi)$. Note that $\Phi(\xi)$ is supported in Q/q, and moreover

$$\varphi(n) = \int_{\mathbb{R}^d} \Phi(\xi)\, e^{-2\pi i n\xi}\, d\xi = q^{-1}\, \widetilde{\Psi}(q^{-1}n)\,,$$

where $\widetilde{\Psi}$ is the Fourier transform of Ψ'. Now $|\widetilde{\Psi}(x)| \le A_N(1+|x|)^{-N}$ for all $N \ge 0$, so that $\sum_{n\in\mathbb{Z}^d} |\varphi(n)| \le A$.

Next, there is the estimate $|G(a/q,\ell)| = O(q^{-d/2})$; this is well-known, but in any case it follows from the standard one-dimensional case merely by observation that $G(a/q,\ell)$ is a d-fold product of these one-dimensional sums.† Moreover, if $\hat{\gamma}_s = \sum_{\ell\in\mathbb{Z}^d/q\mathbb{Z}^d} e^{2\pi i s\cdot\ell/q}\, G(a/q,\ell)$, then,

$$\hat{\gamma}_s = \frac{1}{q^d} \sum_{n\in\mathbb{Z}^d/_{q\mathbb{Z}^d}} \sum_{\ell\in\mathbb{Z}^d/q\,\mathbb{Z}^d} e^{2\pi i s\cdot\ell/q}\, e^{2\pi i(a/q)|n|^2}\, e^{2\pi i n\cdot\ell/q} = e^{2\pi i(a/q)|s|^2}\,.$$

Hence by Proposition 2.2, the norm of the corresponding operator (acting on ℓ^p to itself, $1 \le p \le 2$, with scalar-valued functions) is $O(q^{-(d/2)(2-2/p)}) = O(q^{-d(1-1/p)})$.

Next, the multiplier

$$\sum_{\ell\in\mathbb{Z}^d} \Psi_q\,(\xi-\ell/q)\, d\hat{\sigma}_\lambda(\xi-\ell/q) \tag{3.5}$$

corresponds to a convolution operator from $\ell^p(\mathbb{Z}^d)$ (scalar-valued), to $\ell^p_B(\mathbb{Z}^d)$, where B is the ℓ^∞ space of functions of $\lambda > 0$, for which λ^2 is an integer, and $0 < \lambda^2 \le N$. Notice that $\Phi_q(\xi) = \Phi(q\xi)$ is a bounded multiplier of $L^p(R^d)$ to itself (with norm independent of q). Observe also that $d\hat{\sigma}_\lambda(\xi)$ is a bounded multiplier from $L^p(R^d)$ to $L^p_B(\mathbb{R}^d)$, for $p > \frac{d}{d-1}$, which is a consequence of the spherical maximal theorem in $\mathbb{R}^d$. Finally, note that $m(\xi) = \Phi_q(\xi)\, d\hat{\sigma}_\lambda(\xi)$ is supported in Q/q. Thus, applying the corollary to Proposition 2.1, we see that (3.5) is a bounded multiplier from $\ell^p(\mathbb{Z}^d)$ to $\ell^p_B(\mathbb{Z}^d)$, with norm independent of N (and q). Letting $N \to \infty$, and combining this with the estimate for the first multiplier, we have established conclusion (a) of Proposition 3.1. The second conclusion follows from this because

$$M_\star \le c_d \sum_{1\le q<\infty} \sum_{\substack{(a,q)=1\\ 1\le a\le q}} M_\star^{a/q}\,,$$

†For the one-dimensional estimates, see [W].

so that

$$\Big\|M_\star\Big\|_{\ell^p\to\ell^p} \le A \sum_{1\le q<\infty} q\cdot q^{-d(1-1/p)} < \infty\,,$$

if $1-d(1-1/p)<-1$, i.e. when $p>\frac{d}{d-2}$.

4. Approximations

We now state the assertions which guarantee that M_λ provides an adequate approximation to our operator A_λ. There are two facts; the first is a purely ℓ^2 statement.

Proposition 4.1. *There is a bound A, so that for any $\Lambda>0$,*

$$\| \sup_{\Lambda\le\lambda\le 2\Lambda} |A_\lambda(f) - M_\lambda(f)| \,\|_{\ell^2} \le A\,\Lambda^{2-d/2} \|\, f\,\|_{\ell^2}\,, \quad \text{if } d\ge 5\,. \tag{4.1}$$

The second is a partial result for A_λ which was known previously (see [M]).

Proposition 4.2. *There is a bound A, so that for any $\Lambda>0$*

$$\| \sup_{\Lambda\le\lambda\le 2\Lambda} |A_\lambda(f)\, \|_{\ell^p} \le A\,\|\, f\,\|_{\ell^p}\,, \quad \text{if } d\ge 5\,, \quad p>\frac{d}{d-2}\,. \tag{4.2}$$

Recall that the λ which appear in (4.1) and (4.2) are always restricted to the fact that λ^2 is an integer.

We shall momentarily take these two propositions for granted and see how they, together with Proposition (3.1), prove our main theorem.

Now (4.2) together with Proposition (3.1) yield

$$\| \sup_{\Lambda\le\lambda\le 2\Lambda} (A_\lambda - M_\lambda)f\, \|_{\ell^p} \le A\, \|\, f\,\|_{\ell^p} \qquad \text{for } 2\ge p>\frac{d}{d-2}\,.$$

Interpolating this with (4.1) gives

$$\| \sup_{\Lambda\le\lambda\le 2\Lambda} |(A_\lambda - M_\lambda)(f)|\, \|_{\ell^p} \le A\,\Lambda^{-\varepsilon(p)}\, \|\, f\,\|_{\ell^p}$$

for some $\varepsilon(p)>0$, if $\frac{d}{d-2}<p\le 2$.

Next,

$$\sup_{1\le\lambda<\infty} |(A_\lambda - M_\lambda)\,(f)| \le \sum_{k=0}^{\infty} \sup_{2^k\le\lambda\le 2^{k+1}} |(A_\lambda - M_\lambda)f|\,.$$

Taking the ℓ^p norm we get that

$$\| \sup_{1\le\lambda<\infty} |(A_\lambda - M_\lambda)\,f|\, \|_{\ell^p} \le A'\, \|\, f\,\|_{\ell^p}\,, \text{ for } \frac{d}{d-2} < p \le 2\,,$$

since $\sum_k 2^{-\xi(p)k} < \infty$. Thus, invoking Proposition (3.1) again yields

$$\| \sup_{1\le\lambda<\infty} |A_\lambda(f)| \|_{\ell^p} \le A \| f \|_{\ell^p} \quad \text{when} \quad \frac{d}{d-2} < p \le 2\,.$$

Since the corresponding estimate for $p=\infty$ is trivial, the full range $\frac{d}{d-2} < p \le \infty$ then follows by interpolation, proving the main theorem.

5. The decomposition of A_λ

To prove the crucial approximation property (4.1) we shall decompose the operator A_λ into a sum, each of whose terms corresponds to a fraction a/q, with $1\le q$, $1\le a\le q$, and $(a,q)=1$. It is here we use the ideas of the "circle method" of Hardy, Littlewood, and Ramanujan.

Let us fix $\Lambda>0$, and consider any λ for which $\Lambda\le\lambda\le 2\Lambda$. We shall write $a_\lambda(\xi)$ for the multiplier corresponding to the operator A_λ given by (3.1). We claim that

$$a_\lambda(\xi) = \frac{e^{2\pi\varepsilon\lambda^2}}{\lambda^{d-2}} \sum_{n\in\mathbb{Z}^d} e^{-2\pi\varepsilon|n|^2} e^{2\pi i n\cdot\xi} \int_0^1 e^{2\pi i(|n|^2-\lambda^2)t}\, dt\,.$$

Here ε is positive, but otherwise arbitrary; we will fix it later by setting $\varepsilon = 1/\Lambda^2$. This identity is obvious because $\int_0^1 e^{2\pi i(|n|^2-\lambda^2)t}\,dt = 1$ or 0 according to whether $|n|=\lambda$ or not.

Now we introduce the Θ function

$$\mathcal{F}(z,\xi) = \sum_{n\in\mathbb{Z}^d} e^{-2\pi|n|^2 z} e^{2\pi i n\xi}\,, \tag{5.1}$$

for $\Re(z)>0$, and we make a Farey direction of level $=\Lambda$ of the interval $[0,1]$ of the t integration. That is, for each a/q, $(a,q)=1$ with $1\le a\le q$, and $q\le\Lambda$, we associate the interval $\bar I(a/q) = \left\{t : -\frac{\beta}{q\Lambda} \le t - a/q \le \frac{\alpha}{q\Lambda}\right\}$, where $\alpha=\alpha(a/q,\Lambda)\approx 1$, and $\beta=\beta(a/q,\Lambda)\approx 1$, with α and β chosen appropriately. We denote by $I(a/q)$ the corresponding intervals translated to the origin, $I(a/q) = \left\{\tau : -\frac{\beta}{q\Lambda} \le \tau \le \frac{\alpha}{q\Lambda}\right\}$. Inserting this in the above formula for $a_\lambda(\xi)$ and using identity (5.1) we get

$$a_\lambda(\xi) = \sum_{1\le q\le\Lambda} \sum_{\substack{1\le a\le q\\ (a,q)=1}} a_\lambda^{a/q}(\xi)\,,$$

where

$$a_\lambda^{a/q}(\xi) = \frac{e^{2\pi\varepsilon\lambda^2}}{\lambda^{d-2}} e^{-2\pi i\lambda^2 a/q} \int_{I(a/q)} e^{-2\pi i\lambda^2\tau}\, \mathcal{F}(\varepsilon - i\tau - i\,a/q, \xi)\, d\tau\,. \tag{5.2}$$

Next we use the fundamental identity for the Θ function (5.1). It states that for $\Re(z) > 0$,

$$\mathcal{F}\left(z - i\,\frac{a}{q}, \xi\right) \;=\; \frac{1}{(2z)^{d/2}} \sum_{\ell \in \mathbb{Z}^d} G\left(\frac{a}{q}, \ell\right)\, \exp\left(\frac{-\pi|\xi - \frac{\ell}{q}|^2}{2z}\right). \tag{5.3}$$

Here $G(a/q, \ell)$ is the normalized Gauss sum (3.4). The above is the d-dimensional version of a familiar identity. (For $d = 1$ see, e.g., [SW$_1$, (3.4)]; also [W].) The general case $d \geq 1$ can be proved the same way invoking the Poisson summation formula; alternatively one can observe that (5.3) is merely the d-fold product of the corresponding 1-dimensional identities for each variable $\xi_1, \xi_2, \ldots \xi_d$, separately.

From (5.3) and (5.2) it follows that

$$a_\lambda^{a/q}(\xi) \;=\; e^{-2\pi i \lambda^2 a/q} \sum_{\ell \in \mathbb{Z}^d} G(a/q, \ell)\, J_\lambda\left(a/q, \xi \,-\, \ell/q\right), \tag{5.4}$$

where

$$J_\lambda(a/q, \xi) \;=\; \frac{e^{2\pi\varepsilon\lambda^2}}{\lambda^{d-2}} \int_{I(a/q)} e^{-2\pi i \lambda^2 \tau}\, (2(\varepsilon \,-\, i\tau))^{-d/2}\, e^{\frac{-\pi|\xi|^2}{2(\varepsilon \,-\, i\tau)}}\, d\tau\,. \tag{5.5}$$

6. Approximations, continued

We shall approximate the multipliers $a_\lambda^{a/q}(\xi)$ above by multipliers $b_\lambda^{a/q}(\xi)$ where the cut-off factors $\Phi_q(\xi \,-\, \ell/q)$ have been inserted in (5.4). That is, we define

$$b_\lambda^{a/q}(\xi) \;=\; e^{-2\pi i \lambda^2 a/q} \sum_{\ell \in \mathbb{Z}^d} G(a/q, \ell)\, \Phi_q(\xi \,-\, \ell/q)\; J_\lambda(\xi \,-\, \ell/q)\,. \tag{6.1}$$

Here $\Phi_q(\xi) = \Phi(q\xi)$.

Next we approximate $b_\lambda^{a/q}(\xi)$ by replacing the integral (5.5) that appears in (6.1) by the corresponding integration when taken over the whole real line. So we set

$$c_\lambda^{a/q}(\xi) \;=\; e^{-2\pi i \lambda^2 a/q} \sum_{\ell \in \mathbb{Z}^d} G(a/q, \ell)\; \Phi_q(\xi \,-\, \ell/q)\; I_\lambda(\xi \,-\, \ell/q)\,, \tag{6.2}$$

with

$$I_\lambda(\xi) \;=\; \frac{e^{2\pi\varepsilon\lambda^2}}{\lambda^{d-2}} \int_{-\infty}^{\infty} e^{-2\pi i \lambda^2 \tau}\; (2(\varepsilon \,-\, i\tau))^{-d/2}\; e^{\frac{-\pi|\xi|^2}{2(\varepsilon \,-\, i\tau)}}\, d\tau\,. \tag{6.3}$$

We define the operators $A_\lambda^{a/q}$, $B_\lambda^{a/q}$, $C_\lambda^{a/q}$, as the convolution operators (acting on functions of $\mathbb{Z}^d$), whose Fourier multipliers are respectively, $a_\lambda^{a/q}(\xi)$, $b_\lambda^{a/q}(\xi)$, and $c_\lambda^{a/q}(\xi)$.

PROPOSITION 6.1.

$$\sum_{1\le q\le\Lambda}\ \sum_{\substack{1\le a\le q\\(a,q)=1}} \Big\| \sup_{\Lambda\le\lambda\le 2\Lambda} \left|\left(A_\lambda^{a/q} - B_\lambda^{a/q}\right) f\right| \Big\|_{\ell^2} \le A\,\Lambda^{2-d/2}\,\| f\|_{\ell^2}, \tag{6.4}$$

$$\sum_{1\le q\le\Lambda}\ \sum_{\substack{1\le a\le q\\(a,q)=1}} \Big\| \sup_{\Lambda\le\lambda\le 2\Lambda} \left|\left(B_\lambda^{a/q} - C_\lambda^{a/q}\right) f\right| \Big\|_{\ell^2} \le A\,\Lambda^{2-d/2}\,\| f\|_{\ell^2}\ . \tag{6.5}$$

It is understood that in the above assertions our ε is fixed to be $= 1/\Lambda^2$.

To prove (6.4), let F_τ be the function on $\mathbb{Z}^d$ which is given in terms of its Fourier expansion by $\hat F_\tau(\xi) = \mu(\xi)\,\hat f(\xi)$ where

$$\mu(\xi) = \sum_{\ell\in\mathbb{Z}} (1 - \Phi_q(\xi - \ell/q))\, e^{-\pi|\xi-\ell/q|^2/2(\varepsilon - i\tau)}\,.$$

Note that since each term in the sum is supported where $|\xi - \ell/q| \ge c/q$,

$$\sup_\xi |\mu(\xi)| \le A\,\exp\left(\frac{-c\varepsilon}{q^2(\varepsilon^2+\tau^2)}\right),$$

for some $c > 0$. Thus

$$\| F_\tau \|_{\ell^2} \le A\,\exp\left(\frac{-c\varepsilon}{q^2(\varepsilon^2+\tau^2)}\right) \| f \|_{\ell^2}\ .$$

Now observe that

$$\sup_{\Lambda\le\lambda\le 2\Lambda} \left|\left(A_\lambda^{a/q} - B_\lambda^{a/q}\right) f\right| \le A\,\Lambda^{-d+2}\, q^{-d/2}\int_{I(a/q)} (\varepsilon^2+\tau^2)^{-d/4}\, |F_\tau| d\tau\,,$$

because $G(a/q,\ell) = O(q^{-d/2})$.

As a result,

$$\sup_{\Lambda\le\lambda\le 2\Lambda} \Big\|\left|\left(A_\lambda^{a/q} - B_\lambda^{a/q}\right) f\right|\Big\|_{\ell^2}$$

$$\le A\,\Lambda^{-d+2}\, q^{-d/2}\int_{I(a/q)} (\varepsilon^2+\tau^2)^{-d/4} \exp\left(\frac{-c\varepsilon}{q^2(\varepsilon^2+\tau^2)}\right) d\tau\cdot \| f \|_{\ell^2}\ .$$

Now, because $e^u \le c\,u^{-d/4}$, we get a contribution of $\Lambda^{-d+2}\,\varepsilon^{-d/4}\,|I(a/q)|$ $\| f \|_{\ell^2}$. Taking into account that $\varepsilon = 1/\Lambda^2$, and $\sum\limits_{1\le q\le\Lambda}\sum\limits_{a,q} |I(a/q)| = 1$, we obtain (6.4).

The proof of (6.5) is similar. Notice that we are now integrating over τ in the complement of $I(a/q)$, and thus $|\tau| \ge c/q\Lambda$. We are led in the same way to see that

$$\Big\| \sup_{\Lambda\le\lambda\le 2\Lambda} \left|\left(B_\lambda^{a/q} - C_\lambda^{a/q}\right) f\right| \Big\|_{\ell^2}$$

$$\le A\,\Lambda^{-d+2}\, q^{-d/2}\int_{|\tau|\ge c/q\Lambda} \tau^{-d/2}\, d\tau\cdot \| f \|_{\ell^2}$$

$$\le A\,\Lambda^{-d+2}\, q^{-d/2}\,(q\Lambda)^{d/2-1}\ \| f \|_{\ell^2} = A\,\Lambda^{-d/2+1}\, q^{-1}\ \| f \|_{\ell^2}\ .$$

Now sum in a, then over q, $q \leq \Lambda$. This gives a contribution of

$$O\left(\Lambda^{-d/2+1}\right)\left(\sum_{1\leq q\leq\Lambda}\ \sum_{1\leq a\leq q} 1\right) q^{-1}\ \|\ f\ \|_{\ell^2} =\ O\left(\Lambda^{-d/2+2}\right)\ \|\ f\ \|_{\ell^2},$$

which proves (6.5).

To complete the approximation process (the proof of (4.1)) we now identify $I_\lambda(\xi)$ given by (6.3).

Lemma 6.1.

$$I_\lambda(\xi)\ =\ c_d\, d\hat{\sigma}_\lambda\,(\xi)\,.$$

Taking this temporarily for granted we observe that as a result, $c_d\, M_\lambda^{a/q} = C_\lambda^{a/q}$ (see (3.3) and (6.2), (6.3)). Hence for $\Lambda \leq \lambda \leq 2\Lambda$,

$$\begin{aligned} |(A_\lambda\ - M_\lambda)\ f\,|\ \ &\leq\ \sum_{1\leq q\leq\Lambda}\ \sum_{(a,q)=1}\left|\left(A_\lambda^{a/q}\ - C_\lambda^{a/q}\right)\ f\,\right| \\ &\quad + \sum_{q>\Lambda}\ \sum_{(a,q)=1} c_d\left|M_\lambda^{a/q}\, f\,\right|\,. \end{aligned}$$

However,

$$\left|\,A_\lambda^{a/q}\ - C_\lambda^{a/q}\,\right|\ \leq\ \left|\,A_\lambda^{a/q}\ - B_\lambda^{a/q}\,\right|\ +\ \left|\,B_\lambda^{a/q}\ - C_\lambda^{a/q}\,\right|.$$

Thus, invoking Proposition (6.1), and Proposition (3.1) for $p = 2$, we see that

$$\begin{aligned} &\|\ \sup_{\Lambda\leq\lambda\leq 2\Lambda}\ |\,(A_\lambda\ - M_\lambda)\ (f)\,|\ \|_{\ell^2} \\ &\qquad =\ O\left(\Lambda^{2-d/2}\right)\ \|\,f\|_{\ell_2}\ +\ \left(\sum_{q>\Lambda}\ \sum_{1\leq a\leq q} q^{-d/2}\right)\ \|\,f\|_{\ell^2} \\ &\qquad =\ O\left(\Lambda^{2-d/2}\right)\ \|\,f\,\|_{\ell^2}\ ,\ \text{if}\ \ d\ \geq\ 5\,. \end{aligned}$$

Therefore, Proposition (4.1) is now proved, and with it the proof of our main theorem is complete, save for verification of the lemma above.

7. Proof of Lemma 6.1

The identity

$$\frac{e^{2\pi\varepsilon\lambda^2}}{\lambda^{d-2}}\int_{-\infty}^{\infty} e^{-2\pi i\lambda^2\tau}\ (2(\varepsilon\ - i\tau))^{-d/2}\ e^{\frac{-\pi|\xi|^2}{2(\varepsilon\ - i\tau)}}\ d\tau\ =\ c_d\, d\hat{\sigma}_\lambda(\xi) \tag{7.1}$$

is probably known, but we have not found it in the literature, and so we will give a proof.

First we observe that the left-side of (7.1) is in fact independent of ε, and so we may take $\varepsilon = 1/\lambda^2$. We see that this follows by changing the contour while integrating the function $F(z) = (2z)^{-d/2} e^{2\pi z} e^{-\pi|\xi|^2/2z}$ along lines parallel to the x axis in the upper half-plane. Next, with $\varepsilon = 1/\lambda^2$, and with the change of variables $\lambda^2\tau = t$,

$$I_\lambda(\xi) \;=\; e^{2\pi} \int_{-\infty}^{\infty} e^{-2\pi i t}\, \frac{1}{(2(1\,-\,it))^{d/2}}\; e^{\frac{-\pi\lambda^2|\xi|^2}{2(1\,-\,it)}}\, dt\,.$$

We now insert an extra convergence factor $e^{-\pi\delta t^2}$ in the integral defining I_λ above. Denoting the resulting integral by I_λ^δ we have $I_\lambda^\delta \to I_\lambda^\delta$; moreover if φ is any test function in the Schwartz space, then

$$\int_{\mathbb{R}^d} \widehat{\varphi}(\xi)\, I_\lambda(\xi)\, d\xi \;=\; \lim_{\delta\to 0} \int_{\mathbb{R}^d} \widehat{\varphi}(\xi)\, I_\lambda^\delta(\xi)\, d\xi\,.$$

Also,

$$\int_{\mathbb{R}^d} \widehat{\varphi}(\xi)\, I_\lambda^\delta(\xi)\, d\xi \;=\; \int_{\mathbb{R}^d} \varphi(x)\, \hat{I}_\lambda^\delta(x)\, dx\,. \tag{7.2}$$

Calculating the Fourier transform of the Gaussian $e^{\frac{-\pi|\lambda|^2\xi|^2}{2(i-it)}}$ we see that

$$\hat{I}_\lambda^\delta(x) \;=\; \int_{-\infty}^{\infty} e^{-2\pi i t}\, e^{-\pi\delta t^2}\, e^{-2\pi\frac{|x|^2}{\lambda^2}(1-it)}\, dt\,,$$

which in turn is $e^{-2\pi|x|^2/\lambda^2}\, \delta^{-1/2}\, e^{-\pi(1-|x|^2/\lambda^2)/\delta}$. Inserting this in (7.2), and letting $\delta \to 0$, we obtain

$$\int_{\mathbb{R}^d} \widehat{\varphi}(\xi)\, I_\lambda(\xi)\, d\xi \;=\; c_d \int_{\mathbb{R}^d} \varphi(x)\, d\sigma_\lambda(x)\,,$$

and thus $I_\lambda(\xi) \;=\; c_d\, d\widehat{\sigma}_\lambda(\xi)$, as was to be proved.

Note that

$$c_d \;=\; I_\lambda(0) \;=\; e^{2\pi} \int_{-\infty}^{\infty} e^{-2\pi i t}\, \frac{dt}{(2(1-it))^{d/2}} \;=\; \frac{\pi^{d/2}}{\Gamma(d/2)}\,.$$

8. Counter-examples

Since we shall be dealing with all $d \geq 2$, we return to the original definition of the averages A_λ,

$$A_\lambda(f)(x) \;=\; \frac{1}{N(\lambda)} \cdot \sum_{|m|=\lambda} f(n-m)\,.$$

Let us take f to be the unit mass at the origin; i.e. $f(0) = 1$, and $f(n) = 0$, if $n \in \mathbb{Z}^d$, $n \neq 0$. Then clearly $f \in \ell^p(\mathbb{Z}^d)$, for every p. Next, we observe that

$$A_\lambda(f)(n) \;=\; 1/_{N(\lambda)} \;=\; 1/_{N(|n|)}\,, \;\;\text{if}\;\; |n| \;=\; \lambda\,.$$

Hence,

$$A_\star(f)(n) = \sup_\lambda A_\lambda(f)(n) \geq 1/_{N(|n|)}. \tag{8.1}$$

(Recall that $N(\lambda)$ = number of $m \in \mathbb{Z}^d$, so that $|m| = \lambda$; i.e. $N(\lambda) = r_d(\lambda^2)$.)

Consider now the situation when $d \geq 5$. Then as we have pointed out, $N(\lambda) \approx \lambda^{d-2}$, and so $A_\star(f)(n) \geq c|n|^{-d+2}$. But the latter function belongs to $\ell^p(\mathbb{Z}^d)$ only when $p > \frac{d}{d-2}$, and so the necessity of that condition is proved.

Next assume $d \leq 4$. We shall use the fact that $r_4(2^{2k}) = 24$, for every $k \geq 1$. This follows from the Jacobi formula which states that $r_4(m) = 8 \cdot \sigma_1^*(m)$, where $\sigma_1^*(m)$ is the sum of the divisors of m which are not divisible by 4. (See [HW, Chap. 20].) Thus

$$r_2(2^{2k}) \leq r_3(2^{2k}) \leq r_4(2^{2k}) = 24.$$

Now for each d, $d \leq 4$, we then have $N(\lambda) \leq 24$, if $\lambda = 2^k$. And so for $n \in \mathbb{Z}^d$ with $n = (2^k, 0, \ldots)$, we see that $A_\star(f)(n) \geq 1/24$, by (8.1). Because this happens for infinitely many n, we have $A_\star(f) \notin \ell^p$, for any $p < \infty$, and so the necessity of the condition $d \geq 5$ is established.

University of Wisconsin-Madison, Madison, WI
E-mail addresses: magyar@math.wisc.edu
wainger@math.wisc.edu

Princeton University, Princeton, NJ
E-mail address: stein@math.princeton.edu

References

[BB] P. Bleher and J. Bourgain, Distribution of the error term for the number of lattice points inside a shifted ball, in *Analytic Number Theory*, Vol. 1, 141–153, *Progr. Math.* **138**, Birkhäuser Boston, 1996.

[B1] J. Bourgain, Averages over convex curves and maximal operators, *J. Analyse Math.* **47** (1986), 69–85.

[B2] ———, On the maximal ergodic theorem for certain subsets of the integers, *Israel J. Math.* **61** (1988), 39–72.

[HW] G. H. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, Oxford Univ. Press, New York, 1974.

[M] A. Magyar, L^p-bounds for spherical maximal operators on Z^n, *Rev. Mat. Iberoamericana* **13** (1997), 307–317.

[PP] M. Plancherel and G. Pólya, Fonctions entières et intégrales de Fourier multiple (seconde partie), *Comment. Math. Helv.* **11** (1973), 110–163.

[ShW] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The Univ. of Illinois Press, Urbana, IL (see §19) (1949).

[S] E. M. Stein, Maximal functions I: Spherical means, *Proc. Nat. Acad. Sci.* **73** (1976), 2174–2175.

[SW1] E. M. Stein and S. Wainger, Problems in harmonic analysis related to curvature, *Bull. Amer. Math. Soc.* **84** (1978), 1239–1295.

[SW2] ———, Discrete analogues in harmonic analysis II: fractional integration, *J. d'Analyse* **80** (2000), 335–355.

[W] A. Walfisz, Gitterpunkte in Mehrdimensionalen Kugeln, *Monografie Matematyczne* **33**, Warsaw, 1957.

(Received September 6, 2000)