Diophantine equations and ergodic theorems

Akos Magyar

Abstract

Let (X, μ) be a probability measure space and $T_1, \ldots, T_n$ be a family of commuting, measure preserving invertible transformations on X. Let $Q(m_1, \ldots, m_n)$ be a homogeneous, positive polynomial with integer coefficients, and consider the averages:

$$A_\lambda f(x) = \frac{1}{r_Q(\lambda)} \sum_{Q(m)=\lambda} f(T_1^{m_1}, \ldots, T_n^{m_n} x)$$

where $r_Q(\lambda)$ denotes the number of integer solutions $m = (m_1, \ldots, m_n)$ of the diophantine equation $Q(m) = \lambda$.

We prove that under a certain non-degeneracy condition on the polynomial $Q(m)$ and an ergodic condition on the family of transformations $T = (T_1, \ldots, T_n)$ the pontwise ergodic theorem holds, that is:

$$\lim_{\lambda \to \infty} A_\lambda f(x) = \int_X f \, d\mu$$

for μ a.e. $x \in X$. This means that the solutions sets of the diophantine equation $Q(m) = \lambda$ become uniformly distributed when mapped to the space X via the transformations $T_1, \ldots, T_n$.

The proof uses a variant of the Hardy-Littlewood method of exponential sums developed by Birch and Davenport and techniques from harmonic analysis. A key point is the corresponding maximal theorem, which is a discrete analogue of a maximal theorem on $\Re^n$ corresponding to the level surfaces of the polynomial $Q(x)$.

0. Introduction

A fundamental problem in number theory is to determine asymptotically the number of integer solutions $m = (m_1, \ldots, m_n)$ of a diophantine equation $Q(m_1, \ldots, m_n) = \lambda$ as $\lambda \to \infty$ through the integers, and $Q(m)$ is a positive polynomial with integer coefficients. A general result of this type follows from a variant of the Hardy-Littlewood method of exponential sums developed by Birch [2] and Davenport [4], which is as follows.

Let $Q(m_1, \ldots, m_n)$ be a positive homogeneous polynomial of degree d with integral coefficients, and suppose that it satisfies the non-degeneracy condition

$$(0.1) \qquad n - dim\ V_Q > (d-1)2^d$$

Here $V_Q = \{z \in \mathbf{C}^n \ :\ \partial_1 Q(z) = \ldots \partial_n Q(z) = 0\}$ is the complex singular variety of the polynomial Q. For simplicity we'll refer to polynomials satisfying all the above conditions as non-degenerate forms.

Then the following asymptotic formula holds for the number of integer solutions $r_Q(\lambda) = |\{m \in \mathbf{Z}^n \ :\ Q(m) = \lambda\}|$

$$(0.2) \qquad r_Q(\lambda) = c_Q \lambda^{\frac{n}{d}-1} \sum_{q=1}^{\infty} K(q,0,\lambda) \ +\ O_\delta(\lambda^{\frac{n}{d}-1-\delta})$$

for some $\epsilon > 0$. The expression $K(\lambda) = \sum_{q=1}^{\infty} K(q,0,\lambda)$ is called the singular series, the terms are special cases of $(l=0)$ the exponential sums

$$(0.3) \qquad K(q,l,\lambda) = q^{-n} \sum_{(a,q)=1} \sum_{s \in \mathbf{Z}^n/q\mathbf{Z}} e^{2\pi i \frac{a(Q(s)-\lambda)+s\cdot l}{q}}$$

that is a goes through the reduced residue classes $(mod\ q)$ and s_j goes through all residue classes $(mod\ q)$ for each j. We remark that $K(q,0,\lambda)$ is a Kloostermann sum if $Q(m)$ is a quadratic form.

The asymptotic formula (0.2) can be valid just under a condition of type (0.1). Indeed consider the polynomial $Q(m) = (m_1^2 + \ldots + m_n^2)^{d/2}$ ($d > 2$ even). Then $r_Q(\lambda) = 0$ unless $\lambda = \mu^{d/2}$, $\mu \in \mathbf{N}$, and in that case $r_Q(\lambda) = \mu^{n/2-1} = \lambda^{n/d-2/d}$. Hence formula (0.2) is never valid. The reason is that the complex singular variety: $V_Q = \{z \in \mathbf{C}^n \ :\ z_1^2 + \ldots + z_n^2 = 0\}$ has dimension $n-1$.

It is meaningful only if the singular series is nonzero. It can be shown, that if Q is a non-degenerate form, then there exists an arithmetic progression $\Gamma \subseteq \mathbf{N}$ and a constant $0 < A_Q$ such that

$$(0.4) \qquad A_Q \leq K(\lambda)\ , \quad \text{for every} \ \ \lambda \in \Gamma$$

we'll refer to such sets Γ as *sets of regular values* of the polynomial Q. Inequality (0.4) is true for all large λ, just under additional assumptions modulo primes. Indeed consider the polynomial $Q(m) = m_1^d + pQ_1(m_2, \dots m_n)$. For $\lambda = p\lambda_1 + s$ s being a quadratic non-residue, the equation $Q(m) = \lambda$ has no solution, since d is even. Such conditions will be discussed later.

A crucial observation of the paper is, that a similar approximation formula to (0.2) holds for the Fourier transform of the solution set:

$$\hat{\sigma}_{Q,\lambda}(\xi) = \sum_{m \in \mathbf{Z}^n, Q(m)=\lambda} e^{2\pi i m \cdot \xi} \ , \quad \xi \in \Pi^n$$

Here $\Pi^n = \Re^n / \mathbf{Z}^n$ is the flat torus.

Lemma 1 *Let $Q(m)$ be a non-degenerate form, then there exists $\delta > 0$, s.t.*

$$(0.5) \quad \hat{\sigma}_{Q,\lambda}(\xi) = c_Q \lambda^{\frac{n}{d}-1} \sum_{q=1}^{\infty} K(q,l,\lambda) \sum_{l \in \mathbf{Z}^n} \psi(q\xi - l) d\tilde{\sigma}_Q(\lambda^{\frac{1}{d}}(\xi - s/q))) +$$

$$+ \mathcal{E}_\lambda(\xi) \quad , \quad \textit{and} \quad \sup_\xi |\mathcal{E}_\lambda(\xi)| \le c_\delta \lambda^{\frac{n}{d}-1-\delta}$$

Here $\psi(\xi)$ is a smooth cut-off, $\psi(\xi) = 1$ for $\sup_j |\xi_j| \le 1/8$ and $\psi(\xi) = 0$ for $\sup_j |\xi_j| \ge 1/4$. Moreover

$$(0.6) \qquad d\tilde{\sigma}_Q(\xi) = \int_{\{x \in \Re^n \,:\, Q(x)=1\}} e^{2\pi i x \cdot \xi} \, d\sigma_Q(x)$$

here $d\sigma_Q(x) = \frac{dS_Q(x)}{|Q'(x)|}$, where $dS_Q(x)$ denotes the Euclidean surface area measure of the level surface $Q(x) = 1$, and $|Q'(x)|$ is the magnitude of the gradient of the form Q.

The approximation formula (0.5) means, that the Fourier transform of the indicator function of the solution set $Q(m) = \lambda$ is asymptotically a sum over all rational points, of pieces of the Fourier transform of a surface measure of $Q(x) = \lambda$, multiplied by arithmetic factors and shifted by rationals. This formula in the special case $Q(m) = \sum_j m_j^2$ was proved earlier in [6].

Our main purpose is to study the distribution of the solution sets $\{m \in \mathbf{Z}^n \ : \ Q(m) = \lambda\}$.

Theorem 1 *Let $Q(m)$ be a non-degenerate polynomial and Λ is corresponding set of regular values. Then for a test function $\phi(x) \in S(\Re^n)$ one has*

$$(0.7) \qquad \lim_{\lambda \in \Lambda, \lambda \to \infty} \frac{1}{r_Q(\lambda)} \sum_{Q(m)=\lambda} \phi(\lambda^{-1/d} m) = \int_{Q(x)=1} \phi(x) \, d\sigma_Q(x)$$

That is when the solution sets $Q(m) = \lambda$ are projected to the unit surface $Q(x) = 1$ via the dilations $m \to \lambda^{-1/d} m$, they weakly converge to the surface measure $\frac{dS_Q(x)}{|Q'(x)|}$. This is well-known in case $Q(x)$ is a quadratic form.

The main results of the paper concerns the uniform distribution of the images of the solution sets, when mapped to a measure space via an ergodic family of transformations.

Let (X, μ) be a probability measure space, and $T = (T_1, \ldots, T_n)$ be a family of commuting, measure preserving and invertible transformations. Suppose for every positive integer q the family $T^q = (T_1^q, \ldots, T_n^q)$ is ergodic. We recall this means, that for every $f \in L^2(X, \mu)$

$$T_1^q f = \ldots T_n^q f = f$$

implies $f = constant$. We'll refer to a family of transformations satisfying all the above conditions as a *strongly ergodic family.*

Theorem 2 *Let $Q(m)$ be a non-degenerate form, Γ be a corresponding set of regular values and $T = (T_1, \ldots, T_n)$ a strongly ergodic family of transformations of a measure space (X, μ).*

For $f \in L^2(X, \mu)$ consider the averages

$$A_\lambda f(x) = \frac{1}{r_Q(\lambda)} \sum_{Q(m_1, \ldots m_n) = \lambda} f(T_1^{m_1} T_2^{m_2} \cdots T_n^{m_n} x)$$

Then one has

$$\| \lim_{\lambda \in \Gamma, \lambda \to \infty} (A_\lambda f - \int_X f d\mu) \|_{L^2(X,\mu)} = 0 \tag{0.7}$$

This is an L^2 ergodic theorem, it follows from a non-trivial estimate on the exponential sums $\hat{\sigma}_{Q,\lambda}(\xi)$ at irrational points $\xi \notin \mathbf{Q}^n$. More precisely one needs the following

Lemma 2 *Let $Q(m)$ be a non-degenerate form, Γ be a corresponding set of regular values. Then for $\xi \notin \mathbf{Q}^n$ one has*

$$\lim_{\lambda \in \Lambda, \lambda \to \infty} \frac{1}{r_Q(\lambda)} |\hat{\sigma}_{Q,\lambda}(\xi)| = 0 \tag{0.8}$$

To see the correspondence, suppose that $f \in L^2(x, \mu)$, $f \neq constant$ is a joint eigenfunction of the shifts: $T_j f = e^{2\pi i \xi_j} f$ $(T_j f(x) = f(T_j x))$. Then $A_\lambda f = \frac{1}{r_Q(\lambda)} \hat{\sigma}_{Q,\lambda}(\xi) f$, and the strong ergodicity of the family T implies that $\xi \notin \mathbf{Q}^n$.

The main result of the paper is the corresponding pointwise ergodic

Theorem 3 *Let $Q(m)$ be a non-degenerate form, Γ be a corresponding set of regular values and $T = (T_1, \ldots, T_n)$ a strongly ergodic family of transformations of a measure space (X, μ). Let $f \in L^2(X, \mu)$, Then for μ-almost every $x \in X$ one has*

$$\lim_{\lambda \in \Gamma, \lambda \to \infty} A_\lambda f(x) = \int_X f \, d\mu \tag{0.9}$$

Theorem 3. means, that the images of the solution sets

$$U_\lambda = \{m \in \mathbf{Z}^n \ : \ Q(m) = \lambda\} \tag{0.10}$$

under the transformations $T = (T_1, \ldots, T_n)$:

$$\Omega_{x,\lambda} = \{(T_1^{m_1} T_2^{m_2} \cdots T_n^{m_n} x) \ : \ m \in U_\lambda\} \tag{0.11}$$

become uniformly distributed on X w.r.t. μ for a.e. $x \in X$. Let us mention a special case

Corollary 1 *Let $\alpha_1, \ldots, \alpha_n$ be a set of irrational numbers ($\alpha_j \notin \mathbf{Q}$ $\forall j$). If $Q(m)$ is a non-degenerate form, and Γ is a corresponding set of regular values, then the sets*

$$\Omega_{\lambda,\alpha} = \{(m_1 \alpha_1, \ldots, m_n \alpha_n) \in \Pi^n \ : \ Q(m_1, \ldots, m_n) = \lambda\} \tag{0.12}$$

become uniformly distributed on the torus Π^n w.r.t. the Lebesgue measure.

Indeed, if $X = \Pi^n$ and $T_j(x_1, \ldots, x_j, \ldots x_n) \to (x_1, \ldots, x_j + \alpha_j, \ldots x_n)$ and $\alpha_j \notin \mathbf{Q}$, then the family $T = (T_1, \ldots, T_n)$ is strongly ergodic.

The proof of the pointwise ergodic theorem is based on the L^2 boundedness of a corresponding maximal function

Theorem 4 *Let $Q(m)$ be a non-degenerate form, Γ be a corresponding set of regular values. For $\phi \in l^2(\mathbf{Z}^n)$ we define the maximal function*

$$N^* \phi(m) = \sup_{\lambda \in \Gamma} \frac{1}{r_Q(\lambda)} \Big| \sum_{Q(l)=\lambda} \phi(m-l) \Big| \tag{0.13}$$

Then one has

$$\|N^* \phi\|_{l^2(\mathbf{Z}^n)} \leq C \|\phi\|_{l^2(\mathbf{Z}^n)} \tag{0.14}$$

Theorem 4. is a discrete analogue of a maximal theorem on $\Re^n$, corresponding to the level surfaces of the form $Q(x)$.

Theorem 5 *Let $Q(x)$ be a non-degenerate form and $f \in L^2(\Re^n)$. Then for the maximal function*

$$M^*f(x) = \sup_{\lambda>0} \lambda^{-\frac{n}{d}+1} \left| \int_{Q(y)=\lambda} f(x-y) \frac{dS_{Q,\lambda}(y)}{|Q'(y)|} \right| \tag{0.15}$$

one has

$$\|M^*f\|_{L^2(\Re^n)} \leq C\|f\|_{L^2(\Re^n)} \tag{0.16}$$

For the polynomial $Q(x) = \sum_{j=1}^n x_j^2$ this is the spherical maximal theorem of E.M.Stein [10]. In general, we haven't found this result stated in the literature, nor does it seem to follow easily from the known generalizations of the spherical maximal theorem, see [8], [9]. In fact the proof will use estimates for exponential sums.

Theorem 4. was proved earlier by Magyar, Stein and Wainger [6], in the special case $Q(m) = \sum_{j=1}^n m_j^2$, moreover there the $l^p \to l^p$ boundedness of the discrete maximal operator was shown, for the sharp range of exponents $p > \frac{n}{n-2}$. The non-degeneracy condition (0.1) is also, sharp in the sense, that for the form $Q(m) = m_1^2+m_2^2+m_3^2+m_4^2$ (where *codim* $V_Q = 4 = (d-1)2^d$), Theorem 4. is not true, taking averages on any arithmetic progression Γ, see section 5. below. Hence the present work is the continuation of that paper to some extent.

Also we were motivated by Bourgain's proof of an ergodic theorem, see [3] corresponding to arithmetic subsets of the natural numbers (such as the set of squares), where the Hardy-Littelwood method was used to reduce discrete maximal operators to the corresponding continuous ones.

However in the present case, the averages are over disjoint sets, the *strong ergodicity* condition is also necessary, and is actually a condition on the joint spectrum of the transformations $(T_1, \ldots, T_n)$. Thus we will need the Spectral Theorem even in case of the point-wise convergence, i.e. in the proof of Theorem 3.

1. Exponential sums and oscillatory integrals

We recall some results of Birch [2] on exponential sums, and prove the estimates and properties of oscillatory integrals, needed later. In particular we give a proof of Theorem 5.

Let $Q(m)$ be a non-degenerate form of degree d, that is a positive homogeneous polynomial with integer coefficients, satisfying the non-degeneracy condition (0.1). Let $P > 1$, $0 < \theta \leq 1$ be fixed.

Definition 1 *For $1 \leq q \leq P^{(d-1)\theta}$, $1 \leq a < q$, $(a,q) = 1$ we define the major arcs*

$$(1.1) \qquad L_{a,q}(\theta) = \{\alpha : 2|\alpha - a/q| < q^{-1} P^{-d+(d-1)\theta}\}$$

$$L(\theta) = \bigcup_{q \leq P^{(d-1)\theta},\, (a,q)=1} L_{a,q}(\theta)$$

If $\alpha \notin L(\theta)$ then α belongs to the minor arcs.

The following properties of the major arcs are immediate from the definition, see [2, Sec.4] for the proof.

Proposition 1 *If*

(i) $\theta_1 < \theta_2$ then $L(\theta_1) \subseteq L(\theta_2)$

(ii) $\theta < \frac{d}{3(d-1)}$ then the intervals $L_{a,q}(\theta)$ are disjoint for different values of a and q.

(iii) $\theta < \frac{d}{3(d-1)}$ then $|L(\theta)| \leq P^{-d+3(d-1)\theta}$.

Let $Q_1(m)$ be a polynomial of degree d, such that its d-degree homogeneous part $Q(m)$ is a non-degenerate form.

Throughout the paper we'll use the notation $\kappa = \frac{codim\ V_Q}{2^{d-1}}$, and it is understood that $\frac{\kappa}{d-1} > 2$ which follows from condition (0.1). For a real α, and smooth cut-off function $\phi(x)$, consider the exponential sum

$$(1.2) \qquad S(\alpha) = \sum_{m \in \mathbf{Z}^n} e^{2\pi i \alpha\, Q_1(m)}\, \phi(m/P)$$

This is a Weyl type sum, the trivial estimate is $S(\alpha) \leq P^n$. The following estimates due to Birch [2, Sec.4] are of basic importance

Lemma 3 *Suppose $\alpha \notin L(\theta)$, then for any $\epsilon > 0$, one has*

$$(1.3.1) \qquad |S(\alpha)| \leq C_\epsilon P^{n-\kappa\theta+\epsilon}$$

If $\delta < \frac{\kappa-2(d-1)}{12d(d-1)}$ and $\frac{2\delta\kappa}{d-1} - 2 < \theta < \frac{1}{6d}$ then one has for the average over the minor arcs

$$(1.3.2) \qquad \int_{\alpha \notin L(\theta)} |S(\alpha)| \, d\alpha \leq C_\delta P^{n-d-\delta}$$

The constants C_ϵ, and C_δ depend just on the homogeneous part $Q(m)$, on the cut-off ϕ, on ϵ and δ.

Remark. Estimate (1.3.1) is proved in [4, Lemma 4.3] when the cut-off ϕ is replaced by the characteristic function χ of a cube of side length ≈ 1. Choose χ s.t. $\chi\phi = \phi$ and by Plancherel

$$\sum_{m \in \mathbf{Z}^n} e^{2\pi i \alpha Q_1(m)} \phi(m/P)\chi(m/P) =$$

$$= \int_{\Pi^n} \Big(\sum_{m \in \mathbf{Z}^n} e^{2\pi i \alpha Q_1(m) - m\cdot\xi} \chi(m/P) \Big) \, (P^n \hat{\phi}(P\xi)) \, d\xi$$

Here Π^n is the flat torus and can be identified with $[-1/2, 1/2]^n$. Estimate (1.3) holds for the first term of the integral uniformly in ξ and it is easy to see that $\|P^n \hat{\phi}(P\xi)\|_1 \leq c_\phi$.

To see (1.3.2), one uses (1.3.1) for most most $\alpha \notin L(\theta')$, with $\theta < \theta'$, when it is not valid is a set of small measure by (1.1), giving an improvement in average, see [2, Lemma 4.4]. $\Box$

Corollary 2 *Let $Q(m)$ be a non-degenerate form, and $1 \leq a < q$ be natural numbers s.t. $(a, q) = 1$. Consider the Weyl sum*

$$(1.4) \qquad S(a, q) = \sum_{m \in \mathbf{Z}^n, m_j \ (mod\ q)} e^{2\pi i \frac{a}{q} Q(m)}$$

One has

$$(1.5) \qquad |S(a, q)| \leq c_{Q,\epsilon} q^{n - \frac{\kappa}{d-1} + \epsilon}$$

Proof. Choose $\alpha = a/q$, $P = q$ and notice $\alpha \notin L(\theta)$ for $\theta < \frac{1}{d-1}$. Indeed for $q_1 \leq q^{(d-1)\theta} < q$: $|a/q - a_1/q_1| \geq (qq_1)^{-1} \geq q_1^{-1} q^{-d+(d-1)\theta}$. The estimate follows from (1.3). $\Box$

Corollary 3 *If $|\alpha| < P^{-2d/3}$ then $|S(\alpha)| \leq C_{Q,\epsilon} P^{n+\epsilon} (P^d|\alpha|)^{-\frac{\kappa}{d-1}}$*

Proof. Choose θ s.t. $|\alpha| = P^{-d+(d-1)\theta}$, that is $(P^d|\alpha|)^{\frac{1}{d-1}} = P^\theta$. The major arcs $L_{a,q}(\theta)$ are disjoint since $(d-1)\theta < d/3$, moreover α is an endpoint of the interval $L_{0,1}(\theta)$ hence $\alpha \notin L_{a,q}(\theta - \epsilon)$ for every $\epsilon > 0$. By (1.3)

$$|S(\alpha)| \leq C_{Q,\epsilon} P^{n-\kappa\theta+\epsilon} = C_{Q,\epsilon} P^{n+\epsilon} (P^d|\alpha|)^{-\frac{\kappa}{d-1}} \quad \square$$

The above corollaries can be found in [2, Sec.4-5], however they quickly follow from Lemma 3., hence we've included their proofs.

Let $Q(x)$ be a non-degenerate form of degree d, $\kappa = \frac{codim\, V_Q}{2^{d-1}}$, $L > 0$, and $\eta \in \Re^n$.

Lemma 4 *Consider the oscillatory integral*

$$I_Q(L, \eta) = \int e^{2\pi i (LQ(x) + x\cdot\eta)} \phi(x)\, dx \tag{1.6}$$

One has for every $\epsilon > 0$

$$I_Q(L, \eta) \leq C_{Q,\epsilon} (1 + L)^{-\frac{\kappa}{d-1}+\epsilon} \tag{1.7}$$

where the constant C_ϵ is independent of L and η.

Proof. The estimate is obvious for $L < 1$. Let $L \geq 1$, the gradient of the phase: $|LQ'(x) + \eta| \geq L$ if $|\eta| \geq CL$ on the support of $\phi(x)$ for large enough constant $C > 0$, and (1.7) follows by partial integration.

Suppose $|\eta| \leq CL$ and introduce the parameters P, θ, α s.t. $\alpha = P^{-d}L$, $L = P^{(d-1)\theta}$ and $P > L^{\frac{3\kappa}{d-1}}$. Changing variables $y = Px$ one has

$$I_Q(L, \eta) = P^{-n} \int e^{2\pi i \alpha\, (Q(y) + P^{d-1} y\cdot\eta)} \phi(y/P)\, dy$$

We compare the integral to a corresponding exponential sum

$$P^{-n} S(\alpha) = P^{-n} \sum_{m \in \mathbf{Z}^n} e^{2\pi i \alpha\, (Q(m) + P^{d-1} m\cdot\eta)} \phi(m/P)$$

If $y = m + z$ where $m \in \mathbf{Z}^n$ and $z \in [0, 1]^n$, then

$$|e^{2\pi i \alpha\, (Q(y) + P^{d-1} y\cdot\eta)} - e^{2\pi i \alpha\, (Q(m) + P^{d-1} m\cdot\eta)}| \leq$$

$$\leq C|\alpha|(|Q(m + z) - Q(m)| + P^{d-1}|\eta|) \leq CP^{-1+(d-1)\theta}$$

since $|\alpha| = P^{-d+(d-1)\theta}$ and $|\eta| \leq P^{(d-1)\theta}$.

Thus $|I_Q(L,\eta) - P^{-n}S(\alpha)| \leq C_Q P^{-1+2(d-1)\theta} \leq C_Q P^{-\frac{1}{3}}$. Corollary 3. implies that

$$|P^{-n}S(\alpha)| \leq C_\epsilon (P^d\alpha)^{-\frac{\kappa}{d-1}+\epsilon} C_\epsilon L^{-\frac{\kappa}{d-1}+\epsilon}$$

and (1.7) follows using $P^{-\frac{1}{3}} \leq L^{-\frac{\kappa}{d-1}}$. $\square$

Remarks.

i) It is proved in [2, Sec.4] in case $\eta = 0$, we used a modification of the argument given there.

ii) The proof is based on estimate (1.3), which uses the fact that the polynomial $Q(x)$ has integer coefficients. Does (1.7) remain true assuming the coefficients are real ?

iii) In case $V_Q = \{0\}$, and $\eta = 0$ the integral decays as $(1+L)^{-\frac{n}{d}}$. What is the true decay which holds uniformly in η, in this case ?

The level surfaces of a non-degenerate form $S_{Q,\lambda} = \{x \in \Re^n : Q(x) = \lambda\}$ are compact smooth hypersurfaces (for $\lambda > 0$). Indeed $Q(x) = \lambda$ implies that $|x| \approx \lambda^{1/d}$, moreover $Q'(x) \neq 0$ for every $x \neq 0$.

There is a unique $n-1$-form $d\sigma_Q(x)$ on $\Re^n - 0$ for which

$$dQ \wedge d\sigma_Q = dx_1 \wedge \ldots \wedge dx_n \tag{1.8}$$

called the Gelfand-Leray form, see [1, Sec.7.1]. To see this, suppose that $\partial_1 Q(x) \neq 0$ on some open set U. By a change of coordinates: $y_1 = \partial_1 Q(x), y_j = x_j$ for $j \geq 2$, equation (1.8) takes the form

$$dy_1 \wedge d\sigma_Q(y) = \partial_1 H(y)\ dy_1 \wedge \ldots \wedge dy_n \tag{1.9}$$

where $x_1 = H(y), x_j = y_j$ is the inverse map. Thus the form: $d\sigma_Q(y) = \partial_1 H(y)\, dy_2 \bigwedge \ldots \bigwedge dy_n$ satisfies equation (1.8).

We define the measure $d\sigma_{Q,\lambda}$ as the restriction of the $n-1$ form $d\sigma_Q$ to the level surface $S_{Q,\lambda}$. This measure is absolutely continuous w.r.t. the Euclidean surface are measure $dS_{Q,\lambda}$, more precisely one has

Proposition 2 .

$$d\sigma_{Q,\lambda}(x) = \frac{dS_{Q,\lambda}(x)}{|Q'(x)|} \tag{1.10}$$

Proof. Choose local coordinates y as before, in coordinates y level surface and surface area measure takes the form:

$$S_{Q,\lambda} = \{x_1 = H(\lambda, y_2, \ldots, y_n) : x_j = y_j\}$$

and

$$dS_{Q,\lambda}(y) = (1 + \sum_{j=2}^{n} \partial_j^2 H(\lambda, y))^{1/2}\ dy_2 \wedge \ldots \wedge dy_n$$

Using the identity $F(H(y), y_2, \ldots, y_n) = y_1$ one has

$$\partial_1 F(x) \partial_1 H(y) = 1\ ,\quad \partial_1 F(x) \partial_j H(y) + \partial_j F(x) = 0$$

This implies that $\partial_1 H(y) = (1 + \sum_{j=2}^{n} \partial_j^2 H(y))^{1/2} \cdot |F'(x)|^{-1}$. Then (1.10) follows by taking $y_1 = \lambda$. $\Box$

A key observation of the paper is that the measure $d\sigma_{Q,\lambda}$, considered as a distribution on $\Re^n$, has a simple oscillatory integral representation

Lemma 5 *Let $Q(x)$ be a non-degenerate form and $\lambda > 0$. Then in the sense of distributions*

$$d\sigma_{Q,\lambda}(x) = \int_{\Re} e^{2\pi i (Q(x)-\lambda)t}\, dt \tag{1.11}$$

This means that for any smooth cut-off function $\chi(t)$ and test function $\phi(x)$ one has

$$\lim_{\epsilon \to 0} \int\int e^{2\pi i (Q(x)-\lambda)t} \chi(\epsilon t) \phi(x)\, dx dt\ =\ \int \phi(x) d\sigma_{Q,\lambda}(x) \tag{1.12}$$

Proof. Let U be an open set on which $\partial_1 Q \neq 0$, and by a partition of unity we can suppose, that $supp\, \phi \subseteq U$. Changing variables $y_1 = Q(x), y_j = x_j$ the left side of (1.12) becomes

$$\lim_{\epsilon \to 0} \int\int e^{2\pi i (y_1-\lambda)t} \chi(\epsilon t) \tilde{\phi}(y) |\partial_1 H(y)|\, dy dt = \int \tilde{\phi}(\lambda, y') |\partial_1 H(\lambda, y')| dy'$$

where $y' = (y_2, \ldots y_n)$.

The last equality can be seen by integrating in t and in y_1 first, and using the Fourier inversion formula:

$$\lim_{\epsilon \to 0} \int\int e^{2\pi i (y_1-\lambda)t} \chi(\epsilon t) g(y_1)\, dy_1 dt = g(\lambda)$$

On the other hand $S_{Q,\lambda} \cap U = \{x_1 = H(\lambda, y_2, \ldots y_n) : x_j = y_j\}$ and $d\sigma_{Q,\lambda}(y) = |\partial_1 H(\lambda, y')|\, dy'$ in parameters y'. $\Box$.

Lemma 6 *Let $Q(x)$ be a non-degenerate form of degree d, $\kappa = \frac{codim\, V_Q}{2^{d-1}}$. Then one has for the Fourier transform of the measure $d\sigma_{Q,1} = d\sigma_Q$*

$$|d\tilde{\sigma}_Q(\xi)| \leq C_{Q,\epsilon}(1 + |\xi|)^{-\frac{\kappa}{d-1}+1+\epsilon} \tag{1.13}$$

Proof. Suppose $|\xi| > 1$. Using the fact that $\phi d\sigma_Q = d\sigma_Q$ if $\phi = 1$ on a neighborhood of 0 and formula (1.12), we have

$$d\tilde{\sigma}_Q(\xi) = \int e^{-2\pi i\, x\cdot\xi}\phi(x)\, dx = \tag{1.14}$$

$$= \lim_{\delta\to 0}\int\int e^{-2\pi i\, x\cdot\xi}e^{2\pi i(Q(x)-1)t}\phi(x)\chi(\delta t)\, dxdt$$

We decompose the range of integration into two parts

$$\tilde{d\sigma}_Q(\xi) = \int_{|t|\geq C|\xi|}\int_{\Re^n} \qquad + \int_{|t|\leq C|\xi|}\int_{\Re^n} \qquad = I_1 + I_2$$

Since for fixed $|t| \leq C|\xi|$ the gradient of the phase: $|tQ'(x) - \xi| \geq |\xi|/2$ if $C > 0$ is small enough, integration by parts gives $|I_2| \leq C_N\,(1 + |\xi|)^{-N+1}$ for every $N > 0$.

For $|t| \geq C|\xi|$ Lemma 3. implies

$$|\int e^{2\pi i(tQ(x)-x\cdot\xi)}\phi(x)\, dx| \leq C_\epsilon|\,|t|^{-\frac{\kappa}{d-1}+\epsilon} \quad \text{hence}$$

$$I_1 \leq C_\epsilon\int_{|t|\geq C|\xi|}|t|^{-\frac{K}{d-1}+\epsilon}\, dt \leq C_\epsilon|\xi|^{-\frac{K}{d-1}+1+\epsilon}$$

□

First we prove a dyadic version of Theorem 5., together with a refinement which will be needed in the proof of Theorem 3.

Lemma 7 *Let $\Lambda > 0$ be fixed, $\omega(\xi)$ be a smooth function with supported on the set $\{\Lambda^{-\frac{1}{2d}} \leq \|\xi\| \leq \frac{1}{4}\}$, where $\|\xi\| = max_j|\xi_j|$.*

Let M_λ and $M_{\omega,\lambda}$ be the multipliers acting on $L^2(\Re^n)$ defined by

$$\widetilde{M_\lambda f}(\xi) = d\tilde{\sigma}(\lambda^{1/d}\xi) \quad \text{and} \quad \widetilde{M_{\omega,\lambda} f}(\xi) = \omega(\xi)d\tilde{\sigma}(\lambda^{1/d}\xi$$

Then one has for the maximal operators

$$\|\sup_{\Lambda\leq\lambda<2\Lambda}|M_\lambda f|\,\|_{L^2} \leq C\|f\|_{L^2} \tag{1.15}$$

$$\|\sup_{\Lambda\leq\lambda<2\Lambda}|M_{\omega,\lambda} f|\,\|_{L^2} \leq C\Lambda^{-\frac{1}{2d}}\|f\|_{L^2} \tag{1.16}$$

Note that $M_\lambda f = \lambda^{-\frac{n}{d}+1}\,(f * d\sigma_\lambda)$.

Proof. Using the integral representation (1.11) one has

$$d\tilde{\sigma}(\lambda^{1/d}\xi) = \lambda^{-\frac{n}{d}+1}[d\tilde{\sigma}_\lambda(\xi) =$$

$$= \lambda^{-\frac{n}{d}+1}\int_{\Re}\int_{\Re^n} e^{2\pi i(Q(x)-\lambda)t+m\cdot\xi}\phi(x/\Lambda^{\frac{1}{d}})\,dx\,dt$$

This means

$$M_\lambda f = \lambda^{-\frac{n}{d}+1}\int e^{-2\pi i\lambda t}H_{\Lambda,t}f\,dt$$

where $H_{\Lambda,t}$ is the multiplier corresponding to

$$h_{\Lambda,t}(\xi) = \int e^{2\pi i(Q(x)t+m\cdot\xi)}\phi(x/\Lambda^{\frac{1}{d}})\,dx$$

Then taking the absolute values, and using Minkowski's integral inequality

$$(1.16)\qquad \|\sup_{\Lambda\le\lambda<2\Lambda}|M_\lambda f|\,\|_{L^2} \le C\Lambda^{-\frac{n}{d}+1}\int \|H_{\Lambda,t}f\|_{L^2}\,dt$$

Using again the estimates (see Lemma 6.)

$$|h_{\Lambda,t}(\xi)| \le C\Lambda^{\frac{n}{d}}\min\{(1+\Lambda|\xi|)^{-N},(1+\Lambda|t|)^{-2}\}$$

, (where we used that $-\frac{\kappa}{d-1}+\epsilon < -2$), (1.14) follows from (1.16) because $\Lambda\int(1+\Lambda t)^{-2}\,dt \le C$.

To prove (1.15) we have to replace $h_{\Lambda,t}(\xi)$ by $\omega(\xi)h_{\Lambda,t}(\xi)$. Then one can give better uniform estimates in ξ, indeed for $\Lambda t \le \Lambda^{\frac{1}{2d}}$ it follows

$$|\omega(\xi)h_{\Lambda,t}(\xi)| \le C(1+\Lambda|\xi|)^{-N} \le (1+\Lambda^{\frac{1}{2d}})^{-N} \quad \text{hence}$$

$$\Lambda\int\sup_\xi|\omega(\xi)h_{\Lambda,t}(\xi)|\,dt \le C\Lambda\int_{\Lambda t\le\Lambda^{\frac{1}{2d}}}\Lambda^{-\frac{N}{2d}}dt +$$

$$C\Lambda\int_{\Lambda t\ge\Lambda^{\frac{1}{2d}}}(\Lambda t)^{-2}dt \le\ C\Lambda^{-\frac{1}{2d}}$$

This proves (1.15). $\square$

Proof of Theorem 5. If $Q(x)$ is a non-degenerate form of degree d, then the maximal function: $\bar{M}f(x) = \sup_{\lambda>0} \lambda^{-n/d}|\bar{A}_\lambda f(x)|$, where

$$\bar{A}_\lambda f(x) = \int_{Q(y)\leq\lambda} f(x-y)\,dy$$

is majorized by the standard Hardy-Littlewood maximal function, hence is bounded from $L^2(\Re^n)$ to itself.

Formula (1.8) means, that for a test function $g(y)$

$$\int_{Q(y)\leq\lambda} g(y)\,dy = \int_0^\lambda \int_{Q(y)=s} g(y)\,d\sigma_{Q,s}(y)ds$$

hence

$$\bar{A}_\lambda f(x) = \lambda^{-1}\int_0^\lambda Af(x)\,ds$$

Then the theorem follows by the standard argument of the spherical maximal theorem, see [10]. $\square$

2. The approximation formula

First we rewrite formula (0.5) in the form

$$\hat{\sigma}_{Q,\lambda}(\xi) = c_Q \sum_{q=1}^{\infty} \sum_{(a,q)=1} m_\lambda^{a/q}(\xi) + \mathcal{E}_\lambda(\xi) \tag{2.1}$$

where

$$m_\lambda^{a/q}(\xi) = \sum_{l\in\mathbf{Z}^n} e^{-2\pi i\lambda a/q} G(a/q,l)\,\psi(q\xi-l)d\tilde{\sigma}_{Q,\lambda}(\xi-l/q) \tag{2.2}$$

and $$G(a/q,l) = q^{-n}\sum_{s\in\mathbf{Z}^n/q\mathbf{Z}^n} e^{2\pi i\frac{a(Q(s)-\lambda)+s\cdot l}{q}}$$

Here we used the fact, that $d\tilde{\sigma}_{Q,\lambda}(\eta) = \lambda^{n/d-1}d\tilde{\sigma}_Q(\lambda^{1/d}\eta)$, which follows by scaling, since $|Q'(x)|$ is homogeneous of degree $d-1$.

Note that in the right side of (2.1) there is at most one nonzero term, since the cut-off factor $\psi(q\xi-l)$, and then (1.4) implies

$$|m_\lambda^{a/q}(\xi)| \leq C_\epsilon\lambda^{n/d-1}q^{-\frac{\kappa}{d-1}+\epsilon} \leq C_\epsilon\lambda^{n/d-1}q^{-2-\epsilon} \tag{2.3}$$

by (0.1) if ϵ is small enough, hence the sum in (2.1) is absolutely convergent.

Let N_λ and M_λ denote the convolution operators on $\mathbf{Z}^n$ corresponding to the multipliers $\hat{\sigma}_{Q,\lambda}(\xi)$ and $m_\lambda(\xi) = \sum_q\sum_{(a,q)=1} m_\lambda^{a/q}(\xi)$. The main approximation property we need is the following

Lemma 8 *Let $\Lambda > 0$, $\delta > 0$ be amall, fixed and $f \in l^2(\mathbf{Z}^n)$ then*

$$(2.4) \qquad \| \sup_{\Lambda \leq \lambda < 2\Lambda} |(N_\lambda - M_\lambda) f| \, \|_{l^2} \leq C_\delta \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2}$$

Lemma 5. in the special case $Q(m) = \sum_j m_j^2$ is proved in [6, Prop. 4.1], and the same argument works in the present case, after the preparations made in Section 1.

Also Lemma 1. follows immediately from Lemma 5., since for fixed λ $(\Lambda \leq \lambda < 2\Lambda)$

$$\|(N_\lambda - M_\lambda) f \, \|_{l^2} \leq C \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2} \quad \forall f \in l^2(\mathbf{Z}^n)$$

is equivalent to

$$\sup_\xi |\hat{\sigma}_{Q,\lambda}(\xi) - m_\lambda(\xi)| \leq C \lambda^{\frac{n}{d}-1-\delta}$$

which is the content of (0.5).

Let $P = \Lambda^{1/d}$, and let $\phi(x)$ be smooth cut-off function on $\Re^n$ s.t. $\phi(x) = 1$ for $Q(x) \leq 2$. Then

$$\hat{\sigma}_{Q,\lambda}(\xi) = \sum_{m \in \mathbf{Z}^n} e^{2\pi i m \cdot \xi} \phi(m/P) \int_0^1 e^{2\pi \alpha i (Q(m) - \lambda)} \, d\alpha =$$

$$= \int_0^1 S(\alpha, \xi) e^{-2\pi i \lambda \alpha} \, d\alpha$$

where $S(\alpha, \xi) = \sum_m e^{2\pi i (\alpha Q(m) + m \cdot \xi)} \phi(m/P)$.

Let δ and θ be chosen as in Lemma 3. and integrate separately on the major and minor arcs:

$$(2.5) \qquad \hat{\sigma}_{Q,\lambda}(\xi) = \int_{\alpha \in L(\theta)} S(\alpha, \xi) e^{-2\pi i \lambda \alpha} \, d\alpha + \int_{\alpha \notin L(\theta)} S(\alpha, \xi) e^{-2\pi i \lambda \alpha} \, d\alpha$$

$$= a_\lambda(\xi) + \mathcal{E}^1_\lambda(\xi)$$

The following proposition is a prototype of the error estimates in this section

Proposition 3 *Let $\mathcal{E}^1_\lambda$ be the multiplier corresponding to $\mathcal{E}^1_\lambda(\xi)$ that is: $\widehat{\mathcal{E}^1_\lambda f} = \mathcal{E}^1_\lambda(\xi) \hat{f}(\xi)$. Then one has*

$$(2.6) \qquad \| \sup_{\Lambda \leq \lambda < 2\Lambda} |\mathcal{E}^1_\lambda f| \|_{l^2} \leq C_{Q,\delta} \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2}$$

Proof. Let S_α be defined by $\widehat{S_\alpha f} = S(\alpha,\xi)\hat{f}(\xi)$, then

$$\sup_{\Lambda\le\lambda<2\Lambda} |\mathcal{E}^1_\lambda f| \le \int_{\alpha\notin L(\theta)} |S_\alpha f|\, d\alpha$$

Taking the l^2 norm one gets (2.6) from the minor arc estimate (1.3.2)

$$\int_{\alpha\notin L(\theta)} |S_\alpha(x,\xi)| \le C_\delta \Lambda^{n/d-1-\delta} \quad \square$$

Suppose $\alpha \in L_{a,q}(\theta)$ for some $(a,q)=1$, $q \le P^{(d-1)\theta}$, and write $\alpha = a/q+\beta$, $|\beta| \le P^{-d+(d-1)\theta}$, $m = qm_1+s$. We have

$$S(\alpha,\xi) = \sum_{s\in\mathbf{Z}^n/q\mathbf{Z}^n} e^{2\pi i\frac{aQ(s)}{q}} \sum_{m_1\in\mathbf{Z}^n} e^{2\pi i(\beta Q(qm_1+s)+(qm_1+s)\cdot\xi)}\phi(\frac{qm_1+s}{P})$$

Let $H(x,\beta) = e^{2\pi i\beta Q(m)}\phi(m/P)$, applying Poisson summation for the inner sum

$$\sum_{m_1} H(qm_1+s)e^{2\pi i(qm_1+s)\cdot\xi} = q^{-n}\sum_l e^{2\pi i\frac{l\cdot s}{q}}\tilde{H}(\xi-l/q,\beta)$$

Integrating in β and summing in a,q, one has

$$a_\lambda(\xi) = \sum_{q\le P^{(d-1)\theta}} \sum_{(a,q)=1} a_\lambda^{a/q}(\xi) \tag{2.7}$$

where

$$a_\lambda^{a/q}(\xi) = \sum_{l\in\mathbf{Z}^n} G(a,l,q)J_\lambda(\xi-l/q) \tag{2.8}$$

and

$$J_\lambda(\xi-l/q) = \int_{|\beta|\le P^{-d+(d-1)\theta}} \tilde{H}(\xi-l/q,\beta)e^{-2\pi i\lambda\beta}\, d\beta \tag{2.9}$$

We shall approximate the multipliers $a_\lambda^{a/q}(\xi)$ by multipliers $b_\lambda^{a/q}(\xi)$ where the cut-off function $\psi(q\xi-l)$have been inserted in (2.8), that is let

$$b_\lambda^{a/q}(\xi) = \sum_{l\in\mathbf{Z}^n} G(a,l,q)\psi(q\xi-l)J_\lambda(\xi-l/q) \tag{2.10}$$

Next we extend the integration in β in (2.9) and define

$$c_\lambda^{a/q}(\xi) = \sum_{l \in \mathbf{Z}^n} G(a,l,q)\psi(q\xi - l) I_\lambda(\xi - l/q) \tag{2.11}$$

with

$$I_\lambda(\xi - l/q) = \int_{\Re} \tilde{H}(\xi - l/q, \beta) e^{-2\pi i \lambda \beta}\, d\beta \tag{2.12}$$

Note that the integral in (2.12) is absolute convergent. Indeed by (1.7) and (0.1)

$$|\hat{H}(\eta,\beta)| \leq C_{Q,\epsilon} P^n (1 + P^d|\beta|)^{-\frac{K}{d-1}+\epsilon} \tag{2.13}$$

A crucial point is to identify the the integrals $I_\lambda(\eta)$:

$$I_\lambda(\eta) = \int_{\Re^n} \int_{\Re} e^{-2\pi i (Q(x)-\lambda)\beta} e^{2\pi i x\cdot\eta} \phi(x/P)\, d\beta\, d\eta = \tag{2.14}$$

$$= \int_{\Re^n} d\sigma_{Q,\lambda}(x) e^{2\pi i x\cdot\eta} \phi(x/P)\, d\eta = d\tilde{\sigma}_{Q,\lambda}(\eta)$$

by (1.11). This means that $c_\lambda^{a/q}(\xi) = m_\lambda^{a/q}(\xi)$.

Let $A_\lambda^{a/q}$, $B_\lambda^{a/q}$, $M_\lambda^{a/q}$ denote the multipliers, corresponding to $a_\lambda^{a/q}(\xi)$, $b_\lambda^{a/q}(\xi)$, and $m_\lambda^{a/q}(\xi)$.

Proposition 4 .

$$\sum_{q \leq P^{(d-1)\theta}} \sum_{(a,q)=1} \| \sup_{\Lambda \leq \lambda < 2\Lambda} |(A_\lambda^{a/q} - B_\lambda^{a/q}) f| \, \|_{l^2} \leq C_\delta \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2} \tag{2.15}$$

$$\sum_{q \leq P^{(d-1)\theta}} \sum_{(a,q)=1} \| \sup_{\Lambda \leq \lambda < 2\Lambda} |(B_\lambda^{a/q} - M_\lambda^{a/q}) f| \, \|_{l^2} \leq C_\delta \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2} \tag{2.16}$$

$$\sum_{q \geq P^{(d-1)\theta}} \sum_{(a,q)=1} \| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda^{a/q}) f| \, \|_{l^2} \leq C_\delta \Lambda^{\frac{n}{d}-1-\delta} \|f\|_{l^2} \tag{2.17}$$

Proof. Note that each of the operators $A_\lambda^{a/q}$, $B_\lambda^{a/q}$, $M_\lambda^{a/q}$ are of the form

$$T_\lambda f = \int_I e^{-2\pi i \lambda \beta} U_\beta f \, d\beta$$

where U_β is some convolution operator acting on functions on $\mathbf{Z}^n$: $\widehat{U_\beta f} = \mu_\beta(\xi)\hat{f}(\xi)$, and I is some interval. Then one has the point-wise estimate

$$\sup_{\Lambda \le \lambda < 2\Lambda} |T_\lambda f| \le \int_I |U_\beta f|\, d\beta$$

and taking the l^2 norm

$$\| \sup_{\Lambda \le \lambda < 2\Lambda} |T_\lambda f| \,\|_{l^2} \le \int_I | \sup_\xi | \left(\mu_\beta(\xi)|\right) d\beta \cdot \|f\|_{l^2}$$

For the operator $A_\lambda^{a/q} - B_\lambda^{a/q}$, we have

$$\mu_\beta(\xi) = \sum_{l \in \mathbf{Z}^n} G(a,l,q)(1 - \psi(q\xi - l))\hat{H}(\xi - l/q, \beta) = \mu(\xi)\hat{f}(\xi)$$

and $I = \{|\beta| \le P^{-d+(d-1)\theta}$

Let $\eta = \xi - l/q$ and estimate $\hat{H}(\eta,\beta)$ by partial integration:

$$|\hat{H}(\eta,\beta)| = P^n | \int (e^{2\pi i P^d \beta Q(x)} \phi(x)) e^{2\pi i P x \cdot \eta}\, dx| \le$$

$$C_N P^n |P\eta|^{-N} \int |(d/d\eta)^N (e^{2\pi i P^d \beta Q(x)} \phi(x))|\, dx \le$$

$$\le C_N P^n |P\eta|^{-N} (1 + P^d|\beta|)^N$$

Using the facts that $|P\eta| = P/q|(q\xi - l)| \ge cP^{1-(d-1)\theta}(1 + |(q\xi - l)|)$ on the support of $1 - \psi(q\xi - l)$ (for small $c > 0$), $(d-1)\theta \le 1/3$ and $|G(a,l,q)| \le 1$, one has for $|\beta| \le P^{-d+(d-1)\theta}$

$$|\sup_\xi \mu_\beta(\xi)| \le C_N P^n P^{-N(1-2(d-1)\theta)} \sum_{l \in \mathbf{Z}^n} (1 + |q\xi - l|)^{-N} \le C_N P^{n-N/3}$$

Then choosing N large enough, (2.15) follows since the total length of integration for different values of a a and q is at most 1.

For the operator $B_\lambda^{a/q} - M_\lambda^{a/q}$, we have

$$\mu_\beta(\xi) = \sum_{l \in \mathbf{Z}^n} G(a,l,q)\psi(q\xi - l)\hat{H}(\xi - l/q, \beta) = \mu(\xi)\hat{f}(\xi)$$

but we are integrating now on $|\beta| \geq P^{-d+(d-1)\theta}$. Note that $\psi(q\xi - l) \neq 0$ for at most one values of l, estimate (2.13) and (1.11) : $|G(a,l,q)| \leq Cq^{-2-\epsilon}$. Then

$$|\sup_\xi \mu_\beta(\xi)| \leq C_N P^n (1 + P^d|\beta|)^{-\frac{K}{d-1}+\epsilon}$$

hence by changing variables $\beta_1 = P^d\beta$ one has

$$\| \sup_{\Lambda\leq\lambda<2\Lambda} |(B_\lambda^{a/q} - M_\lambda^{a/q})f| \, \|_{l^2} \leq C_\epsilon P^{n-d} q^{\frac{K}{d-1}} \int_{|\beta_1} |geqP^{(d-1)\theta}|\beta_1|^{-2} d\beta \cdot \|f\|_{l^2} \leq$$

$$\leq C_\epsilon q^{-K/2} P^{n-d-\delta}$$

Summing in $a \leq q$ and in $q = 1$ to ∞ proves (2.16).

For $M_\lambda^{a/q}$ the multiplier $\mu_\beta(\xi)$ is the same, but now the range of integration is the whole real line. Thus

$$\| \sup_{\Lambda\leq\lambda<2\Lambda} |M_\lambda^{a/q} f| \, \|_{l^2} \leq C_\epsilon q^{-2} \int_{\beta\in\Re} (1 + P^d|\beta|)^{-2} d\beta \cdot \|f\|_{l^2} \leq$$

$$\leq C_\epsilon q^{-2} P^{n-d} \cdot \|f\|_{l^2}$$

Summing for $a \leq q$ and $q \geq P^{(d-1)\theta}$ one gets the estimate $P^{n-d-(d-1)\theta} \leq P^{n-d-\delta}$. $\square$

Lemma 8. immediately follows from the above said, indeed for fixed λ

$$|(N_\lambda - M_\lambda)f| \leq \sum_{q\leq P^{(d-1)\theta}} \sum_{(a,q)=1} |(A_\lambda^{a/q} - M_\lambda^{a/q})f| +$$

$$+ \sum_{q\geq P^{(d-1)\theta}} |M_\lambda^{a/q} f| + |\mathcal{E}_\lambda^1 f|$$

We will need the following "dyadic" discrete maximal theorem, (proved in [4] in case $Q(m) = \sum_j m_j^2$)

Proposition 5 *Let $\Lambda > 0$ be fixed, then for the operator:*

$$N_\lambda f(m) = \sum_{Q(l)=\lambda} f(m - l)$$

one has

$$\| \sup_{\Lambda\leq\lambda<2\Lambda} |N_\lambda f| \, \|_{l^2} \leq C\Lambda^{\frac{n}{d}-1} \|f\|_{l^2} \tag{2.18}$$

where the constant C is independent of Λ.

Proof. Note that $\widehat{N_\lambda f}(\xi) = \hat{\sigma}_{Q,\lambda}(\xi)\hat{f}(\xi)$ hence

$$N_\lambda f = \sum_{q,a} M_\lambda^{a/q} f + \sum_{q,a} (A_\lambda^{a/q} - M_\lambda^{a/q}) f + \mathcal{E}_\lambda^1 f$$

By Proposition 4. it is enough to show

$$\sum_{q,a} \| \sup_{\Lambda \le \lambda < 2\Lambda} |M_\lambda^{a/q} f| \,\|_{l^2} \le C \|f\|_{l^2}$$

In proving (2.17) we showed

$$\| \sup_{\Lambda \le \lambda < 2\Lambda} |M_\lambda^{a/q} f| \,\|_{l^2} \le C q^{-\frac{\kappa}{d-1}} P^{n-d} \|f\|_{l^2} = C q^{-\frac{\kappa}{d-1}} \Lambda^{\frac{n}{d}-1} \|f\|_{l^2}$$

The sum in a, q is convergent and the proposition is proved. $\Box$.

3. The singular series

First we show the existence of a regular set of values Γ corresponding to a non-degenerate form Q.

Taking $\xi = 0$ formula (0.5) means that

$$r_Q(\lambda) = c_Q \lambda^{\frac{n}{d}-1} \sum_{q=1}^{\infty} K(q, 0, \lambda) + O(\lambda^{\frac{n}{d}-1-\delta})$$

By the well-known multiplicative property :

$K(q_1, 0, \lambda) K(q_2, 0, \lambda) = K(q_1 q_2, 0, \lambda)$ for q_1 and q_2 being relative primes, we have

$$K(\lambda) = \sum_{q=1}^{\infty} K(q, 0, \lambda) = \prod_{p\ prime} (\sum_{r=0}^{\infty} K(p^r, 0, \lambda)) = \prod_{p\ prime} K_p(\lambda)$$

. Note that $K(1, 0, \lambda) = 1$, then by estimate (1.5) it follows that $K_p(\lambda) = 1 + O(p^{-\frac{\kappa}{d-1}+1+\epsilon})\kappa =$

$$1/2 \le \prod_{p>R\ prime} |K_p(\lambda)| \le 2 \qquad (3.1)$$

We recall that $K_p(\lambda)$ is the density of solutions of the equation $Q(m) = \lambda$ among the p-adic integers, see [2]. More precisely,

Proposition 6 *Let* $r_Q(p^N, \lambda) = |\{m \in \mathbf{Z}^n/p^N\mathbf{Z}^n : \; Q(m) = \lambda \; (mod \; p^N)\}|$, *that is the number of solutions of the equation* $Q(m) = \lambda \; (mod \; p^N)$. *Then one has*

$$(3.2) \qquad \sum_{r=0}^{N} K(p^r, 0, \lambda) = p^{-n(N-1)} r_Q(p^N, \lambda)$$

Proof. First

$$r_Q(p^N, \lambda) = \sum_{m \; (mod \; p^N)} p^{-N} \sum_{b=1}^{p^N} e^{2\pi i (Q(m)-\lambda)\frac{b}{p^N}}$$

since the inner sum is equal to p^N or 0 according to $Q(m) = \lambda \; (mod \; p^N)$ or not. Next one writes $b = ap^{N-r}$, where $(a,p) = 1$, $a < p^r$ and $r = 0, \ldots, N$, and collects the terms corresponding to a fixed r which turn out to be $K(p^r, 0, \lambda)$. $\square$

We remark that this implies: $lim_{n\to\infty} p^{-n(N-1)} r_Q(p^N, \lambda) = K_p(\lambda)$.

To count the number of solutions $(mod \; p^N)$, one uses the p-adic version of Newton's method, see [7].

Lemma 9 *Let* p *be a prime,* λ *and* k, l *be natural numbers s.t.* $l > 2k$. *Suppose there is an* $m_0 \in \mathbf{Z}^n$ *for which*

$$(3.3) \qquad Q(m_0) \equiv \lambda \; (mod \; p^l)$$

moreover suppose, that p^k *is the highest power of* p *which divides all the partial derivatives* $\partial_j Q(m_0)$.

Then for $N \geq l$, *one has* $p^{-N(n-1)} r_Q(p^N, \lambda) \geq p^{-l(n-1)}$

Proof. For $N = l$ this is obvious. Suppose it is true for N, and consider all the solutions $m_1 \; (mod \; p^{N+1})$ of the form $m_1 = m + p^{N-k}s$ where $s \; (mod \; p)$. Then

$$Q(m + p^{N-k}s) - \lambda = Q(m) - \lambda + p^{n-k} Q'(m) \cdot s = 0 \;\; (mod \; p^{N+1})$$

,that is $a + b \cdot s = 0 \; (mod \; p)$ where $ap^N = Q(m) - \lambda$ and $bp^k = Q'(m)$. Then $b_j \neq 0 \; (mod \; p)$ for some j hence there are p^{n-1} solutions of this form. All obtained solutions are different $mod \; (p^{N+1})$, and m_1 satisfy the hypothesis of the lemma. $\square$

We remark that in case of $m = 1$, $k = 0$ the above argument shows that there are exactly $p^{(N-1)(n-1)}$ solutions m for which $m = m_0 \; (mod \; p)$ and $q(m) = \lambda (mod \; p^N)$.

Lemma 10 *let $Q(m)$ be a non-degenerate form, then there exists a set of regular values in the sense of* (0.4).

Proof. Let $\lambda_0 = Q(m_0) \neq 0$ for some fixed $m_0 \neq 0$. Let $p_1, \ldots, p_J$ be the set of primes less then R (R is defined in (3.1)). Let k be an integer s.t. p_j^k does not divide $d\lambda_0$, for all $j \leq J$, where d is degree of $Q(m)$. By the homogeneity relation $Q'(m_0) \cdot m = d\lambda_0$ it follows that p_j^k does not divide some partial derivative $\partial_i Q(m_0)$. Fix l s.t. $l > 2k$ and define the arithmetic progression

$\Gamma = \{\lambda_0 + k \prod_{j=1}^J p_j^l : \ k \geq k_Q\}$. Then we claim that Γ is a set of regular values. Indeed by Lemma 9. one has for $\lambda \in \Gamma$

$$K_{p_j}(\lambda) = \lim_{N \to \infty} p_j^{-n(N-1)} r_Q(p_j^N, \lambda) \geq p_j^{-l(N-1)}$$

This together with (3.1) ensures that the singular series $K(\lambda)$ remains bounded from below, and the error term becomes negligible by choosing k_Q large enough. □

Let us remark that along the same lines it can be shown, that all large numbers are regular values of $Q(m)$, if for each prime $p < R$ and each residue class s $(mod\ p)$, there is a solution of the equations $Q(m) = s$ $(mod\ p)$ s.t. $Q'(m) \neq 0$ $(mod\ p)$. This is the case for example for $Q(m) = \sum_j m_j^d$.

Let us fix a set of regular values Γ, and a rational point $k/p \neq 0$ in Π^n , where $k = (k_1, \ldots, k_n) \in \mathbf{Z}^n$. Define the measure space X to be the set of residue classes $(mod\ p)$, with each element having measure $1/p$. Let $T_j(x) = x + k_j$ $(mod\ p)$, then the family of transformations $T = (T_1 \ldots T_n)$ is commuting, measure preserving and ergodic. Indeed for some j, $k_j \neq 0$ $(mod\ p)$ and then T_j is ergodic. The function $f(x) = e^{2\pi i x/p}$ is a joint eigenfunction : $T_j f = e^{2\pi i k_j/p} f$ hence

$$\text{(3.5)} \qquad A_\lambda f = \frac{1}{r_Q(\lambda)} \hat{\sigma}_{Q,\lambda}(k/p) f$$

where $A_\lambda f$ are the averages defined in (0.7). We'll show below that the mean ergodic theorem (0.7) is not valid in this setting, and hence the condition *strong ergodicity* is necessary (note that $T_1^p = \ldots = T_n^p = Id$).

Lemma 11 *Let Γ be a set of regular values. Let p be a large enough prime: $p > d$, $p > R$, $p > \lambda_0$ (where λ_0 is the smallest element of Γ), and $k \in \mathbf{Z}^n$. Then for $\lambda \in \Gamma$, $\lambda = \lambda_0$ $(mod\ p)$ one has*

$$\text{(3.6)} \qquad \frac{1}{r_Q(\lambda)} \hat{\sigma}_{Q,\lambda}(k/p) = \frac{1}{r_Q(p, \lambda)} \sum_{m \in \mathbf{Z}^n / p\mathbf{Z}^n} e^{2\pi i \frac{m \cdot k}{p}} + O(\lambda^{-\delta})$$

Taking the Lemma granted for a moment, note that the expression:

$$S_k = \sum_{m\in\mathbf{Z}^n/p\mathbf{Z}^n} e^{2\pi i\frac{m\cdot k}{p}} \neq 0$$

for at least one $k \neq 0$, since otherwise the equation $Q(m) = \lambda = \lambda_0 \ (mod\ p)$ would have p^n or no solution, both cases are impossible (p being large enough). This follows from Plancherel's formula: $\sum_k |S_k|^2 = p^n|r_Q(p,\lambda_0)|$ on the group $\mathbf{Z}^n/p\mathbf{Z}^n$. Thus (0.7) is not true, assuming only that the family of transformations is ergodic.

Proof. For a regular value $r_Q(\lambda) = c_Q K(\lambda)\lambda^{n/d-1} + O(\lambda^{n/d-1-\delta})$ where $|K(\lambda)| \gg 1$, hence by (0.5), it is enough to show

$$(3.7) \qquad c_Q^{-1}\frac{1}{K(\lambda)}\sum_{q=1}^{\infty}\sum_{l\in\mathbf{Z}^n} K(q,l,\lambda)\psi(qk/p-l)d\tilde{\sigma}_Q(\lambda^{1/d}(k/p-l/q)) =$$

$$= \frac{1}{r_Q(p,\lambda)}\sum_{m\in\mathbf{Z}^n/p\mathbf{Z}^n} e^{2\pi i\frac{m\cdot k}{p}} + O(\lambda^{-\delta})$$

For q not divisible by p, $|\frac{k}{p} - \frac{l}{q}| \geq \frac{1}{pq}$, hence each term in the sum is bounded by $q^{-\frac{\kappa}{d-1}+\epsilon}\lambda^{-\kappa/(d-1)+1+\epsilon}$ by (1.5) and (1.13). There is at most one nonzero term in the l sum for fixed q, and thus the total sum for q not divisible by p is of $O(\lambda^{-\delta})$.

For $q = bp$, in (3.7) only those terms for which $k/p = l/q$ are nonzero, hence the sum becomes

$$\frac{1}{K(\lambda)}\sum_{b=1}^{\infty} K(bp,bk,\lambda)$$

We write $q = cp^r$ where $(c,p) = 1$ and use the multiplicative property

$$K(cp^{r+1}, ckp^r, \lambda) = K(c,0,\lambda)K(p^{r+1}, kp^r, \lambda)$$

It is a straightforward computation using the chinese remainder theorem. At this point it is enough to show

$$(3.8) \qquad \frac{1}{K(\lambda)}\Big(\sum_{(c,p)=1} K(c,0,\lambda)\Big)\Big(\sum_{r=1}^{\infty} K(p^{r+1}, kp^r, \lambda)\Big) =$$

$$= \frac{1}{r_Q(p,\lambda)}\sum_{m\ (mod\ p)} e^{2\pi i\frac{m\cdot k}{p}}$$

Again by multiplicativity

$$(3.9)\qquad \sum_{(c,p)=1} K(c,0,\lambda)\cdot\sum_{r=1}^{\infty} K(p^r,0,\lambda) = \sum_{q=1}^{\infty} K(q,0,\lambda)$$

For the other factor in (3.8) one has

$$(3.10)\qquad \sum_{r=1}^{\infty} K(p^{r+1},kp^r,\lambda) = p^{-(n-1)} \sum_{m\ (mod\ p)} e^{2\pi i \frac{m\cdot k}{p}}$$

Similarly as in (3.2)

$$\sum_{m\ (mod\ p^N)} p^{-N}\sum_{b=1}^{p^N} e^{2\pi i (Q(m)-\lambda)\frac{b}{p^N}} e^{2\pi i \frac{m\cdot k}{p^N}}$$

and writes $b = ap^{N-r}$, where $(a,p)=1$, $a<p^r$ and $r=0,\dots,N$. Each term corresponding to a fixed r is $K(p^r,kp^{r-1},\lambda)$ for $r\geq 1$, while the term corresponding to $r=0$ is zero.

Next, let m_0 be a solution of $Q(m)=\lambda\ (mod\ p)$. Then by homogeneity $Q'(m_0)\cdot m_0 = d\lambda = d\lambda_0 \neq 0$ it follows by the remark after Lemma 9. that the number of solutions: $m\ (mod\ p^N)$ for which $m=m_0\ (mod\ p)$ and $Q(m)=\lambda\ (mod\ p^N)$ is exactly $p^{(n-1)(N-1)}$. Thus

$$\sum_{m\ (mod\ p^N)} p^{-N}\sum_{b=1}^{p^N} e^{2\pi i (Q(m)-\lambda)\frac{b}{p^N}} e^{2\pi i \frac{m\cdot k}{p^N}} = p^{-(n-1)} \sum_{m\ (mod p)} e^{2\pi i \frac{m\cdot k}{p}}$$

and this proves (3.10).

By the same argument

$$(3.11)\qquad K_p(\lambda) = p^{-(n-1)} r_Q(p,\lambda)$$

and (3.8) follows immediately from (3.9), (3.10) and (3.11). $\square$

4. The L^2 ergodic theorem

In this section, we prove Theorems 1-2. and Lemma 2. First we give the

Proof of Theorem 1. Let $\phi_\lambda(x) = \phi(x/\lambda^{1/d})$, the one has

$$\sum_{Q(m)=\lambda} \phi_\lambda(m) = \int_{\Pi^n} \hat{\sigma}_{Q,\lambda}(\xi)\hat{\phi}_\lambda(\xi)\, d\xi$$

where

$$\hat{\phi}_\lambda(\xi)\,d\xi = \sum_{m\in\mathbf{Z}^n} \phi_\lambda(m) e^{-2\pi i m\cdot\xi} = \sum_{m\in\mathbf{Z}^n} \tilde{\phi}_\lambda(\xi+m)$$

by Poisson summation (here $\tilde{\phi}_\lambda(\xi)$ denotes the Fourier transform on $\Re^n$. Since the exponential sum $\hat{\sigma}_{Q,\lambda}(\xi)$ is a smooth periodic function on $\Re^n$ it follows

$$\sum_{Q(m)=\lambda} \phi_\lambda(m) = \int_{\Re^n} \hat{\sigma}_{Q,\lambda}(\xi)\tilde{\phi}_\lambda(\xi)\,d\xi \tag{4.1}$$

Write $\hat{\sigma}_{Q,\lambda}(\xi) = m_\lambda(\xi) + \mathcal{E}_\lambda(\xi)$ and estimate the contribution of the error term

$$\int_{\Re^n} |\mathcal{E}_\lambda(\xi)\tilde{\phi}_\lambda(\xi)|\,d\xi \le C_\delta \lambda^{n/d-1-\delta} \|\tilde{\phi}_\lambda\|_1 \le C_\delta \lambda^{n/d-1-\delta} \tag{4.2}$$

We used the error estimate in (0.5) and the fact that $\|\tilde{\phi}_\lambda\|_1 = \|\tilde{\phi}\|_1 \le C$. Recall that

$$m_\lambda(\xi) = \sum_{q=1}^{\infty} \sum_l K(q,l,\lambda)\psi(q\xi - l) d\tilde{\sigma}_{Q,\lambda}(\xi - l/q)$$

Next we estimate the contribution of the terms corresponding to $l \neq 0$. For $q \ge \lambda^{\frac{1}{2d}}$ we use

$$\sum_{q\ge\lambda^{\frac{1}{2d}}} \sum_{l\neq 0} |K(q,l,\lambda)\psi(q\xi - l) d\tilde{\sigma}_{Q,\lambda}(\xi - l/q)| \le \tag{4.3}$$

$$\le C\lambda^{n/d-1} \sum_{q\ge\lambda^{\frac{1}{2d}}} q^{-2} \le C_\delta \lambda^{n/d-1-\delta}$$

and after integrating we get the same estimate as in (4.2) ($\frac{\kappa}{d-1} > 2$). For $q \le \lambda^{\frac{1}{2d}}$ we give the estimate

$$\sum_{q\le\lambda^{1/2d}} \sum_{l\neq 0} \int_{\Re^n} |K(q,l,\lambda)\psi(q\xi - l) d\tilde{\sigma}_{Q,\lambda}(\xi - l/q)\, \tilde{\phi}_\lambda(\xi)|\,d\xi \le c_N \lambda^{-N} \tag{4.4}$$

for any $N > 0$ integer. For fixed $l \neq 0$, on the support of the cut-off factor $\psi(q\xi - l)$, one has $\|\xi - l/q\| \leq 1/(4q)$, which implies $\|\xi\| \geq 1/(2q)$, and also $\|\xi\| \geq \|l\|/(2q)$ (here $\|\eta\| = \sup_j |\eta_j|$ denotes the sup-norm on $\Re^n$). Thus

$$(4.5) \quad |\tilde{\phi}_\lambda(\xi)| \leq \ C_N \lambda^{n/d}(1 + \lambda^{1/d}|\xi|)^{-2N} \leq$$

$$_N\lambda^{n/d}(1 + \lambda^{1/d}/2q)^{-N}(1 + c|l|/2q)^{-N}$$

Integrating in ξ over the region $\|\xi - l/q\| \leq 1/(4q)$, and then summing in l and in $q \leq \lambda^{\frac{1}{2d}}$ one obtains (4.4).

Estimates (4.3) and (4.4) imply together that the total contribution of the terms corresponding to $l \neq 0$ in (4.1), is $O(\lambda^{n/d-1-\delta}$.

Finally, we note that

$$(4.7) \quad \sum_{q=1}^{\infty} \int |K(q, 0, \lambda)(1 - \psi(q\xi)) d\tilde{\sigma}_\lambda(\xi)\tilde{\phi}_\lambda(\xi)|\, d\xi \leq C_\delta \lambda^{\frac{n}{d}-1-\delta}$$

by the same argument as used in proving (4.3) and (4.4). Indeed the range of integration is $|\xi| \geq c/q$ where both for $q \geq \lambda^{1/2d}$ and for $q \leq \lambda^{1/2d}$, one has a gain, using the decay of the the factor $K(q, 0, \lambda)$ for small, and the decay of $\tilde{\phi}_\lambda$ for large values of q.

Using (4.3), (4.4) and (4.7) one has

$$(4.8) \quad \int_{\Re^n} \hat{\sigma}_{Q,\lambda}(\xi)\tilde{\phi}_\lambda(\xi)\, d\xi = c_Q K(\lambda) \int_{\Re^n} \tilde{\sigma}_{Q,\lambda}(\xi)\tilde{\phi}_\lambda(\xi)\, d\xi + O(\lambda^{\frac{n}{d}-1-\delta}) =$$

$$= r_Q(\lambda) \int_{Q(y)=1} \phi(y)\, d\sigma_Q(y) + O(\lambda^{\frac{n}{d}-1-\delta})$$

Indeed one replaces the singular series $c_Q K(\lambda)$ by $\lambda^{-n/d+1} r_Q(\lambda)$, use Plancherel's formula, and a change of variables $x = \lambda^{1/d} y$.

This proves the Theorem, since $r_Q(\lambda) \geq C_Q \lambda^{n/d-1}$ for regular values λ. $\square$

Proof of Lemma 2. One writes

$$(4.9) \quad \frac{1}{r_Q(\lambda)} |\hat{\sigma}_{Q,\lambda}(\xi)| \leq C_\delta \lambda^{-n/d+1} |m_\lambda(\xi)| + O(\lambda^{-\delta})$$

For q fixed and $\xi \notin \mathbf{Q}^n$ (i.e. when ξ_j is irrational for some j)

$$(4.10) \quad \lambda^{-n/d+1}|m_{q,\lambda}(\xi)| = c_Q \sum_l |K(q, l, \lambda)\psi(q\xi - l) d\tilde{\sigma}_{Q,\lambda}(\xi - l/q)| \leq$$

$$\leq C_Q q^{-\frac{\kappa}{d-1}+\epsilon} |d\tilde{\sigma}_{Q,\lambda}(\lambda^{1/d}\{q\xi\}/q)|$$

where $\{\xi\} = min\,|\xi - l|$. Indeed in the l sum only term corresponding to the closest lattice point to $q\xi$ is nonzero.

Note that $\{q\xi\} \neq 0$ for every q, since otherwise $\xi \in \mathbf{Q}^n$. Then by (1.13) and (4.10) for $q \leq \lambda^{1/2d}$ we have the estimate $\lambda^{-n/d+1}|m_{q,\lambda}(\xi)| \leq Cq^{-1-\epsilon}\lambda^{-\delta}$, while for $q \geq \lambda^{1/2d}$ one uses the bound $q^{-1-\epsilon}$. The lemma follows by summing in q. $\square$

In both the mean and pointwise ergodic theorem the Spectral theorem will play an essential role. Also, strong ergodidity is a condition on joint spectrum of the shifts T_j ($T_j f(x) = f(T_j x)$). To see that let (X, μ) be a probability measure space, $T = (T_1 \dots T_n)$ be a family of commuting, measure preserving and invertible transformations. By the Spectral theorem there exists a positive Borel measure ν_f on the torus Π^n, s.t.

$$\langle P(T_1, \dots, T_n)f, f\rangle = \int_{\Pi^n} p(\xi)d\nu_f(\xi) \tag{4.11}$$

for every polynomial $P(z_1, \dots, z_n)$, where

$$p(\xi) = p(\xi_1, \dots, \xi_n) = P(e^{2\pi i\xi_1}, \dots, e^{2\pi i\xi_n})$$

and $\langle , \rangle$ denotes the inner product on $L^2(X, \mu)$. We recall two basic facts

i) For $r \in \Pi^n$, $\nu_f(r) > 0$ if and only if r is a joint eigenvalue of the shifts T_j, (i.e. there exists $g \in L^2(X)$ s.t. $T_j g = e^{2\pi i r_j} g$ for each j.

ii) If the family $T = (T_1, \dots, T_n)$ is ergodic, then
$\nu_f(0) = |\langle f, \mathbf{1}\rangle|^2 = |\int_X f d\mu|^2$.

Proposition 7 *Suppose the family $T = (T_1, \dots, T_n)$ is ergodic. Then it is strongly ergodic if and only if $\nu_f(r) = 0$ for every $r \in \mathbf{Q}^n$, $r \neq 0$.*

Proof. Suppose $\nu_f(l/q) > 0$ for some $l \neq 0$, then there exists $g \in L^2(X, \mu)$ s.t. $T_j g = e^{2\pi i l_j/q} g$ $\forall j$. But then $T_j^q g = g$ $\forall j$ but $g \neq$ *constant* since $l \neq 0$.

On the other hand suppose that $T_j^q g = g$, $\forall j$ for some $g \neq$ *constant*. Then the functions $g_{s_1 \dots s_n}$ for $s \in \mathbf{Z}^n/q\mathbf{Z}^n$ defined by

$$g_{s_1 \dots s_n} = \sum_{m \in \mathbf{Z}^n/q\mathbf{Z}^n} e^{-2\pi i \frac{m \cdot s}{q}} T_1^{m_1} \dots T_n^{m_n} g$$

are joint eigenfunctions of with eigenvalues s_j/q. They cannot vanish for all $s \neq 0$ $(mod\ q)$, because then one would have $T_j g = g\ \forall j$, as can be seen easily by expressing $T_j g$ in terms of the functions $g_{s_1 \ldots s_n}$. □

Proof of Theorem 2. We start by

$$\|A_\lambda f - \langle f, \mathbf{1}\rangle \mathbf{1}\|_2^2 = \|A_\lambda f\|_2^2 - |\langle f, \mathbf{1}\rangle|^2 = \int_{\Pi^n/\{0\}} \frac{|\hat{\sigma}_{Q,\lambda}(\xi)|^2}{r_Q(\lambda)^2}\, d\nu_f(\xi)$$

The point is that $\nu_f(\mathbf{Q}^n/\{0\}) = 0$ by the strong ergodicity condition, moreover the integrand pointwise tends to zero on the irrationals by Lemma 2, and is majorized by **1**. It follows from the Lebesgue dominant convergence theorem, that the integral also tends to 0 as $\lambda \to \infty$. This proves the theorem. □

5. The discrete spherical maximal theorem

We prove Theorem 4. now. It plays a crucial role in the proof of the pointwise ergodic theorem.

Let $\phi \in l^2\mathbf{Z}^n$, the averages we are interested in: $\frac{1}{r_Q(\lambda)} \sum_{Q(l)=\lambda} \phi(m-l)$ will be replaced by

$$(5.1) \qquad N_\lambda \phi(m) = \frac{1}{\lambda^{n/d-1}} \sum_{Q(l)=\lambda} \phi(m-l)$$

Indeed it is enough to prove the maximal theorem for the averages N_λ, since for regular values: $r_Q(\lambda) \geq c_Q \lambda^{n/d-1}$. We write

$$(5.2) \qquad N_\lambda \phi = M_\lambda \phi + \mathcal{E}_\lambda \phi = \sum_{q=1}^{\infty} \sum_{(a,q)=1} M_\lambda^{a/q} \phi + \mathcal{E}_\lambda \phi$$

where M_λ, $M_\lambda^{a/q}$, $\mathcal{E}_\lambda$ denote the mulitpliers corresponding to the functions $\lambda^{-n/d+1} m_\lambda(\xi)$, $m_\lambda^{a/q}(\xi)$, $\mathcal{E}_\lambda(\xi)$. We denote by M_*, $M_*^{a/q}$, $\mathcal{E}_*$ the corresponding maximal operators.

By Lemma 8.,

$$(5.2) \qquad \|\mathcal{E}_* \phi\|_{l^2} \leq \sum_{k=0}^{\infty} \| \sup_{2^k \leq \lambda < 2^{k+1}} |\mathcal{E}_\lambda \phi|\, \|_{l^2} \leq C_\delta \sum_{k=0}^{\infty} 2^{-k\delta} \|\phi\|_{l^2} \leq C_\delta \|\phi\|_{l^2}$$

The same shows, that

$$(5.3) \qquad \| \sup_{\Lambda \leq \lambda} |\mathcal{E}_\lambda \phi|\, \|_{l^2} \leq C_\delta \Lambda^{-\delta} \|\phi\|_{l^2}$$

Thus to prove Theorem 4. it is enough to show

Lemma 12 *Let $q \geq 1$, and a s.t. $(a, q) = 1$ be given. The one has*

$$(5.2) \qquad \|M_*^{a/q}\|_{l^2} \leq C_\epsilon q^{-\frac{\kappa}{d-1}+\epsilon} \|\phi\|_{l^2}$$

It is understood that $Q(m)$ is a non-degenerate form, hence $\kappa = \frac{1}{2(d-1)} V_Q > 2$ and $\epsilon > 0$ can be taken arbitrary small. Hence in the right side of (5.3) we can take the bound $Cq^{-2-\epsilon}$, but we'd like to emphasize the explicit dependence on κ.

Assuming the Lemma for a moment, by sub-additivity it follows:

$$\|M_*\phi\|_{l^2} \leq C \sum_{q=1}^{\infty} q \cdot q^{-2-\epsilon} \|\phi\|_{l^2} \leq C\|\phi\|_{l^2}$$

Together with estimate (5.2) this proves Theorem 4.

The proof of the lemma is based on a general result, proved in [6]

Lemma 13 *Let $q \geq 1$ be a fixed integer and B be a finite dimensional Banach space. Let $m(\xi)$ be a bounded measurable function on $\Re^n$, taking values in B, and supported in the cube $[-\frac{1}{2q}, -\frac{1}{2q}]^n$.*

Define the periodic extension by

$$m_{per}^q(\xi) = \sum_{l \in \mathbf{Z}^n} m(\xi - l/q)$$

Let $T : L^2(\Re^n) \to L_B^2(\Re^n)$ (where $L_B^2(\Re^n)$ is the space of square integrable functions taking values in the space B), be the multiplier operator corresponding to the function $m_\lambda(\xi)$.

Similarly let $T_{dis}^q : L^2(\mathbf{Z}^n) \to L_B^2(\mathbf{Z}^n)$ be the multiplier operator corresponding to the periodic function $m_{per}^q(\xi)$.

Then one has

$$(5.4) \qquad \|T_{dis}^q\|_{L^2(\mathbf{Z}^n) \to L_B^2(\mathbf{Z}^n)} \leq C\|T\|_{L^2(\Re^n) \to L_B^2(\Re^n)}$$

where the constant C does not depend on the Banach space B, and is also independent of q.

Proof of Lemma 12. Choose a smooth function ψ' supported in $[-1/2, 1/2]^n$ for which $\psi = \psi'\psi$. Then $m_\lambda^{a/q}(\xi)$ can be written as the product of the functions

$$(5.5) \qquad m^{a/q}(\xi) = \sum_{l \in \mathbf{Z}^n} G(a, l, q)\psi'(\xi - l/q)$$

and

$$(5.6)\qquad m_\lambda^q(\xi) = \sum_{l\in\mathbf{Z}^n} \psi(\xi - l/q) d\tilde{\sigma}_\lambda(\xi - l/q)$$

For the first multiplier operator $M^{a/q}$ it is bounded from l^2 to itself with norm: $\sup_\xi |m^{a/q}(\xi)| \leq C_\epsilon q^{-\frac{\kappa}{d-1}+\epsilon}$.

The sequence of functions $m_\lambda^q(\xi)$ defined by (5.6) can be considered as a function mapping from $\Re^n$ to the banach space B_Λ which is the l^∞ space of functions of $1 \leq \lambda \leq \Lambda$ for some fixed Λ.

The multiplier corresponding to $\psi(q\xi)d\tilde{\sigma}_\lambda(\xi)$ is a bounded operator from $L^2(\Re^n)$ to $L^2_B(\Re^n)$ (B being the l^∞ space of functions of $\lambda > 0$), which is the content of Theorem 5. Then one applies Lemma 13. to see that the multiplier $m_\lambda^q(\xi)$ is bounded from $l^2\mathbf{Z}^n$ to $l^2_{B_\Lambda}\mathbf{Z}^n$ with norm independent of Λ. This implies (5.2). $\square$

6. The pointwise ergodic theorem

The proof of Theorem 3. consists of a number of reductions, the argument was motivated by that of Bourgain's ergodic theorem corresponding to arithmetic subsets of integers (see [3]). However in our case the averages are taken over disjoint sets, a condition on the joint spectrum must be imposed, and the Spectral theorem will play an essential role.

Let $f \in L^2(X,\mu)$, we can suppose $\int_X f\,d\mu = 0$, and then we have to show that $|A_\lambda f(x)| \to 0$ for μ almost every x, as $\lambda \to \infty$ and $\lambda \in \Gamma$. Then again we can replace the factor $r_Q(\lambda)$ by $\lambda^{n/d-1}$ in the averages.

i) We start with a standard reduction to shifts on $\mathbf{Z}^n$. Let (X,μ) be a probability measure space, $T = (T_1,\ldots,T_n)$. For $x \in X$ and $L > 0$ and define: $\phi_{L,x}(m) = f(T^m x)$ if $\|m\| \leq L$ and to be 0 otherwise. Here $m = (m_1,\ldots m_n) \in \mathbf{Z}^n$, $\|m\| = \sup_j |m_j|$ and $T^m x = T_1^{m_1}\cdot\ldots\cdot T_n^{m_n}x$.

Notice that for fixed $\Lambda < L$

$$(6.1)\qquad A_\lambda^* f(T^l x) = \sup_{\lambda\leq\Lambda} |A_\lambda f(T^l x)| =$$

$$= \sup_{\lambda\leq\Lambda} |N_\lambda \phi_{L,x}(l) = |N_\lambda^* \phi_{L,x}(l)$$

for $\|l\| \leq c(L-\Lambda)$ Thus taking the square, summing in l (for $\|l\| \leq c(L-\Lambda)$), and integrating over the space X one obtains

$$(6.2)\qquad c(L-\Lambda)^n \|A_\lambda^* f\|_{L^2(X)} \leq \int_X \|N_\lambda^* \phi_{L,x}\|_{l^2}\,d\mu$$

using the fact that the transformations T^l are measure preserving. Also

$$(6.3) \qquad \int_X \|\phi_{L,x}\|_{l^2}^2 \, d\mu = c_n L^n \|f\|_{L^2(x)}^2$$

Then letting $\Lambda \to \infty$, it follows that the $L^2(X) \to L^2(X)$ norm of the maximal operator A_* is majorized by the $l^2 \to l^2$ norm of the discrete maximal operator N_*. Then it is enough to prove the pointwise ergodic theorem for a dense subset of $L^2(X)$, p.e. for $L^\infty(X)$.

ii) Following [5], one reduces pointwise convergence to L^2 bounds for "truncated" maximal operators. Suppose indirect, that

$$\mu\{x: \ \limsup |A_\lambda f(x)| > 0\} > 0$$

then the same is true with a small constant $\alpha > 0$ inserted:

$$\mu\{x: \ \limsup |A_\lambda f(x)| > 2\alpha\} > 2\alpha$$

and using the definition of the upper limit it is easy to see, that to each λ_k if λ_{k+1} is chosen large enough then

$$\mu\{x: \ A_k^* f(x) = sup_{\lambda_k \le \lambda \le \lambda_{k+1}} |A_\lambda f(x)| > \alpha\} > \alpha$$

which implies $\|A_*^k f\|_2^2 > \alpha^3$, $\forall \ k$. Lets fix such a sequence λ_k which is quickly increasing: $\lambda_{k+1} > 4{\lambda_k}^{4d}$. Then it is enough to prove

$$(6.4) \qquad \frac{1}{K} \sum_{k \le K} \|A_*^k f\|_2^2 < \alpha^3$$

for $K > K_(\alpha)$. This means that the Cesaro averages converges in (6.4) tends to 0 (the terms themselves may not converge to 0).

Now fix K and choose $L > \lambda_{K+1}$. The reasoning in i) leads to

$$(6.5) \qquad c(L-\Lambda)^n \frac{1}{K} \sum_{k \le K} \|A_k^* f\|_2 \le \int_X \frac{1}{K} \sum_{k \le K} \|N_k^* \phi_{L,x}\|_{l^2} \, d\mu$$

where N_*^k is defined analogously to A_*^k. Thus it is enough to prove

$$(6.6) \qquad \int_X (\frac{1}{K} \sum_{k \le K} \|N_k^* \phi_{L,x}\|_{l^2}^2) \, d\mu \le c_n \alpha^3 L^n \|f\|_2^2$$

for $K > K(\alpha)$ and $L > L(K, \alpha)$.

By (6.3), inequality (6.6) would follow, if the same would be true pointwise, that is $1/K \sum_{k\leq K} \|N_k^* \phi_{L,x}\|_{l^2}^2 \to 0$ for every x, however this seems to be true just in average, and has to do with the fact that nearby averages cannot be compared.

i3) We use the approximations to N_λ introduced in Section 2., and the transfer principle (5.4) to reduce the estimates to that of $L^2 \to\ L^2$ norms of the corresponding maximal operators acting on $\Re^n$.

We often use the following notations; if $\gamma_\lambda(\xi)$ are continuous functions on Π^n, then denote by Γ_λ the corresponding multipliers and by Γ_k^* the maximal operator: $\Gamma_k^*\phi{=}\sup_{\lambda_k\leq\lambda<\lambda_{k+1}|\Gamma_\lambda\phi|}$.

Since

$$\lambda^{-n/d+1}\hat{\sigma}_\lambda(\xi) = \sum_{q=1}^{\infty} \lambda^{-n/d+1} m_{q,\lambda}(\xi) + \lambda^{-n/d+1}\mathcal{E}_\lambda(\xi)$$

then by estimates (2.6) and (5.2)

$$(6.7) \qquad \|\mathcal{E}_k^*\|_{l^2\to l^2} \leq C_\delta \lambda_k^{-\delta}$$

and

$$(6.8) \qquad \|\sum_{q\geq q_\alpha} M_{q,k}^*\|_{l^2\to l^2} \leq C q_\alpha^{-\epsilon}$$

If we apply (6.7) and (6.8) to the function $\phi_{L,x}$ integrate the square over X and average for $k \leq K$, the total contribution to the L^2 norm is less then:

$$(q_\alpha^{-\epsilon} + c_\delta K^{-1}) \int_X \|\phi_{L,x}\|_{l^2}^2 \, d\mu(x) \leq \alpha^3 L^n \|f\|_{L^2(X)}$$

by choosing K and q_α large enough w.r.t. α and ϵ .

Thus enough to deal with the finitely many maximal operators attached to the functions $m_\lambda^{a/q}(\xi)$, for $q \leq q_\alpha$ and $a \leq q$, $(a, q) = 1$. Then we can fix a and q, and write

$$(6.9) \qquad \lambda^{-n/d+1} m_\lambda^{a/q}(\xi) = \sum_{l\in\mathbf{Z}^n} G(a,l,q)\psi(q\xi - l) d\tilde{\sigma}(\lambda^{1/d}(\xi - l/q)) =$$

$$= \sum_{s\in\mathbf{Z}^n/q\mathbf{Z}^n} G(a,s,q)\, \psi(q\xi - s) d\tilde{\sigma}_Q(\lambda^{1/d}(\xi - s/q))_{per}$$

where $\gamma_{per}(\xi) = \sum_{l_1 \in \mathbf{Z}^n} \gamma(\xi - l_1)$ denotes the periodization of γ. Indeed write $l = ql_1 + s$ and use the fact that $G(a,l,q) = G(a,s,q)$. Again we can fix s (there are at most $q^n \leq q_\alpha^n$ choice for each q).

We remark that for $\phi \in l^2$ and $\phi_{s/q}(m) = e^{-2\pi ims/q}\phi(m)$ i.e. $\hat{\phi}_{s/q}(\xi) = \hat{\phi}(\xi + s/q)$, one has

$$M^*_{s/q,k}\phi = M^*_k \phi_{s/q}$$

where $M^*_{s/q,k}$ is the maximal operator which corresponds to the function $\psi(q\xi - s)d\tilde{\sigma}(\lambda^{1/d}(\xi - s/q))_{per}$, while M^*_k corresponds to $\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}(\xi))_{per}$. Indeed one changes variables $(\xi - s/q) \to \xi$ in evaluating the multipliers (the factors $e^{2/piims/q}$ vanish when taking absolute values).

We are in a position to apply the continuous spherical maximal theorem, and further decompose the functions $\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}(\xi))$ to get decay estimates. Let

$\mathbf{1} = \omega_{k,0} + \omega_{k,1} + \omega_{k,2}$ be smooth partition of unity on $\|\xi\| = \sup_j |\xi|_j \leq 1/2$ such that

$\omega_{k,0}(\xi) = 0$ unless $\|\xi\| \geq \frac{1}{2}\lambda_{k+1}^{-2}$,

$\omega_{k,1}(\xi) = 0$ unless $\frac{1}{2}\lambda_{k+1}^{-2}\|\xi\| \leq \lambda_k^{-\frac{1}{2d}}$ and

$\omega_{k,2}(\xi) = 0$ unless $\lambda_k^{-\frac{1}{2d}} \leq \|\xi\|$

Accordingly we have the decomposition: $M^*_k \leq M^*_{k,0} + M^*_{k,1} + M^*_{k,2}$ and estimate each term separately.

For fixed λ, using the fact that $|d\tilde{\sigma}(\lambda^{1/d}\xi) - c_Q| \leq \lambda^{1/d}|\xi|$ ($c_Q = d\tilde{\sigma}(0)$), one has

$$(6.10) \qquad |\omega_{k,0}(\xi)\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}\xi) - c_Q\omega_{k,0}(\xi)\psi(q\xi)| \leq C\lambda^{1/d}\lambda_{k+1}^{-2}$$

Thus by the standard square function estimate the $l^2 \to l^2$ norm of the maximal operator (taking the sup over $\lambda_k \leq \lambda < \lambda_{k+1}$) corresponding to the functions in (6.9) is bounded by:

$\sum_{\lambda < \lambda_{k+1}} \lambda^{2/d}\lambda_{k+1}^{-4})^{1/2} \leq \lambda_{k+1}^{-1}$ $(d \geq 2)$.

To estimate the maximal operator $M^*_{k,1}$ corresponding to the functions $\omega_{k,1}(\xi)\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}(\xi))_{per}$, we first use the transfer principle to see that it is bounded by the $L^2(\Re^n) \to L^2(\Re^n)$ norm of the maximal operator corresponding to the functions $\omega_{k,1}(\xi)\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}(\xi))$. Notice that the maximal operator (the sup taken over all $\lambda > 0$) corresponding to the functions $d\tilde{\sigma}(\lambda^{1/d}(\xi))$ is bounded from $L^2 \to L^2$ by Theorem 5.

Thus for $\phi_{s/q} = \phi_{L,x,s/q}$ one has

$$(6.11) \quad \|M^*_{k,1}\phi_{s/q}\|_{l^2} \leq C_Q \int_{\Pi^n} |\omega_{k,1}(\xi)|^2 |\hat{\phi}(\xi + s/q)|^2 \, d\xi$$

The point is that since the sequence λ_k is quickly increasing $\lambda_{k+1} > 4\lambda^{4d}$ each point can belong to at most 3 intervals I_k on which $\omega_{k,1}$ supported. Hence averaging over $k \leq K$ the right side of (6.10), gives a contribution of $3/K\|\phi\|^2_{l^2}$.

Finally, the family of functions $\omega_{k,2}(\xi)\psi(q\xi)d\tilde{\sigma}(\lambda^{1/d}(\xi))$ satisfy the conditions of Lemma 7. Then (1.16) and (5.4) imply the bound

$$(6.10) \quad \|M^*_{k,2}\phi_{s/q}\|_{l^2} \leq C_Q \lambda_k^{-\frac{1}{2d}} \|\phi\|_{l^2}$$

Note that (6.9)-(6.11) mean, that the maximal function

$$1/K \sum_{k \leq K} \|M^*_k \phi_{s/q}\|^2_{l^2} \leq C \int_{\Pi^n} |\psi(q\xi)\omega_{k,1}(\xi)|^2 |\hat{\phi}(\xi + s/q)|^2 \, d\xi +$$

$$+ O(K^{-1}) \|\phi_{s/q}\|_{l^2}$$

i4) It is enough to prove now for fixed $r = s/q$, that

$$(6.11) \quad L^{-n} \int_X \int_{\Pi^n} \omega_{k,1}(\xi) |\hat{\phi}(\xi + s/q)|^2 \, d\xi \, d\mu(x)) < |\alpha|^3 \|f\|^2_2$$

if $k > k(\alpha)$ and $L > L(k, \alpha)$, where we wrote $\omega_k(\xi) = |\omega_{k,1}(\xi)|^2$ for simplicity of notation.

By applying Plancherel for the inner integral in (6.11), one obtains

$$L^{-n} \int_X \sum_{m,m'} \phi_{L,x}(m) \bar{\phi}_{L,x}(m') \hat{\omega}_k(m - m') e^{2\pi i(m-m')s/q} \, d\xi \, d\mu(x) =$$

$$= L^{-n} \sum_{\|m\| \leq L, \, \|m'\| \leq L} \langle T^{m-m'} f, f \rangle \hat{\omega}_k(m - m') e^{2\pi i(m-m')s/q} =$$

$$= L^{-n} \int_{\Pi^n} \sum_{\|m\| \leq L, \, \|m'\| \leq L} \hat{\omega}_k(m - m') e^{2\pi i(m-m')(\theta + s/q)} \, d\nu_f(\theta) =$$

$$= L^{-n} \int_{\Pi^n} \sum_{l \in \mathbf{Z}^n} a_L(l) \hat{\omega}_k(l) e^{2\pi i(\theta + s/q)} \, d\nu_f(\theta)$$

by the spectral theorem, where $a_L(l) = |\{(m, m');\ \|m\| \leq L,\ \|m'\| \leq L,\ m - m' = l\}|$. Finally one gets

$$\int_{\Pi^n} (L^{-n}\hat{a}_L * \omega_k)\,(\theta + s/q)\,d\nu_f \tag{6.12}$$

where $*$ denotes the convolution on Π^n (w.r.t. Lebesgue measure).

Note that

$$L^{-n}\hat{a}_L(\theta) = L^{-n}|\sum_{m=-L}^{L} e^{2\pi i m\theta}|^{2n} \leq L^n\ \min(1, \frac{1}{L\{\theta\}})^{2n}$$

This means that $L^{-n}\hat{a}_L$ is a δ-sequence (i.e. weakly converges to a Dirac delta) as $L \to \infty$. Indeed it is easy to see that: $L^{-n}\hat{a}_L * \omega_k \leq c\omega_k + \epsilon$ for every $\epsilon > 0$ if L is large enough w.r.t. to λ_k and ϵ.

Finally if we substitute this estimate into (6.12), then using the fact that $\omega_k(\theta) = 0$ unless $\|\theta\| \leq \lambda_k^{-1/2d}$, one has

$$\int_{\Pi^n} (L^{-n}\hat{a}_L * \omega_k)\,(\theta + s/q)\,d\nu_f \ \leq c d\nu_f\{\theta :\ \|\theta + s/q\| < \lambda_k^{-1/2d}\} +$$

$$+ \epsilon\, d\nu_f(\Pi^n) \leq \alpha^3 \|f\|^2_{L^2(X)}$$

if k is large enough w.r.t. α and L is large enough w.r.t. k and α.

Indeed $d\nu_f(\Pi^n) = \|f\|^2_{L^2(X)}$, and only here we use the condition strong ergodicity, that is the condition that $d\nu_f\{s/q\} = 0$ for every rational point $s/q \neq 0$ (note that by our assumption $d\nu_f\{0\} = \int_X f\,d\mu = 0$ also), which implies $d\nu_f\{\theta :\ \|\theta + s/q\| < \lambda_k^{-1/2d}\} \to 0$ as $k \to \infty$.

This proves Theorem 4. $\square$.

References

[1] Arnold, V., Varchenko, A.,: *Singularities of differentiable mappings I-II*, Monographs in Math., Birkhauser, Boston (1988)

[2] Birch, B.J.,: *Forms in many variable*, Proc. Roy. Soc. Ser. A, 265. 245-263 (1961)

[3] Bourgain, J.,: *On the maximal ergodic theorem for certain subsets of integers* Israeli J. Math., 61, 39-72 (1988)

[4] Davenport, H.,: *Cubic forms in 32 variables* Phil. Trans. A, 251, 193-232

[5] Magyar, A.,: *L^p-bounds for spherical maximal operators on* $\mathbf{Z}^n$ Rev. Mat. Iberoam., 13, 307-317 (1997)

[6] Magyar, A., Stein, E.M., Wainger S.,: *Discerete analogues in harmonic analysis: spherical averages*, submitted to Annals. of Math.

[7] Serre, J.P.: *A course in arithmetic*, Graduate texts in Math., Springer Verlag, (0

[8] Sogge, C. : *Fourier integrals in classical analysis*, Cambridge University Press, (1993)

[9] Sogge, C., Stein, E.M.: *Averages of functions over hypersurfaces in* $\Re^n$, Invent. Math. 82, 543-556, (1985).

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

[11] Stein, E.M., Wainger, S: *Discerete analogues in harmonic analysis II: fractional integration*, Jour. d'Analyse, 80, 335-354 (2000)