Maximal Operators Associated to Discrete Subgroups of Nilpotent Lie Groups

by

Akos Magyar, Elias M. Stein, Stephen Wainger

§1. Introduction and statement of main theorem

The purpose of this paper is to prove a maximal theorem for averages taken over suitable discrete sub-varieties of nilpotent Lie groups.

The simplest instance of this result arises for the group of 3×3 strictly upper-triangular matrices, (the Heisenberg group, $\mathbb{H}^1$). In this example one considers averages for functions f defined on $\mathbb{Z}^3$ given by

$$(1.1)\quad \mathcal{A}_r(f)(a) = \frac{1}{r^2} \sum_{\substack{b=(b_1,b_2)\in\mathbb{Z}^2 \\ |b|<r}} f(a_1+b_1,\, a_2+b_2,\, a_3+a_1b_2), \text{ with } a=(a_1,a_2,a_3)\in\mathbb{Z}^3.$$

The assertion is then

$$\| \sup_{r>1} |\mathcal{A}_r(f)\, \|_{\ell^2(\mathbb{Z}^3)} \leq A \,\| f \,\|_{\ell^2(\mathbb{Z}^3)} \ .$$

We now formulate our general result and describe the background and some of the ideas of the proof.

a. The main theorem

We start with a simply connected nilpotent Lie group G of step k, with $k \geq 2$. Thus for its Lie algebra $\mathfrak{g}$ we have the descending central series $\mathfrak{g} = \mathfrak{i}_1 \supset \mathfrak{i}_2 \cdots \supset \mathfrak{i}_k \supset \mathfrak{i}_{k+1} = \{0\}$, of ideals $\mathfrak{i}_j$ with $\mathfrak{i}_{j+1} = [\mathfrak{g}, \mathfrak{i}_j]$, and $\mathfrak{i}_k \neq \{0\}$. We let N denote the sub-group of G corresponding to $\mathfrak{i}_k$ and set $G' = G/N$, with $\pi = G \longrightarrow G'$ the corresponding projection.

We also suppose we are given a uniform (i.e. discrete, co-compact) sub-group Γ of G, and denote $\Gamma' = \pi(\Gamma)$ the corresponding sub-group of G'.

We note that the ideals $\{\mathfrak{i}_j\}$ induce a homogeneous structure "at infinity" and a natural corresponding family of balls $\{\mathcal{B}_r\}_{r\geq 1}$, $\mathcal{B}_r \subset G$; we also set $\mathcal{B}'_r = \pi(\mathcal{B}_r)$, and $\mathcal{B}'_r \subset G'$.

Finally, we assume that we are given a polynomial mapping $\mathcal{R}$ of G' to G so that:

(1) $\pi\mathcal{R}=$ identity on G' (i.e. $\mathcal{R}$ is a section).

(2) $\mathcal{R} : \Gamma' \longrightarrow \Gamma$.

(3) for some $c > 0$, $\mathcal{R} : \mathcal{B}'_r \subset \mathcal{B}_{cr}$, for every $r \geq 1$.

Denote $B'_r = \mathcal{B}'_r \cap \Gamma'$ and let $|\mathcal{B}'_r| =$ the number of points in B_r. With these notions specified, we define for a function f on Γ the averages $\mathcal{A}_r(f)$ by

$$(1.2)\quad \mathcal{A}_r(f)(a) \;=\; \frac{1}{|B'_r|}\sum_{b\in B'_r} f(a\cdot\mathcal{R}(b))\;,\;\; a\in\Gamma\,.$$

Our result is then:

Theorem

$$(1.3)\quad \sup_{r\geq 1} \parallel\; |\mathcal{A}_r(f)|\;\parallel_{\ell^2(\Gamma)} \leq\; A\;\parallel f\parallel_{\ell^2(\Gamma)}.$$

b. Background

The proof of the theorem above depends in several aspects on that of a previous abelian analogue obtained by Bourgain [BO, 1-3]. We can formulate that result as follows. Suppose P is a polynomial on $\mathbb{R}$ with values in $\mathbb{R}^d$ that takes $\mathbb{Z}$ to $\mathbb{Z}^d$. Then if f is a function on $\mathbb{Z}^d$, the mapping

$$(1.4)\quad f \mapsto \sup_{r\geq 1} |\tfrac{1}{r}\sum_{n=1}^{r} f(m-P(n))|\,,\;\; m\in\mathbb{Z}^d$$

is bounded on $\ell^2(\mathbb{Z}^d)$ to itself, (as well as on ℓ^p, $1 < p \leq \infty$).

Another antecedent, obtained independently by Arkhipov and Oskolkov [AO], can be restated as a Hilbert transform analogue of (1.4). In effect, they showed that the mapping

$$(1.5)\quad f \mapsto \sum{}' \frac{f(m - P(n))}{n}$$

is bounded on $\ell^2(\mathbb{Z}^d)$.

Common to both the results in [BO] and [AO] is the use of techniques related to the circle method of Hardy-Littlewood-Ramanujan, ideas which will also be important for us.

After these two results it was natural to try to extend the scope of the theory to higher-dimensional sums and also in the non-translation-invariant setting, thus in this way to obtain discrete analogues of singular Radon transforms and the corresponding maximal functions. (For these operators, see [CNSW], and the model situation for nilpotent Lie groups in [RS].) In [SW] an ℓ^2 theory for a class of discrete analogues of singular Radon transforms was developed. Thus we may view the present paper as a first attack on the other branch of the question - that of the maximal theorem in this context.

c. Remarks about the proof

Here we want to discuss briefly two aspects of the proof that differ or go beyond the techniques used in the works mentioned above.

First one decomposes the circle, in the case of Example (1.1), (or the torus $\mathbb{T}^{d(k)}$, where $d(k)$ = dimension of $\mathfrak{i}_k$, in general) into "major arcs" or "minor arcs". In previous situations this was done, very roughly, according to whether the corresponding denominators q belonged to either of the two cases: $1 \le q \le r^\epsilon$, or $r^\epsilon < q \le r^{k-\epsilon}$. Here we shall need a tripartite decomposition: the "small" q, $1 \le q \le (\log r)^\gamma$; the "intermediate" q, $(\log r)^\gamma < q \le r^\epsilon$; and the "large" q, with $r^\epsilon < q$.

The range for small q is dictated by the need to approximate our operator by appropriate tensor products, and the resulting necessity in making maximal estimates of choosing common denominators for those q. These common denominators cannot be too large, requiring a condition like $q \le (\log r)^\gamma$. However, once q exceeds this bound, a similar but much simpler analysis allows us to treat each q at a time. This device works if q is not large ($q \le r^\epsilon$, for some small ϵ). For the remaining q, $r^\epsilon < q$, the method of Weyl sum estimates (i.e. their operator analogues as in [SW2]) are effective.

A second departure relates to the estimates of certain exponential-sum operators that occur. We describe this in a situation which is simpler than in the applications below, but yet illustrates the main point.

Consider the one-dimensional operator, S_r^θ, mapping $\ell^2(\mathbb{Z})$ to itself, given by

$$(1.6)\quad S_r^\theta(f)(m) = \frac{1}{r}\sum_{1\le n\le r} e^{2\pi i\theta mn} f(n)\,.$$

The trivial estimate for S_r^θ is of course $\| S_r^\theta \| = O(1)$, and as $r \longrightarrow \infty$ this is best that can be said as long as θ is "small", that is $|\theta| \le 1/r^2$. Also, when θ is "large", e.g. $|\theta| \approx 1$ one does obtain cancellation, $\| S_r^\theta \| = O(r^{-\delta})$ for some $\delta > 0$, when θ satisfies the usual arithmetic conditions arising in Weyl-sum estimates. However, beyond these standard phenomena, there is an intermediate range: here a (non-arithmetic) cancellation occurs and one has

$$(1.7)\quad \| S_r^\theta \| = O\left(1/|r^2\theta|^{1/2}\right), \text{ when say } 1/r^2 \le |\theta| \le 1/r^{2-\epsilon}.$$

d. Organization of the paper

Given the complexity of the proof of the theorem, we begin in Section 2 by sketching the proof in the special case corresponding to (1.1). This allows us to explain the basic scheme in this simpler case. The goal of Sections 2 to 6 is to reduce the abstract and general form of the averages $\mathcal{A}_r$ given by (1.2) to a concrete version given by (5.1). As a result, we get an operator on $\ell^2(\mathbb{Z}^d)$, with summation taken over $\mathbb{Z}^{d'}$ (with $d = \dim(G)$, $d' = \dim(G')$), and where all polynomials that occur have integer coefficients. An important fact here is the observation of a certain non-degeneracy of the polynomial P_k, which polynomial arises from the multiplication law. The reduction to the concrete picture is made possible by considering an appropriate Malcev basis for a rational structure of the Lie algebra $\mathfrak{g}$.

In Section 7 we decompose the averages $\mathcal{A}_r$ into three parts, corresponding to the division of the sizes of the denominators q mentioned above. Section 8 begins our analysis when q is small, and the corresponding grouping of terms with common denominator Q. The relevant maximal operator that arises is treated in Section 9 by the use of Lemma 9.1, which gives the estimate (1.7) in the general case. Section 10 deals with a Gauss-sum-like operator that arises in the tensor product decomposition.

Finally, Section 11 takes up the estimates for denominators q of intermediate size, and Section 12 for the rest.

We make some additional remarks.

1. It seems reasonable to suppose that it is possible to reformulate our result so as also to contain the full abelian case (1.4), when $p = 2$. However, we have not attempted this because it would not seem to shed further light on the (non-abelian) issues we have dealth with; also because our techniques at present do not give results for $p < 2$.

2. However one can prove the maximal theorem for averages (1.2) for all p, $1 < p$, in the special case when the group G is of Step 2. (See [IMSW]). This uses ideas in the recent work of [IW] that gives the ℓ^p boundedness, for $1 < p < \infty$, of the Hilbert transform (1.5) and its higher-dimensional abelian analogues.

3. The special case of the main theorem in which Γ is the group of strictly upper-triangular matrices with integer coordinates has been presented at several conferences by the authors, and has been announced in [M].

§2. A special case

Here we sketch the special case where G is the group of 3×3 the upper triangular matrices with ones on the main diagonal.

We let Γ be the discrete subgroup

$$\Gamma = \left\{ \alpha \in G : \alpha = \begin{pmatrix} 1 & a_1 & a_3 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

with a_1, a_2 and a_3 integers. We then put $\beta' = \begin{pmatrix} 1 & b_1 & 0 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix}$,

where b_1 and b_2 are integers. Our averaging operator is, for a special choice of $\mathcal{R}$, then

$$\mathcal{A}f(\alpha) \,=\, \frac{1}{r^2} \sum_{\substack{\beta'=(b_1,b_2)\\ |b_1|\le r\\ |b_2|\le r}} f(\alpha \,\cdot\, \beta')$$

for f a function on Γ and α a point in Γ. In this case our main theorem asserts the following estimate.

$$\| \sup_{r>0} |\mathcal{A}_r f| \;\|_{\ell^2(\Gamma)} \le \, C \;\| \, f \, \|_{\ell^2(\Gamma)} \;.$$

The outline of the proof of this special case is as follows.

Let ψ be a smooth function on $\mathbb{R}^2$ which is compactly supported and is one in a neighborhood of the origin. Put

$$\mathcal{M}_j \, f(\alpha) \,=\, \frac{1}{2^{2j}} \sum_{\beta'} f(\alpha \,\cdot\, \beta') \, \psi \, \left(\frac{\beta'}{2^j} \right)$$

where we identify the matrix β' with the point $\beta' = (b_1, b_2)$ in $\mathbb{Z}^2$. It then suffices to prove

$$\| \sup_{j} |\, \mathcal{M}_j f| \;\|_{\ell^2} \le \, C \;\| \, f \, \|_{\ell^2}. \tag{2.1}$$

We take the Fourier transform in the a_3 variable. That is we define

$$\hat{f}(a_1, a_2, \theta) \,=\, \sum_{a_3 \in Z} f(a_1, a_2, a_3) \, e^{2\pi i a_3 \theta} \,.$$

Then with

$$\alpha \,=\, (a_1, a_2, a_3) \,=\, (a, a_3)$$

$$\mathcal{M}_j \, f(\alpha) \,=\, \int_0^1 e^{-2\pi i a_3 \theta} \, S_j^\theta \, \hat{f}(\,\cdot, \theta)(a) \, d\theta \,. \tag{2.2}$$

Here S_j^θ takes functions on $\mathbb{Z}^2$ to functions on $\mathbb{Z}^2$, and is defined for w, a function on $\mathbb{Z}^2$, by

$$S_j^\theta\, w(a_1,a_2) \;=\; \frac{1}{2^{2j}} \sum_{b=(b_1,b_2)} \psi\left(\frac{b}{2^j}\right)\, e^{-2\pi i a_1 b_2 \theta}\, w(b_1+a_1, b_2+a_2)\,.$$

We want to analyze the integral in (2.2) by the circle method of Hardy, Littlewood and Ramanujan. To this end, choose a smooth even function $\chi(\theta)$ which is one for $-1 \le \theta \le 1$ and is supported in $-2 \le \theta \le 2$. We then write for $(\ell, q) = 1$, $1 \le \ell \le q$,

$$\mathcal{M}_j^{(\ell,q)}\, f(\alpha) \;=\; \int e^{-2\pi i a_3 \theta}\, \chi(2^{(2-\epsilon)j}(\theta - \frac{\ell}{q}))\, S_j^\theta\, \hat{f}(\cdot,\theta)\,(a)\, d\theta\,, \tag{2.3}$$

for a suitable small fixed ϵ.

We collect together those terms for which $2^\kappa \le q < 2^{\kappa+1}$. Set

$$\mathcal{M}_{j,\kappa} \;=\; \sum_{\substack{\ell,q \\ 2^\kappa \le q < 2^{\kappa+1} \\ 1 \le \ell \le q \\ (\ell,q)=1\,.}} \mathcal{M}_j^{(\ell,q)}\,.$$

Next we choose a number γ with $1 < \gamma < 2$. Then we put

$$M_j^1 \;=\; \sum_{2^\kappa \le j^\gamma} \mathcal{M}_{j,\kappa}\,, \tag{2.4}$$

and

$$M_j^2 \;=\; \sum_{j^\gamma \le 2^\kappa < 2^{\epsilon j}} \mathcal{M}_{j,\kappa}. \tag{2.5}$$

Then

$$\mathcal{M}_j \;=\; M_j^1 \,+\, M_j^2 \,+\, E_j. \text{ Here} \tag{2.6}$$

$$E_j(\alpha) \;=\; \int_{\mathcal{E}_j} e^{-2\pi i a_3 \theta}\, B(\theta)\, S_j^\theta \hat{f}\,(\,\cdot\,,\,\theta)(a)\, d\theta$$

where $B(\theta) = B_j(\theta)$ is a function bounded informly in j, while each θ in $\mathcal{E}_j$ has an approximation by a rational ℓ/q with $(\ell, q) = 1$ such that

$$\left|\theta - \frac{\ell}{q}\right| \leq \frac{1}{q} 2^{-(2-\epsilon)j}\,,\ 2^{\epsilon j} \leq q \leq 2^{(2-\epsilon)j}\,.$$

This can be seen as follows. Recall Dirichlet's approximation, whereby for any real θ and positive integer N, there is a fraction ℓ/q, with $(\ell, q) = 1$ (or $\ell = 0$, and $q = 1$), such that

$$|\theta - \ell/q\,| \leq \frac{1}{qN}\,,\ \text{ and } \ 1 \leq q \leq N\,.$$

We now apply this with N the greatest positive integer not exceeding $2^{(2-\epsilon)j}$, and consider the intervals that arise in this way divided in two cases: either with $q \leq 2^{\epsilon j}$, or with $2^{\epsilon j} < q \leq 2^{(2-\epsilon)j}$. Note that in the former case, any two such intervals are *disjoint* for large j, because if this were not the case, we would have $|\ell_1/q_1 - \ell_2/q_2| \leq 2 \cdot 2^{-(2-\epsilon)j}$. But then $\frac{1}{q_1 q_2} \leq |\ell_1/q_1 - \ell_2/q_2| \leq 2 \cdot 2^{-(2-\epsilon)j}$, which contradicts the fact that $q_1 \leq 2^{\epsilon j}$ and $q_2 \leq 2^{\epsilon j}$. (Here we use the fact that $2^{2\epsilon j} \ll 2^{(2-\epsilon)j}$, which happens when $\epsilon < 2/3$.) Now the intervals in M_j^1 are those for $q \leq j^\gamma$, and the intervals for M_j^2 are those for $j^\gamma < q \leq 2^{\epsilon j}$. The complement in $[0,1]$ of the union of the intervals arising for both M_j^1 and M_j^2 is the set $\mathcal{E}_j$. Also $B(\theta) = 1 - \sum \chi(2^{(2-\epsilon)j}(\theta - \ell/q))$.

We next point out that for $\theta \,\epsilon\, \mathcal{E}_j$, we have the "Weyl-sum" estimate

$$\| S_j^\theta \| \leq C 2^{-\eta j} \tag{2.7}$$

for some $\eta > 0$.

The estimate (2.7) is contained in [SW3], but it can also easily be seen directly; (see also Section 12).

Given (2.7) we may apply Plancherel's theorem to see

$$\| E_j \| \leq C\, 2^{-\eta j}\,.$$

Then a square function argument ($\sup_j |E_j\, f| \leq (\sum_j |E_j\, f|^2)^{1/2}$) disposes of $\sup_j |E_j|$.

Let us turn to the most difficult term in (2.6), M_j^1. To treat this term, we will prove

$$(2.8)\quad \| \sup_{\substack{j \\ 2^\kappa \le j^\gamma}} |\mathcal{M}_{j,\kappa} f| \| \le C 2^{-\eta\kappa} \| f \|$$

for some $\eta > 0$. To treat the various denominators q, $2^\kappa \le q < 2^{\kappa+1}$, that arise in $\mathcal{M}_{j,\kappa}$, it would be convenient to have a common denominator and for this reason we begin by taking $Q = \prod_{2^\kappa \le q < 2^{\kappa+1}} q$.

Next it is easy to see that the subgroup $Q\Gamma$ of Γ is a normal subgroup. ($Q\Gamma$ consists of those $\alpha \in \Gamma$ whose coordinates are multiples of Q). For any $\alpha \in \Gamma$, we may write $\alpha = (\mu Q) \cdot \sigma$ where the coordinates of σ are between 0 and $Q-1$. We then wish to write $\mathcal{M}_{j,\kappa} = \mathcal{N}_j \circ H +$ small error, where $\mathcal{N}_j$ is an operator acting on the μQ variables, and H acts on the σ variables. In order to approximate $\mathcal{M}_{j,\kappa}$ by $\mathcal{N}_j \circ H$, it would be necessary to know that for any $\eta > 0$, $Q < C_\eta 2^{\eta j}$. Unfortunately, this is not the case. So, following Bourgain [BO1], we divide the q's with $2^\kappa \le q < 2^{\kappa+1}$ into groups $\mathcal{F}$, with each group containing at most $2^{\delta\kappa}$ elements where $\frac{1}{2} < \delta < 1/\gamma$. If we now define $Q = \prod_{q \in \mathcal{F}} q$ then one can see that for any $\eta > 0$

$$Q = \prod_{q \epsilon \mathcal{F}} q < 2^{(\kappa+1)\, 2^{\delta\kappa}} < (j^{\gamma+1})^{j^{\gamma\delta}} < 2^{C(\log j)\, j^{\gamma\delta}} < C_\eta\, 2^{\eta j}$$

since $\gamma\delta < 1$. Also the number of such groups we need to cover all q's with $2^\kappa \le q < 2^{\kappa+1}$ is at most $2^{(1-\delta)\kappa}$. So to obtain the estimate (2.8) it will suffice to obtain the estimate

$$(2.9)\quad \| \sup_{\substack{j \\ 2^\kappa \le j^\gamma}} \mathcal{M}^{\mathcal{F}}_{j,\kappa} f \| \le \frac{C}{2^{\kappa/2}} \| f \|,$$

where

$$(2.10)\quad \mathcal{M}^{\mathcal{F}}_j = \sum_{q \epsilon \mathcal{F}} \sum_{\substack{a=1 \\ (u,q)}}^{q} \mathcal{M}^{(\ell,q)}_j .$$

Note that in the integral (2.3) for $\mathcal{M}^{(\ell,q)}_j$ we may insert a factor $\chi(c2^{2\kappa}(\theta - \frac{\ell}{q}))$ for c a large fixed constant, since $2^\kappa \lesssim j^\gamma$. Thus in the expression for $\mathcal{M}^{(\ell,q)}_j f$ we may replace f by $f^{(\ell,q)}$ where

$$\hat{f}^{(\ell,q)}(\theta) = \chi\left(c2^{2\kappa}\left(\theta - \frac{\ell}{q}\right)\right) \hat{f}(\theta) .$$

The advantage is that the functions $\hat{f}^{(\ell,q)}(\theta)$ have disjoint support, and thus the $f^{(\ell,q)}$ are orthogonal at least for $2^\kappa \le q < 2^{\kappa+1}$; the disjointness of the supports is due to the fact that $\frac{1}{q_1 q_2} \le (c2^{2\kappa})^{-1}$.

Now in the integral (2.3) we expand the periodic extension of $\chi(2^{(2-\epsilon)j}(\theta - \frac{\ell}{q}))$ in a Fourier series. So

$$\chi(2^{(2-\epsilon)j}(\theta - \frac{\ell}{q})) \;=\; \sum_{b_3} \frac{1}{2^{(2-\epsilon)j}}\, \widehat{\chi}\left(\frac{b_3}{2^{(2-\epsilon)j}}\right)\, e^{2\pi i b_3(\epsilon - \frac{\ell}{q})}$$

where $\hat{\chi}$ is the Fourier transform of χ on R. We also expand $\hat{f}(b_1, b_2, \theta)$ in its Fourier series and then perform the θ integration. If we identify $(b_1, b_2, b_3) = (b, b_3)$ with the matrix

$$\beta \;=\; \begin{pmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix}\,,$$

we find

$$\mathcal{M}_j^{(\ell,q)}\, f(\alpha) \;=\; \frac{1}{2^{2j}2^{(2-\epsilon)j}} \sum_{\beta=(b,b_3)} e^{-2\pi i b_3 \frac{\ell}{q}}\, \psi\left(\frac{b}{2^j}\right)\, \widehat{\chi}\left(\frac{b_3}{2^{(2-\epsilon)j}}\right)\, f^{(\ell,q)}\,(\alpha\cdot\beta)\,.$$

With Q as above, we write

$$\alpha \;=\; \mu Q\,\cdot\,\sigma$$

and

$$\beta \;=\; \nu Q\,\cdot\tau\,.$$

We use the notation

$$\begin{aligned} \mu \quad &= \quad (m_1, m_2, m_3) \quad = \quad (m, m_3)\,, \\ \sigma \quad &= \quad (s_1, s_2, s_3) \quad = \quad (s, s_3)\,, \\ \nu \quad &= \quad (n_1, n_2, n_3) \quad = \quad (n, n_3) \end{aligned}$$

and

$$\tau \quad = \quad (t_1, t_2, t_3) \quad = \quad (t, t_3)\,.$$

We will then have $0 \le s_j \le Q-1$ and $0 \le t_j \le Q-1$. Since $q|Q$, the factor $e^{-2\pi i b_3 \frac{\ell}{q}}$ equals $e^{-2\pi i t_3 \frac{\ell}{q}}$. This will be essential in approximating $\mathcal{M}_j^{\mathcal{F}}$ by $\mathcal{N}_j \,\cdot\, H$.

Now with a little care, we can show that except for a small error

$$\mathcal{M}_j^{\mathcal{F}}(\mu Q\cdot\sigma) = \frac{1}{2^{2j}}\,\frac{1}{2^{(2-\epsilon)j}} \sum_{\nu Q=(nQ,n_3Q)} \psi\left(\frac{nQ}{2^j}\right)\, \widehat{\chi}\left(\frac{n_3Q}{2^{(2-\epsilon)j}}\right)\cdot \sum_{\substack{\ell,q \\ \tau=(t,t_3)}} e^{-2\pi i t_3 \frac{\ell}{q}}\, f^{(\ell,q)}\,(\mu Q\cdot\nu Q\cdot\{\sigma\cdot\tau\})$$

where $\{\sigma \cdot \tau\}$ is in the residue class of $(\sigma \cdot \tau) \bmod Q$. That is, $\sigma \cdot \tau \equiv \{\sigma \cdot \tau\} \bmod Q$ and more particularly the coordinates of $\{\sigma\tau\}$ are between 0 and $Q-1$. To control the error, as we have said before, it is crucial to use the fact that for any $\eta > 0$, $Q \leq C_\eta 2^{\eta j}$. The analysis of the error is carried out in the general case at end of Section 8.

Once this error is controlled, we have succeeded in getting, up to a small error that

$$\mathcal{M}_j^{\mathcal{F}}(\mu a \cdot \sigma) = \mathcal{N}_j \cdot H^\sigma ,$$

where for F a function on $Q\Gamma$,

$$\mathcal{N}_j F(\mu Q) = \sum_{\nu Q} \frac{1}{2^{2j}} \frac{1}{2^{(2-\epsilon)j}} \psi\left(\frac{nQ}{2^j}\right) \widehat{\chi}\left(\frac{n_3 Q}{2^{(2-\epsilon)j}}\right) F(\mu Q \cdot \nu Q) ,$$

and

$$H^\sigma(\mu Q) = \sum_{\tau=(t,t_3)} \sum_{q \epsilon \mathcal{F}} \sum_{\substack{\ell=1 \\ (\ell,q)=1}}^{q} e^{-2\pi i t_3 \frac{\ell}{q}} f^{(\ell,q)}(\mu Q \cdot \{\sigma \cdot \tau\}) .$$

This plays the role of a crucial "tensor product" decomposition. Then to obtain the estimate (2.9), it will suffice to prove

$$\| \sup_j \mathcal{N}_j F \|_{\ell^2(Q\Gamma)} \leq \frac{C}{Q^3} \| F \|_{\ell^2(Q\Gamma)} \tag{2.11}$$

and

$$\sum_\sigma \sum_{\mu Q} |H^\sigma(\mu Q)|^2 \leq \frac{CQ^6}{2^\kappa} \| f \|^2_{\ell^2(\Gamma)}. \tag{2.12}$$

Let us begin with the estimate (2.11). Let

$$\mathcal{B}_j = \{g \in G : |x_1| < 2^j , |x_2,| < 2^j , |x_3| < 2^{2j}\} .$$

It is well known that (group) translates of the balls $\mathcal{B}_r$ form a standard Vitali family of balls on G. It follows that translates of the balls on $Q\Gamma$ defined by

$$B_j^Q = \{\mu Q \in Q\Gamma : |m_1 Q| < 2^j , |m_2 Q| < 2^j , |m_3 Q| < 2^{2j}\}$$

forms a standard Vitali family of balls on $Q\Gamma$ with all constants uniform in Q, since $B_j^Q = \mathcal{B}_j \cap Q\Gamma$, and $|B_j^Q| \sim \frac{2^{4j}}{Q^3}$. It therefore follows by a well known argument that

$$\| \sup_j |\mathcal{N}_j^0 F| \|_{\ell^2(Q\Gamma)} \leq \frac{C}{Q^3} \| F \|_{\ell^2(Q\Gamma)} .$$

So we must now consider the difference $\mathcal{N}_j - \mathcal{N}_j^0$. To obtain (2.11) we have to show

$$\| \sup_j |\mathcal{N}_j F - \mathcal{N}_j^0 F| \| \leq \frac{C}{Q^3} \| F \| .$$

We will estimate $\mathcal{N}_j - \mathcal{N}_j^0$ by using Fourier analysis on the group $Q\mathbb{Z}$, and we shall see the gain of 2^κ in (2.12) will come from an oscillation arising due to the "twist" of the group multiplication. For F defined on QZ, we set

$$\hat{F}(\theta) = \sum F(nQ)\, e^{2\pi i n Q\theta} .$$

Then

$$F(nQ) = Q \int_{-\frac{1}{2Q}}^{\frac{1}{2Q}} e^{-2\pi i n Q\theta} \hat{F}(\theta) d\theta ,$$

and

$$\sum |F(nQ)|^2 = Q \int_{-\frac{1}{2Q}}^{1/2Q} |\hat{F}(\theta)|^2 d\theta .$$

In particular, we point out that

$$\frac{1}{2^{(2-\epsilon)j}} \hat{\chi}\left(\frac{n_3 Q}{2^{(2-\epsilon)j}}\right) = \int_{-\frac{1}{2Q}}^{\frac{1}{2Q}} e^{-2\pi i n_3 Q\theta} \chi(2^{(2-\epsilon)j}\theta)\, d\theta .$$

Thus the $Q\mathbb{Z}$ Fourier transform of $\frac{1}{2^{(2-\epsilon)j}} \hat{\chi}\left(\frac{n_3 Q}{(2^{2-\epsilon)j}}\right)$ equals $\frac{1}{Q}\chi(2^{(2-\epsilon)j}\theta)$. Now

$$\mathcal{N}_j F(\mu Q) = Q \int_{-\frac{1}{2Q}}^{\frac{1}{2Q}} e^{-2\pi 2 m_3 Q\theta} \frac{\chi(2^{(2-\epsilon)j}\theta)}{Q} S_j'^{\theta} \hat{F}(mQ, \theta) d\theta ,$$

where for a function w defined on $Q\mathbb{Z}$

$$S_j'^{\theta} w(mQ) = \frac{1}{2^{2j}} \sum_{nQ} \psi\left(\frac{nQ}{2^j}\right) e^{-2\pi i \theta m_1 n_2 Q^2} w(mQ + nQ) .$$

We will show that for $2^{-2j} \leq \theta \leq 2^{-(2-\epsilon)j}$,

(2.13) $\| S_j^{\prime\theta} \| \leq C \left(\frac{1}{Q^2 |2^{2j}\theta|^{1/2}} \right)$

Let us assume (2.13). Then

$$\sum_{m_3 Q} \left| \left(\mathcal{N}_j - \mathcal{N}_j^0 \right) F(\mu Q) \right|^2$$

$$= \frac{1}{Q} \int |\chi(2^{2-\epsilon)j}\theta) - \chi(2^{2j}\theta)|^2 \, |S_j^{\prime\theta} \hat{F}(mQ, \theta)|^2 d\theta \, .$$

So

$$\sum_{\mu Q} |(\mathcal{N}_j - \mathcal{N}_j^0) F(\mu Q)|^2 \leq$$

$$\frac{C}{Q^5} \int |\frac{1}{|2^{2j}\theta|} \left(\chi \left(2^{(2-\epsilon)j}\theta \right) - \chi \left(2^{2j}(\theta) \right) \right|^2 \sum_{mQ} |\hat{F}(mQ, \theta)|^2 \, d\theta \, .$$

Thus

$$\sum_j |(\mathcal{N}_j - \mathcal{N}_j^0) F(\mu Q)|^2$$

$$\leq C \frac{1}{Q^6} \sum_{\mu Q} (F(\mu Q)|^2 \, .$$

This is because the sum $\sum \frac{1}{|2^{2j}\theta|}$ is bounded as j ranges over $2^{2j}|\theta| \geq c$.

Thus by using a square function argument, we arrive at (2.11). So the estimate (2.13) implies (2.11), and we turn to a sketch of the proof of (2.13). To simplify the writing we shall assume $Q = 1$ here. In proving (2.13), it suffices to consider functions w that are supported in a ball of radius 2^j, $\{n : |n - k| < 2^j\}$, for an arbitrary center k.

Thus we replace the free variable m by $m + k$ and we are led (up to a factor of $e^{2\pi i \theta m_2 k_1}$) to estimating the sum

$$\frac{1}{2^{2j}} \sum_n \psi(n/2^j) \, e^{-2\pi i \theta m_1 n_2} \, w_k(m + n)$$

where $w_k(u) = e^{-2\pi i k_1 u_2} \, w(u + k)$, and now $w_k(u)$ is supported in the ball $|u| < 2^j$, and this restricts m to the ball $|m| < 2 \cdot 2^j$.

We can then compare the above sum with the integral

$$\frac{1}{2^{2j}} \int \psi(y/2^j) \, e^{-2\pi i \theta x_1 y_2} \, f(x + y) \, dy \, ,$$

where f is defined by $f(y) = w_k(n)$ if y belongs to the unit cube centered at n.

Since $|y| \leq 2^j$ and $|x| < 2 \cdot 2^j$ the error committed by this comparison gives a bound $O(\theta 2^j) + O(2^{-j})$, the first coming from the variation over each cube of $e^{-2\pi i \theta x_1 y_2}$, and the second from the variation of $\psi(y/2^j)$. Moreover, by rescaling and using stationary phase one can observe that the integral operator has a norm which is $O(|\theta 2^{2j}|^{-1/2})$. Altogether, this gives (2.13), since $\theta = O(2^{-(2-\epsilon)j})$. (In the setting of more general nilpotent Lie groups as in Section 9, where the multiplicative twist is not bilinear, it is harder to approximate $S_j^{\prime\theta}$ by an integral operator. However, this approximation can be made instead for the operator $(S_j^{\prime\theta})\ (S_j^{\prime\theta})^*$.)

We turn to the proof of (2.12), that is the estimate for $H^\sigma(\mu Q)$. After a change of variables, we see

$$H^\sigma(\mu Q) \ = \sum_{\substack{\tau = (t_1,t_2,t_3) \\ 0 \leq t_j \leq Q-1 \\ \ell,q}} e^{-2\pi i (t_3 - s_3 - s_1 t_2)\, \frac{\ell}{q}}\, f^{\ell,q}\,(\mu Q \cdot \tau)\,.$$

For w a function on $\mathbb{Z}^2/Q$, define

$$Tw(s_1,s_2) \ = \sum_{t_1=0}^{Q-1} \sum_{t_2=0}^{Q-1} e^{2\pi i s_1 t_2\, \frac{\ell}{q}}\, w(t_1,t_2)\,.$$

The main step in establishing (2.11) is to show

$$(2.14)\ \ \| Tw \|^2_{\ell^2(\mathbb{Z}^2/Q)} \leq \tfrac{Q^4}{q}\ \| w \|^2_{\ell^2(\mathbb{Z}^2/Q)}.$$

This is an operator "Gauss sum" estimate of the kind considered in [SW3].

Granting (2.14), we see

$$\sum_{s_3} |H^\sigma(\mu Q)|^2 \ \leq\ Q \sum_{\ell,q} \left| \sum_{\tau} e^{-2\pi i (t_3 - s_1 t_2)\, \frac{\ell}{q}}\, f^{(\ell,q)}\,(\mu Q, \tau) \right|^2$$

$$\leq\ Q^2 \sum_{\ell,q} \sum_{t_3=1}^{Q} \left| \sum_{t_1=1}^{Q} \sum_{t_2=1}^{Q} e^{2\pi i t_2 s_1\, \frac{\ell}{q}}\ f^{(\ell,q)}\,(\mu Q \cdot \tau) \right|^2 .$$

So using (2.14), we find

$$\sum_{\tau} |H^\sigma(\mu Q)|^2 \ \leq\ \frac{Q^6}{q} \sum_{\tau} |f^{\ell,q}(\mu Q \cdot \tau)|^2\,.$$

Thus

$$\sum_{\sigma,\mu Q} |H^\sigma(\mu Q)|^2 \;\leq\; C\,\tfrac{Q^6}{2^\kappa} \sum_{\mu Q,\tau} |f^{(\ell,q)}(\mu Q\,\cdot\,\tau)|^2$$

$$\leq\; C\,\tfrac{Q^6}{2^\kappa}\;\|\; f^{(\ell,q)}\;\|^2$$

Since $|\sum \hat{f}^{(\ell,q)}| \leq C\,|\hat{f}|$ and the $\hat{f}^{\ell,q}$ have disjoint supports, we arrive at (2.12).

To prove (2.14), let us denote by $k(s,s')$ the kernel of TT^*. Then

$$\sum_s |k(s,s')| \leq \sum_{s_1=0}^{Q-1}\sum_{s_2=0}^{Q-1} \left| \sum_{t_1=0}^{Q-1}\sum_{t_2=0}^{Q-1} e^{2\pi i(s_1-s_1')\,t_2\,\frac{Q}{q}} \right|$$

$$= \frac{Q^4}{q^4} \sum_{s_1=0}^{q-1}\sum_{s_2=0}^{q-1} \left| \sum_{t_1=0}^{q-1}\sum_{t_2=0}^{q-1} e^{2\pi i(s_1-s_1')\,t_2\,\frac{Q}{q}} \right| .$$

The sum on t_2 is 0 unless $s_1 = s_1'$.

Thus

$$\sum_s |k(s,s')| \leq \frac{Q^4}{q}\,.$$

Similarly

$$\sum_{s'} |k(s,s')| \leq \frac{Q^4}{q}\,.$$

This gives (2.14) and completes our discussion of M_j.

The treatment of M_j^2 is similar to but much easier than the argument for M_j^1. In discussing M_j^2 it suffices to show that for each κ with $j^\gamma < 2^\kappa < 2^{\epsilon j}$,

$$\|\; \mathcal{M}_{j,\kappa}\;\| \leq\; C2^{-\kappa/2}$$

uniformly in j. In fact, then $\|\sup_{j^\gamma<2^\kappa} |\mathcal{M}_{j,\kappa}|\;\|^2 \leq \sum_{j^\gamma<2^\kappa} \|\;\mathcal{M}_{j,\kappa}\;\|^2 \leq c\sum_{j=1}^{\infty} j^{-\gamma} < \infty$, since $\gamma > 1$.

§3. Some preliminaries on uniform subgroups, canonical coordinates and corresponding balls

We now turn to the general situation.

The main facts we need to know about uniform subgroups are summarized in the following proposition. For the background material needed we shall refer to [CG].

Proposition 3.1: *Let Γ be a uniform subgroup of a simply connected nilpotent Lie group G. Then $\mathfrak{g}$, the Lie algebra of G has a basis*

$$X_1^1, \ldots X_1^{d(1)}, \ldots, X_k^1, \ldots, X_k^{d(k)}$$

such that

(a)

$$X_u^1, \ldots, X_u^{d(u)}, \ldots, X_k^1, \ldots, X_k^{d(k)}$$

is a basis for the ideal $\mathfrak{i}_u$.

(b)

$$\Gamma = \Big\{\alpha \in G : \alpha = \exp(a_1^1 X_1^1) \cdot \exp(a_1^2 X_1^2) \ldots \exp(a_1^{d(1)} X_1^{d(1)}) \ldots$$

$$\ldots \exp(a_k^1 X_k^1) \ldots \exp(a_k^{d(k)} X_k^{d(k)})$$

where the a_u^v are integers$\}$.

(c) $\left[X_u^v, X_{u'}^{v'}\right]$ *is a finite linear combination of the X_u^v with rational coefficients.*

The ideals $\mathfrak{i}_u$ are defined in §1 and are the span of all commutators in $\mathfrak{g}$ of length $\geq u$.

In particular, if $\mathfrak{s}_u$ is the subspace of $\mathfrak{g}$ spanned by $X_u^1, X_u^2, \cdots X_u^{d(u)}$, then the ideal $\mathfrak{i}_u$ equals $\mathfrak{s}_u \oplus \mathfrak{s}_{u+1} \cdots \oplus \mathfrak{s}_k$; also $d(u)$ is the dimension of $\mathfrak{s}_u$.

The proof of Proposition 3.1 is contained in [CG]. For the convenience of the reader, we restate the definitions and results in [CG] which imply this proposition.

Let $\mathfrak{g}$ be a real nilpotent Lie algebra. A basis $\{X_1, \cdots X_n\}$ for $\mathfrak{g}$ is called a strong Malcev basis if for each m, $1 \leq m \leq n$, the span of $\{X_1, \cdots X_m\}$ is an ideal of $\mathfrak{g}$.

Suppose $\mathfrak{h}_1 \subseteq \mathfrak{h}_2 \subset \cdots \subset \mathfrak{h}_k = \mathfrak{g}$ is a nested sequence of ideals in $\mathfrak{g}$, and that $\{X_1, \cdots X_n\}$ is a basis for $\mathfrak{g}$. We say $\{X_1, \cdots X_n\}$ is a strong Malcev basis for $\mathfrak{g}$ passing through $\mathfrak{h}_1, \cdots \mathfrak{h}_k$ if

$\{X_1, \cdots X_n\}$ is a strong Malcev basis for $\mathfrak{g}$, and for each j, $1 \leq j \leq k$, there is a positive integer m_j such that $\{X_1, \cdots X_{m_j}\}$ is a basis for $\mathfrak{h}_j$. Theorem 1.13(b) asserts that for any nested sequence of ideals $\mathfrak{h}_1 \subseteq \mathfrak{h}_2 \subseteq \cdots \subseteq \mathfrak{h}_k = \mathfrak{g}$, there is a strong Malcev basis passing through $\mathfrak{h}_1, \cdots \mathfrak{h}_k$. (See also the second note after the proof of Theorem 1.13(b) in [CG].)

To say that $\mathfrak{g}$ has a rational structure means that $\mathfrak{g}$ has a basis $\{X_1, \cdots X_n\}$ such that $[X_i, X_j] = \Sigma\, C_{i,j,k} X_k$ with $C_{i,j,k}$ rational. Then $\mathfrak{g}_Q =$ span of $\{X_1, \cdots X_n\}$ over the rationals is called a rational structure for $\mathfrak{g}$. (See Lemma 5.11 (a)). Let $\mathfrak{h}$ be a subalgebra of $\mathfrak{g}$. (For us $\mathfrak{h}$ will be an ideal). Let $\mathfrak{g}_Q$ be a fixed rational structure, and put $\mathfrak{h}_Q = \mathfrak{h} \cap \mathfrak{g}_Q$. Then $\mathfrak{h}$ is said to be rational if the span of $\mathfrak{h}_Q$ over the reals is $\mathfrak{h}$. (This definition is given between Lemmas 5.1.1 and 5.1.2 in [CG]). Let G be a connected, simply connected, nilpotent Lie group, and suppose H is a closed subgroup. H is said to be rational if the Lie algebra of $\mathfrak{h}$ is a rational subalgebra of $\mathfrak{g}$.

Assume G is a connected, simply connected Lie group with Lie algebra $\mathfrak{g}$, and suppose Γ is a discrete subgroup of G. A strong Malcev basis $\{X_1, \cdots X_n\}$ of $\mathfrak{g}$ is said to be strongly based on Γ if

$$\Gamma = \{y \in G : \, y \, = \, \exp m_1 X_1 \cdot \exp m_2 X_2 \cdots \exp m_n X_n \,, \text{ with } m_j \text{ integers}\}$$

Theorem 5.1.6 in [CG] asserts, among other things, that if Γ is a uniform subgroup of G, $\mathfrak{g}$ has a strong Malcev basis strongly based on Γ. In the proof of Theorem 5.1.8 (a), it is shown that for any such basis $[X_i, X_j] = \Sigma\, C_{i,j,k} X_k$ with $C_{i,jk}$ rational. Thus any such basis determines a rational structure, $\mathfrak{g}_Q$, for $\mathfrak{g}$. Namely, $\mathfrak{g}_Q$ is the rational span of $\{X_i \cdots X_n\}$. Denote the inverse of the exponential mapping from $\mathfrak{g}$ to G by log. Then by the Campbell-Hausdorff formula $\log \gamma \in \mathfrak{g}_Q$ for every $\gamma \in \Gamma$. Also, for any such basis $\exp X_j \in \Gamma$. Thus, $\mathfrak{g}_Q$ is the rational span of $\log \gamma$, $\gamma \in \Gamma$, so $\mathfrak{g}_Q$ is determined by Γ. We call this $\mathfrak{g}_Q$ the rational structure of $\mathfrak{g}$ determined by Γ. Corollary 5.2.2 asserts that the ideals $\mathfrak{i}_u$ of the descending central series of $\mathfrak{g}$ are rational subalgebras, or equivalantly, if $\{\mathfrak{i}_u\}$ are the closed normal subgroups of $\mathfrak{g}$ corresponding to $\{\mathfrak{i}_u\}$, the $\{\mathfrak{i}_u\}$ are rational subgroups. Thus Proposition 3.1 is implied by the following statement contained in Proposition 5.3.2 of [CG].

Statement: Let Γ be a uniform subgroup of a connected, simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$. Let $H_1 \underset{\neq}{\subset} H_2 \underset{\neq}{\subset} \cdots \underset{\neq}{\subset} H_k = G$ be closed normal subgroups of G. Denote the Lie algebra of H_j by $\mathfrak{h}_j$, $1 \leq j \leq k$. Suppose the H_j are rational subgroups of $\mathfrak{g}$ (with respect to the rational structure determined by Γ). Then there exists a strong Malcev basis for $\mathfrak{g}$ strongly

based on Γ passing through $\mathfrak{h}_j$, $1 \le j \le k$.

We now fix some important notation. With $d(j)$ as in the statement of Proposition 3.1, then $d = d(1) + \cdots + d(k)$ is the dimension of $\mathfrak{g}$, and $d' = d(1) + \cdots + d(k-1)$ is the dimension of $\mathfrak{g}' = \mathfrak{g}/\mathfrak{i}_k$. We also set

$$D \;=\; d(1) + 2d(2) + \cdots + kd(k)$$

and

$$D' \;=\; d(1) + 2d(2) + \cdots + (k-1)\, d(k-1)\,.$$

D and D' are called the homogeneous dimensions of G and G' respectively.

For g and h in G we write $g \cdot h$ or gh for their product. If g' and h' are elements of G', we use the notation $g' \circ h'$ for their product in G'.

Suppose $x_u = (x_u^1, \ldots x_u^{d(u)})$ is a point in $\mathbb{R}^{d(u)}$. We write X_u for $(X_u^1, \ldots X_u^{d(u)})$, and $x_u \cdot X_u$ for $x_u^1 X_u^1 + \cdots + x_u^{d(u)} X_u^{d(u)}$; also we set

$$\exp(x_u * X_u) \;=\; \exp(x_u^1 X_u^1) \,\cdot\, \exp(x_u^2 X_u^2) \,\ldots\, \exp(x_u^{d(u)} X_u^{d(u)})\,.$$

Since G is a simply connected nilpotent Lie group, every $g \in G$ can be written uniquely in the form

$$(3.1)\quad g \;=\; \exp(z_1 \cdot X_1 + \cdots + z_k \,\cdot\, X_k)$$

with $z_u \in \mathbb{R}^{d(u)}$. The $z_1, \ldots z_k$ are called canonical coordinates of the first kind.

In view of part (b) of proposition (3.1) it is useful to see that each g in G can be written in the form

$$(3.2)\quad g \;=\; \exp(x_1 * X_1)\, \exp(x_2 * X_2) \,\ldots\, \exp(x_k * X_k)$$

and to try to clarify the relation between the x's in (3.2) and the z's in (3.1). The x's in (3.2) are called canonical coordinates of the *second kind*.

For this purpose it is convenient to introduce the notion of the homogeneous degree of a polynomial, and consider the associated $\mathbb{R}^d$ and its corresponding homogeneity.

Suppose we are given a gradation of $\mathbb{R}^d$. That is, each $x \in \mathbb{R}^d$ can be written $x = (x_1, \dots x_k)$ with each $x_u = (x_u^1, \dots x_u^{d(u)})$ a point in $\mathbb{R}^{d(u)}$. If $e_u = (e_u^1, \dots e_u^{d(u)})$ with each e_u^v a non-negative integer, we write $x_u^{e_u}$ for $(x_u^1)^{e_u^1} \dots (x_u^{d(u)})^{e_u^{d(u)}}$. Also, we put

$$|e_u| \,=\, e_u^1 \,+ \cdots +\, e_u^{d(u)}\,.$$

If $\mathcal{P}(x)$ is a monomial so that

$$\mathcal{P}(x) \,=\, A\, x_1^{e_1} \,\cdots x_k^{e_k}$$

with $A \neq 0$, we say that $\mathcal{P}$ has *homogeneous degree* j if $|e_1| \,+\, 2|e_2| + \cdots k|e_k| \,=\, j$. If

$$\mathcal{P}(x) \,=\, \sum_{\xi} M_\xi(x)$$

is a finite sum of non-zero distinct monomials we say that the homogeneous degree of $\mathcal{P}$ is the maximum homogeneous degree of the $M_\xi(x)$. We write the non-isotropic dileation $x \longrightarrow \lambda \circ x$ as $\lambda \circ x = (\lambda x_1, \lambda^2 x_2, \cdots, \lambda^k x_k)$.

We collect some useful facts about homogeneous degrees of polynomials in the following lemma.

Lemma 3.2:

(a) *Suppose $\mathcal{P}(x_1, \dots x_k)$ is a polynomial of homogeneous degree j with $j < k$. Then $\mathcal{P}$ is a function of $x_1, \dots x_j$.*

(b) *If $\mathcal{P}(x_1, \dots x_k)$ is a polynomial of homogeneous degree j, then for every x, $\mathcal{P}(\lambda \circ x)$ is a polynomial in λ of degree at most j.*

(c) *Let $\mathcal{P}(x_1, \dots x_k)$ be a polynomial of homogeneous degree j. Then for at least one x, with $|x_u^v| \leq 1$, for all u and v, $\mathcal{P}(\lambda \circ x)$ is a polynomial in λ of degree j.*

(d) *Assume $\mathcal{P}(x_1, \dots x_k)$ is a polynomial of homogeneous degree j. Let*

$$\mathcal{Q}_u \,=\, \left(\mathcal{Q}_u^1, \dots \mathcal{Q}_u^{d(u)}\right), 1 \,\leq\, u \,\leq\, k$$

and set

$$\mathcal{P}'(x) \,=\, \mathcal{P}(\mathcal{Q}_1(x), \dots \mathcal{Q}_k(x))\,.$$

Then if each $\mathcal{Q}_u^v$ as of homogeneous degree at most u, $\mathcal{P}'(x)$ has homogeneous degree of most j.

Proof: The statements (a) and (b) are obvious, and we turn to (c).

Suppose

$$\mathcal{P}(x) \;=\; \sum_{\xi} M_\xi(x)$$

with the M_ξ distinct monomials. If $M_\xi(x)$ has homogeneous degree r, then

$$M_\xi(\lambda \circ x) \;=\; \lambda^r M_\xi(x)\,.$$

Thus, it suffices to observe that there is at least one x with $|x_u^v| \leq 1$ such that

$$\sum_{\substack{\xi \\ \deg M_\xi = \lambda}} M_\xi(x)$$

is not zero.

To prove (d) we may assume $\mathcal{P}(x)$ is a monomial

$$\mathcal{P}(x) \;=\; x_1^{e_1} \cdots x_k^{e_k}\,.$$

$$\mathcal{P}'(\lambda \circ x) \;=\; \mathcal{P}_1^{e_1}(\lambda \circ x) \cdots \mathcal{P}_k^{e_k}(\lambda \circ x)\,.$$

Now each $\mathcal{P}_u^v$ has homogeneous degree at most u. So each $\mathcal{P}_u^v(\lambda \circ x)$ has degree at most u in λ. Thus $\mathcal{P}_u^{e_u}(\lambda \circ x)$ has degree at most $u|e_u|$ as a polynomial in λ. So the degree of $\mathcal{P}'(\lambda \circ x)$ as a polynomial in λ is at most $|e_1| + 2|e_2| + \cdots + |k|e_k|$ which is at most j. Conclusion (d) now follows from conclusion (c).

We shall also use the same notion of homogeneous degree for polynomials in two variables x and y.

Thus if $M(x,y)$ is a monomial with $M(x,y) \;=\; x_1^{e_1} \cdots x_k^{e_k} y_1^{e'_1} \cdots y_k^{e'_k}$, we say the homogeneous degree of $M(x,y)$ is $|e_1| + |e'_1| + \cdots k|e_k| + k|e'_k|$.

We now relate the coordinates of the first and second kind.

Proposition 3.4: *Let $g \in G$ with*

$$g \;=\; \exp(x_1 * X_1) \cdots \exp(x_k * X_k)\,.$$

Then

$$g \;=\; \exp(z \cdot X_1 + \cdots + z_k \cdot X_k)$$

with

(3.3) $z_u^v = x_u^v + q_u^v(x_1, \ldots x_{u-1})$, $1 \leq u \leq k$.

Here each q_u^v is a polynomial of homogeneous degree at most u. Moreover, q_u has rational coefficients and $q_u(0) = 0$.

Conversely, if

$$g = \exp(z_1 \cdot X_1 + \cdots + z_k \cdot X_k)$$

g has a unique representation

$$g = \exp(x_1 * X_1) \cdots \exp(x_k * X_k)$$

with

(3.4) $x_u = y_u + q'_u(y_1, \ldots y_{u-1})$

where each q'^v_u has homogeneous degree of most u, rational coefficients and $q'_u(0) = 0$. In particular, each $g \in G$ has a unique expression of the form (3.2).

Proof: We first note that

$$[\mathfrak{i}_u, \mathfrak{i}_{u'}] \subseteq \mathfrak{i}_{u+u'}\,.$$

This is because $\mathfrak{i}_u$ consists of all elements of $\mathfrak{g}$ which are finite linear combinations of commutators of length at least u.

By the Campbell-Hausdorff formula

$$\begin{aligned}\exp(x_1 * X_1) \cdot \exp(x_2 * X_2) \cdots \exp(x_k \cdot X_k) \\ = \exp\{x_1 \cdot X_1 + \cdots + x_k \cdot X_k + \sum_\xi M_\xi\}\,.\end{aligned}$$

The sum on ξ is finite, and for each ξ

$$M_\xi = A_\xi\, x_{u_1}^{v_1} \ldots x_{u_r}^{v_r} [X_{u_1}^{v_1}, [X_{u_2}^{v_2}, \ldots [X_{u_r}^{v_{r-1}}] \ldots]$$

where the A_ξ are rational, and $r \geq 2$. Each M_ξ has homogeneous degree $u_1 + \cdots + u_r$ and $[X_{u_1}^{v_1}[X_{u_2}^{v_2} \ldots [X_{u_{r-1}}^{v_{r-1}}, X_{u_r}^{v_r}] \ldots]$ belongs to $\mathfrak{i}_{u_1 \cdots + u_r}$ and so can therefore be written as

$$\sum_{u \geq u_1 + \cdots + u_r} A_{u,v}\, \mathcal{P}_{u,v}\,(x_{u_1}^{v_1}, \ldots x_{u_r}^{v_r})\, X_u^v$$

with $A_{u,v}$ rational. Since $u_1 + \cdots + u_r \leq u$ and $r \geq 2$ each u_j is strictly less than u. Also, $\mathcal{P}_{u,v}(0) = 0$. This gives (3.3). (3.4) now follows from (3.3) by inductively solving for x_u in terms of z_u and using Lemma 3.3.

We now turn to the description of multiplication in exponential coordinates of the second kind.

Proposition 3.5: *Suppose g and h are in G with*

$$g = \exp(y_1 * X_1) \ldots \exp(y_k * X_k)$$

and

$$h = \exp(x_1 * X_1) \ldots \exp(x_k * X_k)\,.$$

Then

$$g \cdot h = \exp(z_1 * X_1) \exp(z_2 * X_2) \ldots \exp(z_k * X_k)$$

where

$$z_1 = y_1 + x_1$$

and for $u \geq 2$

$$z_u = y_u + x_u + P_u(y_1, \ldots y_{u-1}, x_1 \ldots x_{u-1})\,.$$

$P_u = (P_u^1, \ldots P_u^{d(u)})$ are polynomials having the following properties:

(a) *$P_u(0,x) = P_u(y,0) = 0$; P_u has rational coefficients; and each P_u^v has homogeneous degree (in x and y) at most u.*

(b) *Let*

$$P_k = (P_k^1, \ldots P_k^{d(k)})$$

then for $1 \leq v \leq d(k)$

$$P_k^v(y,x) = \sum_{\substack{e,f \\ 1 \leq e \leq d(k-1) \\ 1 \leq f \leq d(k)}} A_{e,f}^v \, y_{k-1}^e \, x_1^f$$

+ a polynomial not involving y_{k-1} or x_{k-1}.

(c) *The $d(k-1) \cdot d(1)$ by $d(k)$ matrix $(A_{e,f}^v)$ with rows parameterized by e and f and columns parameterized by v, has a left inverse. That is, the matrix has rank $d(k)$.*

Note: While conclusions (a) and (b) are more-or-less standard facts, the rank property expressed in (c) is of crucial use in Section 9 below.

Proof: Note first that

$$g \cdot h \,=\, \exp(z_1 X_1 \,+\, \cdots + \, z_k X_k)$$

with

$$z_u \,=\, y_u \,+\, x_u \,+\, P'_u(y_1, \ldots y_{u-1}\,, x_1\,, \ldots x_{u-1})$$

with P'_u satisfying conditions (a). This is proved by using the Campbell-Hausdorff formula in the same way that (3.3) was proved. The conclusions (a), (b) and (c) now follow from (3.4) and part (d) of Lemma 3.2.

We turn to the proof of (b). We assume $k \geq 3$. The proof in the case $k = 2$ is easier but requires a slight change in the argument.

Let

$$w \,=\, \exp(y_{k-1} * X_{k-1}) \cdot \exp(x_1 * X_1) \cdot \exp(-y_{k-1} * X_{k-1})\,.$$

$$w \,=\, \exp(x_1 * X_1) \cdot \exp \sum_{e,f} y^e_{k-1} x^f_1 [X^e_{k-1}, X^f_1] \,=\, \exp(x, * X_1) \cdot w'$$

with

$$w' \,=\, \exp \sum_{e,f} y^e_{k-1} x^f_1 [X^e_{k-1}, X^f_1]\,.$$

In particular, w' is in the center of $\mathfrak{g}$. Thus

$$g \cdot h \,=\, \exp(y_1 * X_1) \cdots \exp(y_{k-2} * X_{k-2}) \cdot w \cdot [\exp(-y_{k-1} * X_{k-1})]^{-1} \exp(x_2 * X_2) \ldots$$

$$\exp(x_{k-1} * X_{k-1}) \exp(y_k * X_k) \exp(x_k * X_k)$$

$$= \exp(y_1 * X_1) \ldots \exp(y_{k-2} * X_{k-1})$$

$$\cdot \exp(x_1 * X_1)(\exp(-y_{k-1} * X_{k-1}))^{-1} \exp(x_2 * X_2) \ldots \exp(x_{k-1} * X_{k-1})$$

$$\cdot \exp(\sum_v X^v_k (y^v_k + x^v_k \,+\, \sum_{e,f} A^v_{e,f}\, y^e_{k-1} x^f_1)$$

where $A^v_{e,f}$ is defined by the relation

$$[X^e_{k-1}, X^f_1] \,=\, \sum_v A^v_{e,f} X^v_k\,.$$

Also, since $k \geq 3$,

$$\exp(-y_{k-1} * X_{k-1})^{-1} = \exp(y_{k-1} * X_{k-1})$$

and that term commutes with all the terms on its right in the above formula.

Thus

$$\begin{aligned} g \cdot h &= \exp(y_1 * X_1) \dots \exp(y_{k-2} * X_{k-2} \dots \\ \dots \exp(y_{k-1} + x_{k-1}) * X_{k-1} &\exp\left\{ \sum_v \left(y_k^v + x_k^v + \sum_{e,f} A_{e,f}^v \, y_{k-1}^e x_1^f \right) x_{k-1}^v \right\} \\ &= \exp((y_1 + x_1) * X_1) \exp(Q_2(y_1 \dots y_{k-2}) \cdot X_2 \\ &\quad + \dots + Q_k(y_1, \dots y_{k-2}, x_1, \dots x_{k-2}) \cdot X_k) \\ \cdot \exp(y_{k-1} + x_{k-1}) X_{k-1} &\exp\left(\sum_v \left(y_k^v + x_k^v + \sum_{e,f} A_{e,f}^v y_{k-1}^e x_1^f \right) X_k^v \right) \end{aligned}$$

Since $[X_{k-1}, X_u] = 0$ for $u \geq 2$ ($k \geq 3$) the conclusion follows.

We turn to the proof of (c).

Recall that $\mathfrak{s}_u + \mathfrak{s}_{u+1} \cdots + \mathfrak{s}_k = \mathfrak{i}_k$, where $\mathfrak{s}_u$ is the subspace of the Lie algebra spanned by X_u^1, $X_u^2 \cdots X_u^{d(u)}$. Then since $\mathfrak{s}_u \subset \mathfrak{i}_u$, if $u + u' > k$ we have that $[\mathfrak{s}_u, \mathfrak{s}_{u'}] = 0$, because $[\mathfrak{i}_u, \mathfrak{i}_{u'}] = 0$. Therefore, $\mathfrak{i}_k = [\mathfrak{g}, \mathfrak{i}_{k-1}] = [\mathfrak{s}_1 + \mathfrak{s}_2 \cdots + \mathfrak{s}_k \, , \mathfrak{s}_{k-1} + \mathfrak{s}_k] = [\mathfrak{s}_1, \mathfrak{s}_{k-1}]$, which gives $\mathfrak{i}_k = [\mathfrak{s}_1, \mathfrak{s}_{k-1}]$.

So $\mathfrak{i}_k$ is contained in the span of the commutators X_1^v and X_{k-1}^v of length 2. That is, we can write each X_k^r as

$$X_k^r = \sum_{e,f} B_{e,f}^r \, [X_{k-1}^e, \, X_1^f] \, .$$

Thus the matrix $B_{e,f}^r$ is a left inverse for $A_{e,f}^v$.

Corollary 3.6: *If $g \in G$ and*

$$\begin{aligned} g &= \exp(x_1 * X_1) \quad \cdots \quad \exp(x_k * X_k) \\ g^{-1} &= \exp(z_1 * X_1) \quad \cdots \quad \exp(z_k * X_k) \end{aligned}$$

where $z_1 = -x_1$

and for $2 \le u \le k$

$$z_u \;=\; -x_u \;+\; P''_u(x_1, \ldots x_{\ell-1})\,.$$

Here P''_u *are polynomials satisfying the following conditions.*

(a) P''_u *has rational coefficients*

(b) P''_u *has homogeneous degree at most* u.

(c) $P''_u(0) \;=\; 0$.

To prove Corollary 3.6 put $y_u + z_u = 0$ in Proposition 3.5 and solve for y_u recursively. Also use part (d) of Lemma 3.2 to check the homogeneous degree of the P''_u.

Corollary 3.7: *If*

$$g \;=\; \exp(y_1 * X_1) \ldots \exp(y_k * X_k)$$

and

$$h \;=\; \exp(x_1 * X_1) \ldots \exp(x_k * X_k)$$

then

$$g^{-1}h \;=\; \exp(z_1 * X_1) \ldots \exp(z_k * X_k)$$

where

$$z_1 \;=\; x_1 - y_1$$

and for $2 \le u \le k$

$$z_u \;=\; x_u - y_u \;+\; P'_u(y_1, \ldots y_{u-1}, x_1, \ldots x_{u-1})$$

Here the P'_u *are polynomials satisfying the following conditions.*

(a) P'_u *has rational coefficients*; $P'_u(y, y) \;=\; 0$; $P'_u(0, x) \;=\; 0$ *and* $P'_u(y, x)$ *has homogeneous degree at most* u.

(b) $P'_k \;=\; (P'^1_k\,, \ldots\, P'^{d(k)}_k)$ *where each* P'^v_ℓ *has the form*

$$P'^v_k(y, x) \;=\; -\sum_{e,f} A^h_{e,f}\, y^e_{k-1} x^f_1$$

$+$ *polynomial in* y

$+$ *polynomial not involving* y_{k-1} *or* x_{k-1} .

Corollary 3.7 is a consequence of Corollary 3.6 together with Proposition 3.5 and conclusion (d) in Lemma 3.2.

We need analogues of Propositions 3.4 and 3.5 and its corollaries for G'. Since π is a homorphism onto and $\pi_*(X_k) = 0$, each element g' in G' can be written in the form

$$(3.5) \quad g' = \exp(z_1 \cdot \pi_*(X_1) + \cdots + z_{k-1} \cdot \pi_*(X_{k-1}))$$

and

$$(3.6) \quad g' = \exp(x_1 * \pi_*(X_1)) \circ \cdots \circ \exp(x_{k-1} * \pi_*(X_{k-1}))$$

Proposition 3.8: *If g' has expressions (3.5) and (3.6), then for $1 \leq u \leq k-1$,*

$$(3.7) \quad y_u = x_u + q_u(x_1, \ldots x_{u-1})$$

and

$$x_u = y_u + q'_u(x_1, \ldots x_{u-1})$$

where the q_u and q'_u are the same polynomials as in Proposition 3.4.

Proof: Proposition 3.8 follows from Proposition 3.4 because π is a homomorphism onto.

The fact that π is a homomorphism onto also gives us the following analogues of Proposition 3.5, Corollary 3.6 and Corollary 3.7.

Proposition 3.9: *Suppose g' and h' are in G' with*

$$g' = \exp(y_1 * \pi_*(X_1)) \circ \exp(y_2 * \pi_*(X_2)) \circ \cdots \circ \exp(x_{k-1} * \pi_*(X_{k-1}))\,.$$

Then

$$g' \circ h' = \exp(z_1 * \pi_*(X)_1)) \circ \exp(z_2 * \pi_*(X_2)) \circ \cdots \circ \exp(z_{k-1} * \pi_*(X_{k-1}))\,,$$

$$g'^{-1} = \exp(z''_1 * \pi_*(X_1)) \circ \cdots \exp(z''_{k-1} * \pi_*(X_{k-1}))$$

and

$$(g'^{-1} \circ h) = \exp(z'_1 * \pi_*(X_1)) \circ \cdots \circ \exp(z'_{k-1} * \pi_*(X_{k-1}))$$

where for $1 \le u \le k-1$

$$z_u \;=\; y_u \,+\, x_u \,+\, P_u(x_1 \dots x_{u-1}, y_1, \dots y_{u-1})$$

$$z''_u \;=\; -y_u \,+\, P''_u(y_1, \dots y_{u-1})$$

and

$$z'_u \,=\, x_u - y_u \,+\, P'_u(x_1, \dots x_{u-1}, y_1, \dots y_{u-1})$$

where the polynomials P_u, P''_u *and* P'_u *are the same as in Proposition 3.5, Corollary 3.6 and Corollary 3.7.*

Next we want to define the balls $\mathcal{B}^0_r$ and $\mathcal{B}^{0\prime}_r$ on G and G' in terms of coordinates of the second kind; these will be "equivalent" to the balls $\mathcal{B}_r$ and $\mathcal{B}'_r$ introduced earlier in Section 1. Let

$$\mathcal{B}^0_r \,=\, \{g \in G :\, g \,=\, \exp(x_1 * X_1) \cdots \exp(x_k * X_k)$$

with

$$\sup_{1 \le v \le d(u)} |x^v_u| \,<\, r^u$$

for

$$1 \,\le\, u \,\le\, k\}\,.$$

We set $\mathcal{B}^{0\prime}_r \,=\, \pi(\mathcal{B}^0_r)$. So

$$\mathcal{B}^{0\prime}_r \,=\, \{g' \in G' :\, g' \,=\, \exp(x_1 * \pi_*(X_1)) \circ \cdots \circ \exp(x_{k-1} * \pi_*(X_{k-1}))$$

with

$$\sup_{1 \le v \le d(u)} |x^v_u| \,<\, r^u$$

for

$$1 \,\le\, u \,\le\, k-1\}$$

The balls $\mathcal{B}_r$ and $\mathcal{B}^0_r$ are equivalent in the following sense.

Proposition 3.10:

(a) *There exists a positive constant* c *such that whenever* $r \le 1$,

$$\mathcal{B}_r \,\subset\, \mathcal{B}^0_{cr}$$

and

$$\mathcal{B}_r^0 \subset \mathcal{B}_{cr}\,.$$

(b) *There exists a positive constant c such that*

$$\mathcal{B}0'_r \subset \mathcal{B}_{cr}^{0\prime}$$

and

$$\mathcal{B}_r^{0\prime} \subset \mathcal{B}'_{cr}$$

(c) *The ratios*

$$\frac{|\mathcal{B}_r^0 \cap \Gamma|}{r^D} \text{ and } \frac{|\mathcal{B}_r^{0\prime} \cap \Gamma'|}{r^{D'}}$$

are bounded above and below.

(d) *The ratios*

$$\frac{|\mathcal{B}_r \cap \Gamma|}{r^D} \text{ and } \frac{|\mathcal{B}'_r \cap \Gamma'|}{r^{D'}}$$

are bounded above and below.

Proof: Suppose first $g \in \mathcal{B}_r$ so that

$$g = \exp(t_1 Y + \cdots + t_k Y_k)$$

with $Y_u \in \mathcal{I}_u \cap \mathcal{O}$. ($\mathcal{O}$ a fixed neighborhood of 0 in $\mathcal{G}$), and $|t_u| < r^u$. Then by writing Y_u in terms of the X_u^v, we see

$$g = \exp(z_1 \cdot X_1 + \cdots + z_k \cdot X_k)$$

with

$$|z_u| < Cr^u\,.$$

We then use Proposition 3.4 to write

$$g = \exp(x_1 * X_1) \cdots \exp(x_k * X_k)\,.$$

The homogeneity conditions on the polynomials q'_u in Proposition 3.4, then imply $|x_u^v| < Cr^u$. Thus

$$\mathcal{B}_r \subset \mathcal{B}_{cr}^0\,.$$

Similarly the homogeneity conditions on the polynomials q_u in Proposition 3.4 imply $\mathcal{D}_r \subset \mathcal{B}_{cr}$. This proves (a). The proof of (b) is the same. The conclusion (c) is clear, and then (d) follows from (c) and the inclusion relations in (a) and (b).

Notes:

- Since we have not assumed that the dilations $\{x_u^v\} \longrightarrow \{\delta^u X_u^v\}$, $\delta > 0$, are automorphisms, the properties of the balls $\mathcal{B}_r$ in the above proposition, as well in Proposition (3.12) below, can only be asserted for n strictly bounded from below.
- The definition of the balls $\mathcal{B}_r^0$ makes them more convenient for calculation than their equivalent balls $\mathcal{B}_r$. However not to further encumber the notation, we shall from now on designate the former balls as $\mathcal{B}_r$, (and designate the projected balls $\mathcal{B}_r^{0\prime}$ as $\mathcal{B}_r'$).

We now consider the appropriate norm and distance functions.

If g is in G and

$$g \;=\; \exp(x_1 * X_1) \cdots \exp(x_k * X_k)$$

we define

$$\| \, g \, \| = \sup_{\substack{1 \le v \le d(u) \\ 1 \le u \le k}} |x_u^v|^{1/u} \, .$$

If g and h are in G, we set

$$\rho(g,h) \;= \| \, h^{-1} g \, \| \;\; .$$

We gather together the basic properties of $\| \;\; \|$ and ρ in the next proposition.

Proposition 3.11 *Suppose g and h are in G.*

(a) $\| \, g \, \| = 0$ *if and only if g is the identity in G.*

(b) $\rho(g,h) \;=\; 0$ *if and only if $g = h$.*

For the remaining conclusion we assume $\| \, g \, \|$ and $\| \, h \, \|$ are ≥ 1.

(c) $\| \, g^{-1} \, \| \leq \; C \;\| \, g \, \|$

(d) $\rho(g,h) \leq C\rho(h,g)$

(e) $\| g \cdot h \| \leq C(\| g \| + \| h \|)$

(f) $\rho(g,h) \leq C(\rho(g,w) + \rho(w,h))$ *for and* $w \in G$.

(g) $\rho(g \cdot h, h) \leq C(\| g \| + \| h \|^{1-1/k} \| g \|^{1/k})$.

Proof: Statements (a) and (b) are clear. Conclusion (c) follows from Corollary 3.6. Furthermore, conclusion (c) implies conclusion (d). Property (e) is a consequence of Proposition 3.5 - that is the homogeneity properties of the polynomials p_u. If we write $h^{-1}g = h^{-1}w \cdot w^{-1}g$, we see that (c) and (e) imply (f). We turn to the proof of (g). If $\| g \| \geq \| h \|$, we see

$$\begin{aligned} \rho(g \cdot h, h) &= \| h^{-1}gh \| \leq C(\| g \| + \| h \|) \\ &\leq C(\| g \| + \| h \|^{1-1/k} \| h \|^{1/k}) \\ &\leq C(\| g \| + \| h \|^{1-1/k} \| g \|^{1/k}) . \end{aligned}$$

So we may assume $\| g \| \leq \| h \|$. Let

$$g = \exp(y_1 * X_1) \cdots \exp(y_k * X_k)$$

and

$$h = \exp(x_1 * X_1) \cdots \exp(x_k * X_k) .$$

Then

$$h^{-1}g\,h = \exp(z_1 * X_1) \cdots \exp(z_k * X_k)$$

where

$$z_u = y_u + Q_u(y_1, \ldots y_{u-1}, x_1, \ldots x_{u-1})$$

where each Q_u is a polynomial of homogeneous degree at most u, and $Q_u(0,x) = 0$ So

$$|Q_u(y,x)| \leq \sum_{s=1}^{u} |y_s| \, | \, Q'_u(y_1, \ldots y_{u-1} x_1, \ldots x_{u-1})|$$

where each Q'_u is for homogeneous degree at most $u - s$. (Here $|y_s|$ denotes the Euclidean norm of the vector y_s.)

So if $\| g \| = A$ and $\| h \| = B$

$$|z_u| \leq C(A^u + A^s B^{u-s}) .$$

Thus

$$\begin{aligned} |z_u|^{1/u} &\leq C(A + A^{s/u} B^{1-\frac{s}{u}}) \\ &\leq C(A + A^{1/k} B^{1-1/k}) \end{aligned}$$

since $A \leq B$.

This completes the proof of Proposition 3.11.

Notice that $\mathcal{B}_r = \{g \in G , \| g \| < r\}$.

More generally, for $g \in G$, we now set

$$\mathcal{B}_r(g) = \{h \in G : \rho(h, g) < r\} .$$

Proposition 3.2 implies the following Vitali properties of the balls $\mathcal{B}_r(g)$.

Proposition 3.12: *There is a constant $c > 0$ so that if $r \geq 1$*

(a) *If $\mathcal{B}_r(g) \cap \mathcal{B}_r(h) \neq 0$*

$$\mathcal{B}_r(g) \subset \mathcal{B}_{cr}(h) .$$

(b) *If $h \in \mathcal{B}_r(g)$, then*

$$g \in \mathcal{B}_{cr}(h) .$$

Let us define averages $\mathcal{A}'_r f$ for functions f on Γ as follows.

For $\alpha \in \Gamma$

$$(3.8) \quad \mathcal{A}'_r f(\alpha) = \frac{1}{r^{\mathcal{B}'}} \sum_{\beta' \in \mathcal{B}'_r \cap \Gamma'} f(\alpha \cdot \mathcal{R}(\beta')) .$$

Then in view of Proposition 3.10, we have the following assertion:

To prove our main theorem, it suffices to prove

$$\| \sup_{r>0} |\mathcal{A}'_r f| \|_{\ell^2} \leq C \| f \|_{\ell^2} .$$

§4. Coordinates

The use of the cannonical coordinates of the second kind allows us to identify the group G with $\mathbb{R}^{d'} \times \mathbb{R}^{d(k)}$ or $G' \times \mathbb{R}^{d(k)}$ as follows.

Every $g \in G$ can be written as a pair (g', x_k)

with $g' = \exp(x_1 * \pi(X_1)) \cdot \exp(x_2 * \pi(X_2)) \cdots \exp(x_{k-1} * \pi(X_{k-1}))$

and $g = \exp(x_1 * X_1) \exp(x_2 * X_2) \cdots \exp(x_{k-1} * X_{k-1}) \cdot \exp(x_k * X_k)$.

It will then be convenient to write

$g = (x, x_k)$, where $x = (x_1, \ldots x_{k-1}) \in \mathbb{R}^{d'}$, with $d' = d(1) + d(2) \cdots d(k-1)$, $x_k \in \mathbb{R}^{d(k)}$. In this way we identify G with $\mathbb{R}^{d'} \times \mathbb{R}^{d(k)}$ and $G' \times \mathbb{R}^{d(k)}$, and identify G' with $\mathbb{R}^{d'}$. It then follows that $g = (x, x_k) = (x, 0) \cdot (0, x_k)$ and $\pi(x, x_k) = x$.

Similarly, any $\alpha \in \Gamma$ can be represented as $\alpha = (a, a_k)$, with $a \in \mathbb{Z}^{d'}$ and $a_k \in \mathbb{Z}^{d(k)}$, and Γ' is then identified with $\mathbb{Z}^{d'}$.

Suppose

$$\begin{aligned} g &= \exp(y_1 * X_1) \ldots \exp(y_{k-1} * X_{k-1}) \cdot \exp(y_k * X_k) \\ &= (y, y_k) \end{aligned}$$

and

$$\begin{aligned} h &= \exp(x_1 * X_1) \ldots \exp(x_{k-1} * X_{k-1}) \exp(x_k * X_k) \\ &= (x, x_k) . \end{aligned}$$

Then we have the following relation between multiplication in G' and G.

Proposition 4.1:

(a) $g \cdot h = (z, z_k)$

where

$$z_k \;=\; y_k \;+\; x_k \;+\; P_k\,(y_1, \ldots y_{k-1}\, x_1, \ldots x_{k-1})\,.$$

(b) $g^{-1} \;= (y^{-1}, z''^{k})$

where y^{-1} is the inverse of y with respect to the group structure of G' and

$$z''_k \;=\; -y_k \;+\; P''_k(y_1, \ldots y_{k-1})\,.$$

(c)

$$g^{-1} \;\cdot\; h \;=\; (y^{-1} \circ x, z'_k)$$

and

$$z'_k \;=\; x_k - y_k \;+\; P'_k(y_1, \ldots y_{k-1},\; x_1,\; \ldots\; x_1, \ldots x_{k-1})\,.$$

Here P_k, P'_k and P''_k are as in Proposition 3.5. The proof of Proposition 4.1 is an immdiate consequence of Proposition 3.5, Proposition 3.9 and the fact that $\exp(x_k * X_k)$ is in the center of G.

We now consider the polynomial mapping $\mathcal{R} : G' \longrightarrow G$. Recall that by assumption $\pi\mathcal{R}$ is the identity on G'. Using the coordinates above it therefore follows that

$$\mathcal{R}(x) \;=\; (x, R(x))\,.$$

Since multiplication is given by polynomials, R is therefore a polynomial map (from $\mathbb{R}^{d'}$ to $\mathbb{R}^{d(k)}$), and so we write $R = (R^1, R^2, \cdots R^{d(k)})$. We then have:

Proposition 4.2 *The polynomials $R^1, R^2, \cdots R^{d(k)}$ each have homogeneous degree $\leq k$. Each polynomial has rational coefficients, and is of the form*

(4.1)

$$\begin{aligned} R^v(x) \;&=\; \sum_{\substack{e,f \\ 1 \leq e \leq d(k-1) \\ 1 \leq f \leq d(1)}} B^v_{e,f}\; x^e_{k-1}\; x^f_1 \\ &+\; \sum_e B'^v_e\; x^e_{k-1} \\ &+\; \Lambda(x_1, \ldots x_{k-2}) \end{aligned}$$

where Λ is a polynomial of homogeneous degree at most k.

Proof: Note that the R^v are polynomials in the variables $x_1, x_2, \ldots x_{k-1}$only. If R^v had homogeneous degree greater than k there would be an $x \neq 0$ with $|x_u^{v'}| \leq 1$ such that $R^v(\lambda \circ x)$ would be a polynomial in λ of degree at least $k+1$, according to Lemma 3.2. But then

$$\mathcal{R}(\mathcal{B}_1') \not\subset \mathcal{B}_c$$

for any $c > 0$ contradicting the assumption on $\mathcal{R}$. Given that R_k^v is of homogeneous degree at most k, it must have the form (4.1).

Finally, we deal with the rationality assertions in Proposition (4.2).

The fact that $\mathcal{R}$ takes Γ' into Γ implies that for $m = (m_1 \ldots m_{d(k-1)})$ with each $m_u \in \mathbb{Z}^{d(u)}$, $R(m)$ is an element of $\mathbb{Z}^{d(k)}$. Thus, the rationality assertions concerning the coefficients of R in Proposition 4.2 follow from the following lemma.

Lemma 4.3: *Let R be a polynomial of degree at most k in d variables.*

Assume that for every lattice point m in $\mathbb{R}^d$, $R(m)$ is an integer. Then the coefficients of R are rational, and each coefficient may be written with denominator that divides some fixed integer that depends on d and k.

Proof: Let us first suppose $d = 1$, and suppose

$$R = a_k x^k + \cdots .$$

Let $(\Delta R)(x) = R(x) - R(x-1)$. Then $\Delta^k R = k!\, a_k$. So $k! a_k$ is an integer.

Consider

$$\begin{aligned} k!R(x) &= k! a_k x^k + k!\, a_{k-1}\, x^{k-1} + \ldots \\ &= k! a_k x^k + R_{k-1} \end{aligned}$$

Since $R(x)$ takes integers into integers and $k! a_k$ is an integer, R_{k-1} takes integers into integers. Arguing inductively we see $k!(k-1)! \ldots 2! a_j$ is an integer for each $j, 1 \leq j \leq k$. Also a_0 must be an integer since $R(0)$ is an integer.

Now suppose

$$R(x,y) \,=\, \sum_{\ell=0}^{k} y^\ell \, p_\ell(x)$$

where p_ℓ is a polynomial with degree $p_\ell \le k$. Then the preceding argument shows that for each integer $m, p_\ell(m)$ is a rational number and the denominator divides $C(k)$ for some integer $C(k)$. (In fact, $C(k) = k!(k-1)!\cdots 1$.) Then the polynomials $C(k)p_\ell(x)$ takes Z into Z. Thus the polynomials $C(k)p_\ell(x)$ have coefficients which are rational numbers with denominators that divide $C(k)$. Thus, the coefficients of $R(x,y)$ are rationals with denominators that divide $C(k)^2$. This proves the lemma for $d = 2$, and the lemma follows by an inductive argument, which shows that the demoninators that divide $C(k)^d$.

§5. The standard picture

We now want to reformulate our main theorem in a more concrete fashion.

First we recall the identifications above.

We have

$$\Gamma \,=\, \{\alpha :\, \alpha \,=\, \exp(a_1 * X_1) \,\ldots \exp(a_k * X_k)$$

with

$$a_u \,=\, (a_u^1, \ldots, a_u^{d(u)}) \,\in \mathbb{Z}^{d(u)}\,, a_u^v \in \mathbb{Z}\,, 1 \,\le\, v \,\le\, d(u)\}\,.$$

We set $d \,=\, d(1) + \cdots + d(k)$, $D \,=\, d(1) + 2d(2) + \cdots k \cdot d(k)$,

$$d' \,=\, d(1) \,+\, \cdots \,+ d(k-1)\,, \text{ and } D' \,=\, d(1) \,+\, 2d(2) \,+\, \cdots (k-1)\, d(k-1)\,,$$

and use the notation

$$\alpha \,=\, (a_1, a_2, \ldots,\, a_{k-1}, a_k) \,=\, (a, a_k)\,.$$

Our averages $\mathcal{A}'_r$ in (3.8) are now recast as follows.

$$(5.1)\;\; \mathcal{A}'_r(f)(\alpha) \,=\, \frac{1}{r^{D'}} \sum_{b \in \mathcal{B}'_r \cap \mathbb{Z}^{d'}} f(a \circ b, a_k + P_k(a,b) + R(b))\,,$$

where P_k and R are the polynomials described in Propositions 4.1 and 4.2. Also the polynomials involved in the multiplication have 0, as well as P_k and R have rational coefficients.

Proposition 5.1 *The assertion of the theorem is equivalent with the inequality for (5.1)*:

$$\| \sup_{r\geq 1} \mathcal{A}'_r(f) \|_{\ell^2(\mathbb{Z}^d)} \leq C \| f \|_{\ell^2(\mathbb{Z}^d)}$$

§6. A last reduction

Now that we have obtained the form (5.1) with rational polynomials we shall show that we may assume that coefficients of the polynomials are integers. For this reduction we need first of all to consider

$$Q_0 \Gamma = \{\alpha = (a_1, \dots a_k) \epsilon \Gamma : Q_0 | a_u^v \quad \text{for each } a_u^v .\}$$

where Q_0 is an integer so large that all of the denominators of the coefficients of the polynomials P_u^v divide Q_0.

Proposition 6.1: $Q_0\Gamma$ *is a subgroup of* Γ.

Proof: Let $Q_0\alpha = (Q_0\, a_1, \dots Q_0 a_k)$ and $Q_0\beta = (Q_0 b_1, \dots, Q_0 b_k)$ be points of $Q_0\Gamma$. Then

$$(Q_0\alpha \cdot Q_0\beta)_u = Q_0\alpha_u + Q_0\beta_u + P_u(Q_0\alpha, Q_0\beta)\,.$$

But the monomials in P_u^v have at least degree 1 in α *and* β. So

$$P_u^v(Q_0\alpha, Q_0\beta) = Q_0 \sum_{\substack{|e|\geq 1 \\ |f|\geq |}} C_{e,f}^v\, Q_0^{|e|-1}\, \alpha^e \cdot Q_0^{|f|}\, \beta^f$$

where each $C_{e,v}^v$ is a rational number with denominator which divides Q_0. Thus $P_u^v(Q_0\alpha, Q_0\beta)$ is an integer multiple of Q_0. So $Q_0\Gamma$ is closed under multiplication. The coefficients of $(\alpha Q)^{-1}$ are found inductively by solving the equations $Q_0 a_u + b_u + P_u(Q_0 a_1, \dots Q_0 a_{u-1}, b_1, \dots b_{u-1}) = 0$ so we see inductively that the components of $(Q_0 a_u)^{-1}$ are of the form Q_0 times an integer. So $Q_0\Gamma$ is closed under inverse and thus $Q_0\Gamma$ is a subgroup of Γ.

Proposition 6.2: *With the re-parametrization above, the multiplications in* $Q_0\Gamma$ *is expressed by polynomials with integer coefficients. That is, the polynomials* P_u^v, $P^{v\prime\prime}$ *and* $P_u^{v\prime}$ *have integer coefficients. Moreover 5.1, 5.2, 5.3 and 5.4 still hold.*

Proof: This is essentially done in the proof of Proposition 6.1. Note the polynomials expressing the multiplication in $Q_0\Gamma$ differ from those expressing the multiplication in Γ only in that each coefficient gets multiplied by a power of Q_0.

Proposition 6.3: *Each $\alpha \in \Gamma$ has a unique decomposition*

$$\alpha \;=\; \sigma \,\cdot\, Q_0\mu$$

where $Q_0\mu \,\in\, Q_0\Gamma\,, \sigma \,\in \Gamma$ and $\sigma = (s_1 \,\ldots\, s_k) \in [0, Q-1]^d$.

$$0 \,\leq\, s_u^v \,\leq\, Q_0{-}1,\; 1 \,\leq\, v \,\leq\, d(u),\; 1 \,\leq\, u \,\leq\, k\,,$$

Proof: Suppose α has coordinates $(a_1, \ldots a_k)$. Then we want to find lattice points $(m_1, \ldots m_k)$ and $s_1, \ldots s_k$ such that the $0 \leq s_u^v \leq Q-1$, for $1 \leq v \leq d(u)$, $1 \leq u \leq k$ such that

$$a_1 \;=\; s_1 \,+\, Q_0 m_1$$

and

$$a_u \;=\; s_u \,+\, Q_0 m_u \,+\, P_u(s_1, \ldots s_{u-1}\,,\, Q_0 m_1 \cdots Q_0 m_{u-1})\;.$$

Since $P_u(s, 0) = 0$ and since the denominators at the coefficients of P_u divide Q_0, we can solve the equations inductively for s_u and m_u.

We have a similar result for Γ'. We let $Q_0\Gamma'$ denote the points in Γ' whose coordinates are divisible by Q_0.

Proposition 6.4: *$Q_0\Gamma'$ is a subgroup of Γ'. If the points with coordinates $(Q_0 a_1, \ldots Q_0 a_{k-1})$ are parametrized by $(a_1, \ldots a_{k-1})$, the polynomials expressing the product and inverse in $Q_0\Gamma'$ are the same as the first $k-1$ arising in Proposition 5.2.*

Moreover, each $\beta' \in \Gamma'$ has a unique decomposition

$$\beta' \;=\; Q_0\nu' \,\circ\, \tau'$$

where $Q_0\nu' \in Q_0\Gamma'$ and the coordinates t_u^v of τ' satisfy the inequality $0 \leq t_u^v \leq Q_0 - 1$.

Next for each $\tau' \in \mathbb{Z}^{d'}$ with $\tau' \in [0, Q-1]^{d'}$ we define the polynomial $R^{\tau'}$ on $Q_0\Gamma'$ by setting

$$R^{\tau'}(Q_0\nu') \;=\; R((Q_0\nu)' \circ \tau') \;-\; R(\tau') \;-\; P_k(Q_0\nu', \tau')\,,$$

with $\nu' \in \mathbb{Z}^{d'}$.

Also if $\sigma \in \mathbb{Z}^d$, with $\sigma \in [0, Q-1]^d$ and f a function on $\Gamma(= \mathbb{Z}^d)$, then we define the function $f_{\sigma,\tau'}$ on $Q_0\Gamma$ by

$$f_{\sigma,\tau'}(Q_0\mu) \;=\; f(\sigma \cdot (Q_0\mu)) \,\cdot\, (\tau', R(\tau'))\,.$$

Note that

$$
\begin{aligned}
& f(\alpha \cdot \mathcal{R}(\beta')) \\
&= f((\sigma \cdot Q_0\mu) \cdot (Q_0\nu' \circ \tau,\, R(Q_0\nu' \circ \tau')) \\
&= f((\sigma \cdot Q_0\mu) \cdot (Q_0\nu',\, R(Q_0\nu' \circ \tau') - P_k(Q_0\nu', \tau'))) \cdot (\tau', 0) \\
&= f(\sigma \cdot Q_0\mu) \cdot (Q_0\nu',\, R(Q_0\nu' \circ \tau') - P_k(Q_0\nu', \tau' - R(\tau'))\, (\tau', R(\tau')) \\
&= f(\sigma \cdot [Q_0\mu \cdot (Q_0\nu', R^{\tau'}(Q_0\nu')) \cdot (\tau', R(\tau')) \\
&= f_{\sigma,\tau'}\, (Q_0\mu \cdot (Q_0\nu', R^{\tau'}(Q_0\nu'))\, .
\end{aligned}
$$

Thus $\mathcal{A}'_r$ appearing in (5.1) can be written as a finite sumof such expressions (with f replaced by $f_{\sigma,\tau'}$) and in each the corresponding polynomials for the multiplication low $\circ$ and P_k have integer coefficients

We wish next to show that we may assume the polynomial R has integer coefficients. We first observe there is a number Q_1 such that the denominators of all the coefficients appearing in the polynomials $R^{\sigma'}(x) = R(x \circ \sigma') \cdot R^{-1}(\sigma')$ divide Q_1. (We know the coefficients of R are rational so the same is true for $R^{-1}(\sigma')$.)

Now we form groups $Q_1 \circ \Gamma$ and $Q_1\Gamma'$. The group $Q_1\Gamma'$ is as before but

$$
Q_1 \circ \Gamma \,=\, \{\alpha \in \Gamma :\, \alpha \,=\, (a_1Q_1, a_2Q_1, \ldots a_{k-1}Q_1, a_k)\, ,\, a_u \in Z^{d(\mu)}\}\, .
$$

In other words, $Q_1 \circ \Gamma$ consists of points whose first $k-1$ coordinates are divisible by Q_1. We then have the following analogue of Propositions 6.1 and 6.3.

Proposition 6.5: *$Q_1 \circ \Gamma$ is a subgroup of Γ, and for each $\alpha \in \Gamma$, there exists a unique decomposition $\alpha = \sigma \cdot (Q_1\mu)$ with $\sigma \in [0, Q_1 - 1]^d$.*

Now

$$
R^{\tau'}(Q_1x) \,=\, R(Q_1x \circ \tau') \,-\, R(\tau') \,-\, P_k(Q_1x, \tau')\, .
$$

Then what is important to notice is that $R^{\sigma'}(Q_1x)$ has integer coefficients. Thus, if we re-parametrize $Q_1 \circ \Gamma$ by identifying $(a_1Q_1, a_2Q_2, \ldots a_{k-1}Q_1, a_k)$ by $(a_1, a_2, \ldots a_{k-1})$, the expression

for $R^{\sigma'}$ has integer coefficients. Moreover, $Q_1\Gamma'$ is a subset of $Q_1 \circ \Gamma'$ and all the structure of Section 5 is preserved.

Thus, we have achieved the following result.

Proposition 6.6: *In proving*

$$\| \sup_{r>0} |\mathcal{A}'_r f| \ \|_{\ell^2(\Gamma)} \leq \ C \ \| \ f \ \|_{\ell^2(\Gamma)}$$

we may assume that all the polynomials arising in (5.1) have integer coefficients.

§7. The basic decomposition

Let us fix a smooth compactly supported function ψ on $\mathbb{R}^{d'}$ which is one in a neighborhood of the origin.

Then for f a function on Γ and $\alpha \in \Gamma$, $\alpha = (a, a_k)$, we set

$$\mathcal{M}_j f(\alpha) \ = \ \sum_{b \in \mathbb{Z}'} \psi_j(b)\, f(a \circ b, a_k + P_k(a,b) + R(b))\,,$$

and

$$\psi_j(b) \ = \ 2^{-jD'}\, \psi(2^{-j} \circ b)\,.$$

It then suffices to prove

$$\| \sup_j \mathcal{M}_j f \ \|_{\ell^2(\mathbb{Z}^d)} \leq \ C \ \| \ f \ \|_{\ell^2(\mathbb{Z}^d)} \ .$$

Now

$$f(a \circ b, a_k \ + \ P_k(a,b) \ + \ R(b))$$

$$= \int\limits_{\mathbb{T}^{d(k)}} e^{-2\pi i a_k \cdot \theta}\, e^{-2\pi i \theta \cdot [P_k(a,b) + R(b)]}\, \hat{f}(a \ \circ \ b, \theta)\, d\theta$$

where

$$\hat{f}(a,\theta) \ = \ \sum_{a_k} f(a, a_k)\, e^{2\pi i a_k \cdot \theta}\,.$$

Thus for $\alpha = (a, a_k)$

$$(7.1) \ \ \mathcal{M}_j(f)(\alpha) \ = \ \int e^{-2\pi i a_k \cdot \theta}\, S_j^\theta\, \hat{f}(\cdot, \theta)(a)\, d\theta.$$

Here S_j^θ is an operator acting on functions on $\mathbb{Z}'$ defined as follows: if F is a function on $\mathbb{Z}'$

$$(7.1^*)\quad S_j^\theta F(b) \;=\; \sum_{b'\in\mathbb{Z}'} \psi_j(b')\, e^{-2\pi i[P_k(b,b')+R(b')]\cdot\theta}\; F(b\circ b')\,.$$

We wish to analyze the integral 7.1 by the circle method.

We let χ' be an even C^∞ function on R which is supported in $[-2,2]$ and which is one on the interval $[-1,1]$ and $0\le\chi'\le 1$. For $\theta=(\theta_1,\dots\theta_{d(k)})$, set $\chi(\theta)=\chi'(\theta_1)\dots\chi'(\theta_{d(k)})$. So for $\lambda>4$, $\chi(\lambda\theta)$ is supported in $[-\frac12,\frac12]$. If $\lambda>4$, we let $\chi_\lambda(\theta)$ denote the periodic extension of $\chi(\lambda\theta)$. Then $\chi_\lambda(\theta)=\sum\limits_{v\epsilon Z^{d(k)}}\hat\chi_\lambda(v)\,e^{-2\pi iv\cdot\theta}$ where

$$\hat\chi(\xi) \;=\; \int\limits_{\mathbb{R}^{d(k)}} e^{2\pi i\xi\cdot y}\,\chi(y)\,dy\,,$$

and

$$\hat\chi_\lambda(v) \;=\; \frac{1}{\lambda^{d(k)}}\,\hat\chi\left(\frac{v}{\lambda}\right)$$

$(\lambda>4)$.

If $\ell=(\ell_1,\dots\ell_{d(k)})$ is in $\mathbb{Z}^{d(k)}$, we write $(\ell,q)=1$ when the only positive integer dividing $q,\ell_1,\dots$ and ℓ_k is 1. We also write $\frac{\ell}{q}$ to represent $\left(\frac{\ell_1}{q}\;\dots\frac{\ell_k}{q}\right)$ in $\mathbb{R}^{d(k)}$. Further we set

$$\mathcal{M}_j^{(\ell,q)}\,f(\alpha)=\int e^{-2\pi i a_k\cdot\theta}\;\cdot\;\chi_\lambda(\theta-\frac{\ell}{q})\;\cdot\;S_j^\theta\,\hat f(\cdot,\theta)\,(a)\,d\theta\,,$$

where $\lambda\;=\;2^{j(k-\epsilon)}$ for an appropriate small ϵ positive. The choice of the ϵ is fixed throughout. In fact we will see below that it suffices to require $0<\epsilon\;\le\;1/k(k+3/2)$.

We shall also *systematically write* $\lambda=2^{j(k-\epsilon)}$ in what follows. Here we note that $2^{j\epsilon}<(2\lambda)^{1/2}$, and in fact this holds on the basis of our requirement since what is needed here is $\epsilon<k/3$. (The stronger restriction $\epsilon\le 1/k(k+3/2)$ is needed in Section 11).

We collect together all the $\mathcal{M}_j^{(\ell,q)}$ with $2^\kappa\;\le\;q\;<\;2^{\kappa+1}$ for some integer κ. So we define

$$\mathcal{M}_{\kappa,j}\,f(\alpha)\;=\;\sum_{2^\kappa\le q<2^{\kappa+1}}\;\sum_{\substack{\ell_v=1\\(\ell,q)=1}}^{q}\mathcal{M}_j^{(\ell,q)}\,f(\alpha)\,.$$

Next we fix a γ with $1 < \gamma < 2$. We put

$$M_j^1 \;=\; \sum_{2^{\kappa+1} \le j^\gamma} \mathcal{M}_{\kappa,j}$$

$$M_j^2 \;=\; \sum_{j^\gamma \le 2^{\kappa+1} \le 2^{\epsilon j}} \mathcal{M}_{\kappa,j}$$

and write

(7.2) $\mathcal{M}_j f \;=\; M_j^1 f \;+\; M_j^2 f \;+\; E_j f\,,$

where

$$E_j f(a, a_k) \;=\; \int\limits_{\mathcal{E}_j} e^{-2\pi i a_k \cdot \theta} \;\cdot\; B(\theta)\, S_j^\theta\, \hat{f}(\cdot, \theta)\,(a)\, d\theta\,.$$

Let us make the following clarifying remarks about this basic decomposition.

(i) For each fixed j we have split the possible demoniators into essentially three classes: the small q, for which $q \le j^\gamma$; the intermediate q, for which $j^\gamma < q \le 2^{\epsilon j}$; and the remainder, for which $2^{\epsilon j} < q$.

(ii) The restriction $q \le j^\gamma$ is crucial in having common demononators Q that are $O(2^{\eta j})$, for all $\eta > 0$, as was already seen in Section 2.

(iii) For both the small q and the intermediate q (these for which $q \le 2^{\epsilon j}$), we have that the supports of the cut-off functions $\chi_\lambda(\theta - \ell/q)$ are disjoint. In fact if two such supports intersect one would have $|\frac{\ell_v}{q} - \frac{\ell' v}{q'}| \le 4/\lambda$, which implies $\frac{1}{qq'} \le 4/\lambda$, and this contradicts the assumption that q, q' are both $\le 2^{j\epsilon}$, while $\lambda = 2^{j(k-\epsilon)}$.

(iv) Since $B(\theta) = 1 - \sum\limits_{q \le 2^{\epsilon j}} \chi_\lambda(\theta - \ell/q)$ and the supports are disjoint, it follows that $0 \le B(\theta) \le 1$. Moreover, if $|\theta_v - \ell_v/q| \le 1/\lambda$ for q and ℓ_v $1 \le v \le d(k)$, with $(q, \ell_1, \ell_2, \cdots \ell_{d(k)}) = 1$ and $q \le 2^{\epsilon j}$, then $\chi_\lambda(\theta - \ell/q) = 1$. So $B(\theta) = 0$ and thus $\theta \ne \mathcal{E}_j$, where $\mathcal{E}_j$ is defined as the support of B.

(v) An additional remark about the tri-partite range of the q's described in (i) above and the choice of $\lambda = 2^{j(k-\epsilon)}$. One could have chosen seperate ϵ's, one for the size of the q's

(i.e. $2^{j\epsilon_1}$), and the other for λ, (i.e. $\lambda = 2^{j(k-\epsilon_2)}$). Now the restriction that arises for ϵ_2 (in Section 9), is $0 < \epsilon_2 < 1/k$. While the restriction that is needed for ϵ_1 in Section 11 is then $\epsilon_1(k + 1/2) \leq 1/k - \epsilon_2$. If for simplicity we take $\epsilon_1 = \epsilon_2$, we then get $\epsilon \leq \frac{1}{k(k+3/2)}$.

Returning to the decomposition (7.2), our main theorem will be a consequence of the following 3 estimates:

$$(7.3)\quad \| \sup_{2^{\kappa+1} \leq j^\gamma} |\mathcal{M}_{\kappa,j} f| \|_{\ell^2} \leq C 2^{-\eta\kappa}, \qquad \text{for some } \eta > 0$$

(from which it obviously follows that $\| \sup_j |M_j^1 f| \|_{\ell^2} \leq C \| f \|_{\ell^2}$),

$$(7.4)\quad \| M_j^2 f \|_{\ell^2} \leq C j^{-\gamma/2} \| f \|_{\ell^2},$$

and

$$(7.5)\quad \| E_j f \|_{\ell^2} \leq C 2^{-\eta j} \| f \|_{\ell^2}, \qquad \text{for some } \eta > 0.$$

§8. The splitting of $\mathcal{M}_{\kappa,j}$

In this section we consider M_j^1 and thus assume $2^{\kappa+1} \leq j^\gamma$. We choose a $\delta > 0$ with $\frac{1}{2} < \delta < \frac{1}{\gamma}$, (where the inequalities are strict.) We let $\mathcal{F}$ be a subset of the q's with $2^\kappa \leq q < 2^{\kappa+1}$ containing at most $2^{\kappa\delta}$ of the q's, and let

$$\mathcal{M}_{\kappa,j}^{\mathcal{F}} = \sum_{q \in \mathcal{F}} \sum_{\substack{\ell_v = 1 \\ (\ell,q)=1}}^{q} \mathcal{M}_j^{(\ell,q)}.$$

Since we may divide the q's with $2^\kappa \leq q < 2^{\kappa+1}$, into $2^{(1-\delta)\kappa}$ such groups $\mathcal{F}$, the estimate (7.3) will follow from the estimate

$$(8.1)\quad \| \sup_j \mathcal{M}_{\kappa,j}^{\mathcal{F}} f \|_{\ell^2} \leq C 2^{-\kappa/2} \| f \|_{\ell^2}$$

The advantage of decomposing the q's into these smaller classes $\mathcal{F}$, as we have shown in Section 2, is that for any $\eta > 0$, we have

$$(8.2)\quad Q \;=\; \prod_{q\in\mathcal{F}} q \;\le\; C_\eta 2^{\eta j}.$$

For $2^\kappa \le q < 2^{\kappa+1}$ (since $2^\kappa < j^\gamma$, thus $2^{2\kappa} \ll \lambda = 2^{j(k-\epsilon)}$) we can write for $\alpha = (a, a_k)$

$$\mathcal{M}_j^{(\ell,q)} f(\alpha) \;=\; \int e^{-2\pi i a_k\cdot\theta}\,\chi_\lambda(\theta-\frac{\ell}{q})\,\chi_{c2^{2\kappa}}\left(\theta-\frac{\ell}{q}\right)\,S_j^\theta\,\hat f(\cdot,\theta)\,(a)\,d\theta\,.$$

(The factor $\chi_{c\,2^{2\kappa}}\,(\theta-\frac{\ell}{q})$ has been inserted.) We now define functions $f^{(\ell,q)}$ by the relation

$$\hat f^{(\ell,q)}\,(\theta) \;=\; \hat f(\theta)\,\chi_{c\,2^{2\kappa}}(\theta-\frac{\ell}{q})\,.$$

So we may write

$$(8.4)\quad \mathcal{M}_j^{(\ell,q)} f(\alpha) \;=\; \int e^{-2\pi i a_k\cdot\theta}\,\chi_\lambda\left(\theta-\frac{\ell}{q}\right)\,S_j^\theta\,\hat f^{(\ell,q)}(\cdot,\theta)(a)\,d\theta.$$

The advantage of (8.4) is that the functions $f^{(\ell,q)}$ are orthogonal since their Fourier transforms have disjoint support.

Next we expand $\chi_\lambda(\theta-\frac{\ell}{q})$ in its Fourier series.

$$\chi_\lambda\left(\theta-\frac{\ell}{q}\right) \;=\; \sum_{b_k\,\in\,Z^{d(k)}} \frac{1}{\lambda^{d(k)}}\;\cdot\;\hat\chi\left(\frac{b_k}{\lambda}\right)\,e^{-2\pi\,i\,b_k\,\cdot\,\theta}\,e^{-2\pi\,i\,b_k\,\cdot\,\frac{\ell}{q}}\,.$$

We also expand $\hat f^{(\ell,q)}(\theta)$,

$$\hat f^{(\ell,q)}(\theta) \;=\; \sum_{u\in Z^{d(k)}} f^{(\ell,q)}(u)\,e^{2\pi i u\cdot\theta}\,.$$

Then if we perform the integration in (8.4), we find

$$(8.5)\quad \mathcal{M}_j^{(\ell,q)} f(\alpha) \;=\; \sum_{\substack{\beta\in\Gamma\\ \beta\,=\,(b,b_k)}} e^{2\pi i(b_k-R(b))\,\cdot\,\frac{\ell}{q}}\,\psi_j(b)\,\frac{1}{\lambda^{d(k)}}\,\hat\chi\left(\frac{b_k-R(b)}{\lambda}\right)\,f^{(\ell,q)}\,(\alpha\cdot\beta)\,.$$

(We have identified (b,b_k), $b\in\Gamma'$ and $b_k\in Z^{d(k)}$ with points $\beta\in\Gamma$, and we have used the fact that if $\beta=(b,b_k)$ and $\beta'=(b',b'_k)$, $\beta\cdot\beta' = (b\circ b',\, b_k+b'_k+P_k(b,b'))$.

The next step will be to write

$$\alpha = Q\mu \cdot \sigma \text{ and } \beta = Q\nu \cdot \tau$$

with Q as in (8.2), where σ and τ have their coordinates in $[0, Q-1]$, i.e. belong to $[0, Q-1]^d$. The advantage is that the factor $e^{2\pi 2(b_k - R(b)) \cdot \frac{\ell}{q}}$ will depend only on τ. This will enable us to write $\mathcal{M}^{\mathcal{F}}_{\kappa,j}$ approximately as a composition of operators - one an averaging operator in the $Q\nu$ variables, and the other arithmetic in the τ variables.

We shall find it convenient to use the following notation. For each $\gamma \in \Gamma$ (thought as a point in $\mathbb{Z}^d$) we denote by $\{\gamma\}$ the element of Γ whose elements are congruent to those of γ modulo Q, and where $\{\gamma\}$ lies in $[0, Q-1]^d$. Also for each γ and τ in Γ we set $\bar{\gamma}(\tau) = \{\gamma \cdot \tau\}$.

Proposition 8.1: *Let $Q\Gamma$ denote the points in Γ whose coordinates are divisible by Q. Then*

(a) *$Q\Gamma$ is a normal subgroup of Γ.*

(b) *Each $\alpha \in \Gamma$ can be written uniquely in the form*

$$\alpha = Q\mu \cdot \sigma$$

with $Q\mu$ in $Q\Gamma$ and the coordinates of σ in the interval $[0, Q-1]$.

(c) *If*

$$\alpha = Q\mu \cdot \sigma \text{ and } \beta = Q\nu \cdot \tau$$

$$\alpha \cdot \beta = Q\mu \cdot (Q\nu)^* \{\sigma \cdot \tau\}$$

where for each fixed σ and τ the mapping

$$Q\nu \longrightarrow (Q\nu)^*$$

is a bijection from $Q\Gamma$ onto $Q\Gamma$.

(d) *$\rho(Q\nu, (Q\nu)^*) \leq C(\| Q\nu \|^{1-1/k} Q^{1/k} + Q)$.*

(e) *For each $\gamma \in \Gamma$, the mapping $\bar{\gamma}$ is a bijection of $[0, Q-1]^d$.*

Proof: We already know from Proposition (6.1) that $Q\Gamma$ is a subgroup of Γ. Now we know in addition that the polynomials expressing multiplication and inverse have integer coefficients. Thus, if $\beta \in \Gamma$ and $Q\alpha$ is in $Q\Gamma$, with say $\beta = (b_1, \dots b_k)$ and $Q\alpha = (Qa_1, \dots Qa_k)$

$$(\beta^{-1}(Q\alpha)\beta)_u \;=\; Q\alpha_u \;+\; P_u(Q\alpha, \beta)\,,$$

where P_u is a polynomial with integer coefficients. Also, $P_u(0, \beta) = 0$. Thus, $P_u^v(Q\alpha, \beta) = Q$ times an integer, for $1 \le v \le d(u)$. Hence $Q\Gamma$ is normal. The proof of (b) is the same as the proof of Proposition 6.3.

We turn to the proof of part (c).

$$\begin{aligned} \alpha \cdot \beta \;&=\; Q\mu \cdot \sigma \cdot Q\nu \cdot \tau \\ &=\; Q\mu \cdot \sigma \cdot Q\nu \cdot \sigma^{-1} \cdot \sigma \cdot \tau\,. \end{aligned}$$

Since $Q\Gamma$ is normal in Γ, $\sigma \cdot Q\nu\, \sigma^{-1} = (Q\nu)'$ for some $(Q\nu)' \in Q\Gamma$. Moreover, $(Q\nu)'$ is uniquely determined by $Q\nu$ and σ.

We want to see that the mapping $Q\nu \longrightarrow \sigma \cdot Q\nu \cdot \sigma^{-1}$ is onto $Q\Gamma$. Given $Q\nu'$ in Γ, take $Q\nu = \sigma^{-1} Q\nu' \sigma$. Now by (b), $\sigma\tau = Q\nu'' \cdot \{\sigma \cdot \tau\}$ with $Q\nu''$ uniquely determined by σ and τ. We now put $(Q\nu)^* = (Q\nu)' \cdot (Q\nu)''$. Since multiplication by $Q\nu''$ is a bijection of $Q\Gamma$ onto $Q\Gamma$ and the mapping $Q\nu$ to $Q\nu'$ is a bijection, it follows that the mapping $Q\nu \longrightarrow (Q\nu)^*$ is a bijection from $Q\Gamma$ onto $Q\Gamma$. Conclusion (d) follows from Proposition (3.11), (g).

To prove (e) it suffices to see that the mapping is surjective. Suppose, therefore, that $\alpha = (\alpha_1, \dots, \alpha_k) \in \Gamma$ and also $\alpha \in [0, Q-1]^d$. We want to find a $\tau \in [0, Q-1]^d$ so that $\bar\gamma(\tau) = \alpha$. Now here is a $\tau_1 \in \mathbb{Z}^{d(1)}$, whose coordinates are $[0, Q-1]$, so that the coordinates of $\gamma_1 + \tau_1 - \alpha_1$ are divisible by Q. Next choose $\tau_2 \in \mathbb{Z}^{d(2)}$, with coordinates in $[0, Q-1]$ so that the coordinates of $\gamma_2 + \tau_2 + P_2(\gamma_1, \tau_1) - \alpha_2$ are divisible by Q, and then proceed inductively to determine $\tau = (\tau_1, \tau_2, \cdots \tau_u)$.

We now want to write $\mathcal{M}_{j,\epsilon}^{(\ell,q)}$ in (8.5) in terms of the decomposition of Proposition 8.1. Recall that we are using the notation

$$\alpha \;=\; (a_1, \dots a_{k-1}, a_k) \;=\; (a, a_k)$$

and

$$\beta \;=\; (b_2, \dots b_{k-1}, b_k) \;=\; (b, b_k)$$

for α and β in Γ.

We will write

$$\begin{aligned} Q\mu &= (Qm_1, \ldots, Qm_{k-1}, Qm_k) \,=\, (Qm, Qm_k) \\ Q\nu &= (Qn_1, \ldots, Qn_{k-1}, Qn_k) \,=\, (Qn, Qn_k) \\ \sigma &= (s_1, \ldots s_{k-1}, s_k) \,=\, (s, s_k) \end{aligned}$$

and

$$\tau \;=\; (t_1, \ldots t_{k-1}, t_k) \,=\, (t, t_k)$$

Then employing the decomposition in Proposition (8.1), we see

$$b_k \,=\, Qn_k \,+\, t_k \,+\, P_k(Qn, t) \,\equiv\, t_k \pmod{\mathrm{Q}}$$

since $P_k(0, \nu) \,=\, 0$. Then since R has integer coefficients

$$R(nQ \circ t) \,\equiv\, R(t) \pmod{\mathrm{Q}}\,.$$

Thus the factor

$$e^{2\pi i(b_k - R(b))\,\cdot\,\frac{a}{q}} \,=\, e^{2\pi i(t_k - R(t))\,\cdot\,\frac{a}{q}}\,.$$

Thus for $\alpha = Q\mu \cdot \sigma$ we have

$$\mathcal{M}_j^{(\ell,q)}\, f(\alpha) \,=\, \sum_{\tau\,=\,(t,t_k)} \;\sum_{Q\nu\,=\,(Qn,Qn_k)} e^{2\pi i((t_k - R(t))\,\cdot\,\frac{\ell}{q}}\; \psi_j(Qn \circ t)\, \frac{1}{\lambda^{d(k)}}$$

$$\hat{\chi}\,\left(\tfrac{(Q\nu\cdot\tau)_k \,-\, R(Qn\circ t)}{\lambda}\right) \,\times\, f^{(\ell,q)}(Q\mu \cdot (Q\nu)^* \,\cdot\, \{\sigma \,\cdot\, \tau\})\,.$$

Let

$$\mathcal{M}_j^{\prime(\ell,q)}\, f(Q\mu \,\cdot\, \sigma) \,=\, \sum_{\tau\,=\,(t,t_k)} \;\sum_{Q\nu\,=\,(Qn,Qn_k)} e^{2\pi i(t_k - R(t))\,\cdot\,\frac{\ell}{q}}\; \psi_j((Qn)^*) \,\cdot\, \frac{1}{\lambda^{d(k)}}$$

$$\hat{\chi}\,\left(\tfrac{(Q\nu)_k^* \,-\, R((Qn)^*)}{\lambda}\right)\; f^{(\ell,q)}\,(Q\mu \,\cdot\, (Q\nu)^*\,\{\sigma \,\cdot\, \tau\})\,.$$

We shall try to replace $\mathcal{M}_j^{(\ell,q)}$ by $\mathcal{M}_j^{\prime(\ell,q)}$, but to do this we first transform $\mathcal{M}_j^{\prime(\ell,q)}$. We recall from Proposition (8.1) that the mapping $Qv \longrightarrow (Qv)^*$ is a bijection on $Q\Gamma$. So we may drop the $*$'s in the sum above.

Next we set $h(\tau) = t_k - R(t)$, $\tau = (t, t_k)$. Then $h(\tau) = h(\sigma^{-1} \cdot \sigma\tau)$, and since the multiplication has integer coefficients $\sigma^{-1} \cdot \sigma\tau = \sigma^{-1} \cdot \{\sigma \cdot \tau\}$ modulo Q. Therefore since R also has integer coefficients we get $h(\tau) = h(\sigma^{-1} \cdot \{\sigma \cdot \tau\})$ modulo Q. Thus since q divides Q the term $e^{2\pi i(t_k - R(t))\ell/q}$ can be replaced by $e^{2\pi i[(\sigma^{-1} \cdot \{\sigma\cdot\tau\})_k - R(s^{-1} \circ \{\sigma\cdot\tau\}')} \cdot \ell/q$. Finally by (e) of Proposition (8.1) we may replace $\{\sigma \cdot \tau\}$ by τ and thus find that

$$\mathcal{M}'^{(\ell,q)}_{j,\epsilon} f(Q\mu \cdot \sigma) = \sum_{\tau} \sum_{Q\nu} e^{2\pi i((\sigma^{-1} \cdot \tau)_k - R(\sigma^{-1} \circ \tau))} \cdot \psi_j(Q\nu) \frac{1}{\lambda^{d(k)}}$$

$$\hat{\chi}\left(\tfrac{(Q\nu)_k - R(Qn)}{\lambda}\right) f^{(\ell,q)}(Q\mu \cdot Q\nu \cdot \tau).$$

This would allow us to realize $\mathcal{M}^{\mathcal{F}}_{\kappa,j}$ as a composition of two operators ("tensor product") $\mathcal{N}_j$ acting on the $Q\Gamma$ variables and H acting on the $[0, Q-1]^d$ variables. In fact, we define $\mathcal{N}_j$ for a function F on $Q\Gamma$, by setting

$$(8.5^*) \quad \mathcal{N}_j F(Q\mu) = \sum_{Q\nu} \psi_j(Q\nu) \cdot \frac{1}{\lambda^{d(k)}} \chi\left(\frac{Qn_k - R_k(Qn)}{\lambda}\right) F(Q\mu \cdot Q\nu),$$

and

$$H^\sigma(Q\mu) = \sum_{\substack{\tau,\ell,q \\ q\in\mathcal{F}}} e^{2\pi i((\sigma^{-1}\cdot\tau)_k - R(s^{-1}\circ t))\cdot\frac{\ell}{q}} f^{\ell,q}(Q\mu \cdot \tau).$$

Thus, if we could replace $\mathcal{M}^{(\ell,q)}_j$ by $\mathcal{M}'^{(\ell,q)}_j$, we would have, in effect

$$\mathcal{M}^{\mathcal{F}}_{\kappa,j} f(Q\mu \cdot \sigma) = \mathcal{N}_j H^\sigma(Q\mu).$$

Then to prove (8.1), it would suffice to obtain the following estimates

$$\| \sup_j \mathcal{N}_j F \|_{\ell^2(Q\Gamma)} \leq C Q^{-d} \| F \|_{\ell^2(Q\Gamma)} \tag{8.6}$$

and

$$\sum_{\substack{\sigma \\ Q\mu}} |H^\sigma(Q\mu)|^2 \leq \frac{C Q^{2d}}{2^\kappa} \| f \|^2_{\ell^2(\Gamma)} . \tag{8.7}$$

Therefore let us now consider the error in replacing $\mathcal{M}_j^{(\ell,q)}$ by $\mathcal{M}_j'^{(\ell,q)}$.

We have to replace $\frac{1}{\lambda^{d(k)}}\,\hat\chi\left(\frac{a_k}{\lambda}\right)\,\psi_j(a)$ by

$$\frac{1}{\lambda^{d(k)}}\,\hat\chi\left(\frac{b_k}{\lambda}\right)\,\psi)j(b)\,,$$

where

$$\begin{aligned} a &= Qn\,, \qquad a_k = (Q\nu)_k - R(Qn)\,,\\ b &= (Qn)\circ t\,, \quad b_k = (Q\nu\cdot\tau)_k - R(Qn\circ t)\,. \end{aligned}$$

Now by the support conditions on ψ_j we have $\parallel Qu \parallel \le c2^j$ and $\parallel Qn\circ t \parallel \le c2^j$, and therefore since $|t|\le cQ$, Proposition 3.11 insures that $\rho(Qn\circ t, Qn)\le Q^{1/k}\,2^{j(1-1/2)} + Q$.

Next we invoke the following simple observation. If $\alpha=(a_1,\dots a_k)$, and $\beta=(b_1,\dots b_k)$ have the property that $\parallel\alpha\parallel\le N$, $\parallel\beta\parallel\le N$, and $\rho(\alpha,\beta)\le\delta$ (with $N,\delta\ge$), then

$$(8.8)\quad |a_u-b_u| \le c(\delta^u + \delta N^{u-1}), \qquad \text{for} \quad 1\le u\le k.$$

This can easily be verified by an induction in u.

Applying this to $N\approx 2^j$ and $\delta = Q^{1/k}2^{j(1-1/k)}+Q\le cQ2^{j(1-1/k)}$, and being somewhat wasteful in the powers of Q, we see that

$$|a_u-b_u| \le cQ^u2^{ju}\cdot 2^{-j/k}\,, \quad 1\le u\le k-1\,.$$

Therefore by the mean-value theorem

$$|\psi_j(u)-\psi_j(b)| \le c2^{-jD'}\,Q^k\sum_{u=1}^{k-1}2^{ju}\,2^{-j/k}2^{-ju} \le cQ^k\,2^{-jD'}\,2^{-j/k}\,.$$

Similarly $|a_k-b_k|\le Q^k\,2^{jk}\,2^{-j/k}$. We now invoke the fact that $\epsilon<1/k$ (which is a consequence of our assumption $\epsilon\le\frac{1}{k(k+3/2)}$) to get that

$$|\hat\chi\left(\frac{a_k}{\lambda}\right) - \hat\chi\left(\frac{b_k}{\lambda}\right)| \le \frac{c}{\lambda}\,Q^k\,2^{jk}\,2^{-j/k} \le cQ^k2^{-j(1/k-\epsilon)}\,.$$

Altogether then we see that

(8.9) $\| \Delta_j^{\ell,q)} \| \leq cQ^k\, 2^{-\eta j}$,

for some $\eta > 0$ (in fact $\eta = (1/k) - \epsilon$).

Since $\mathcal{M}_{k,j}^{\mathcal{F}}$ is obtained by adding at most Q terms of the kind $\mathcal{M}_j^{\ell,q}$, we see that we have produced an error which is $O(Q^{k+1}\, 2^{-\eta j})$, which however is much smaller $2^{-\kappa/2}$, since $Q = O(2^{\eta_1 j})$ for any $\eta_1 > 0$, and $2^\kappa \leq j^\gamma$. Thus the error is consistent with the inequality (8.1).

§9. The maximal estimate for $\mathcal{N}_j$

We write $\mathcal{N}_j = \mathcal{N}_j^0 + (\mathcal{N}_j - \mathcal{N}_j^0)$, where $\mathcal{N}_j^0$ is defined as follows: for $Q\mu = (Qm, Qm_k)$ in $Q\Gamma$ and F a function on $Q\Gamma$,

$$\mathcal{N}_j^0\, F(Q\mu) \;=\; {\scriptstyle \frac{1}{2^{jD'}}}\,,\, {\scriptstyle \frac{1}{2^{jkd_k}}} \sum_{Q\nu\,=\,(Qn,Qn_k)} \psi(\frac{1}{2^j} \circ (Qm)^{-1} \circ Qn)$$

$$\hat{\chi}\left(\tfrac{Qn_k - Qm_k + P_k'(Qm,Qn) - R((Qm)^{-1} \circ nQ)}{2^{jk}}\right) \cdot\; F(Qn)\,.$$

Since R has homogeneous degree of most k, $|R((Qm)^{-1} \circ Qn)| \leq C\, 2^{jk}$, whenever $\frac{1}{2^j} \circ (Qm)^{-1} \circ Qn$ is in the support of ψ. Thus,

$$|\mathcal{N}_j^0\, F(Q\mu)| \;\leq\; \frac{1}{2^{jD}} \sum_{Q\nu} \Phi(2^{-j} \circ ((Q\mu)^{-1}\, Q\nu))\, F(Q\nu)$$

with Φ decaying rapidly at infinity. For $Q\mu \in Q\Gamma$, set

$$\mathcal{B}_r^Q(\mu Q) \;=\; \mathcal{B}_r(\mu Q) \cap Q\Gamma\,.$$

Then since the balls $\mathcal{B}_r$ satisfy the conclusions of Propositions 3.10 and 3.12, so do the balls $\mathcal{B}_r^Q$. Also

$$\mathcal{N}_j^0\, F(Q\mu) \;\leq\; \frac{C}{2^{jD}} \sum_{Q\nu\epsilon\, \mathcal{B}_{2^j}^Q(Q\mu)} |F(Q\nu)|\,.$$

Moreover,

$$|\mathcal{B}^Q_{2^j}(Q\mu)| \sim \frac{2^{jD}}{Q^d}$$

since

$$\begin{aligned}\mathcal{B}^Q_{2^j}(Q\mu) = \{&Q\nu : |Qm_1 - Qn_1| < 2^j \\ &|Qm_2 - Qn_2 + P_2'(Qm_1, an_1)| < 2^{2j} \\ &\vdots \\ &\vdots \\ &|Qm_k - Qn_k + P_k'(Qm_1, \dots Qm_{k-1}, Qn_1 \dots Qn_{k-1})| < 2^{jk}\end{aligned}$$

with P_u' homogeneous of degree at most u and $P_u'(Q\mu, Q\mu) = 0$. Thus by the standard argument, we have the estimate

$$\| \sup_j |\mathcal{N}_j^0 F| \|_{\ell^2(Q\Gamma)} \leq C\, Q^{-d} \| F \|_{\ell^2(Q\Gamma)} . \tag{9.1}$$

We are going to analyze $\mathcal{N}_j - \mathcal{N}_j^0$ by using Fourier analysis on the group $Q \cdot \mathbb{Z}^{d(k)}$. For U a function on $Q\,\mathbb{Z}^{d(k)}$, we put, for θ in the $d(k)$ dimensional torus,

$$\widehat{U}^Q(\theta) = \sum_{Qn_k \in Q\,\mathbb{Z}^{d(k)}} U(Qn_k)\, e^{2\pi i Qn_k \cdot \theta}. \tag{9.2}$$

Then

$$U(Qn_k) = Q^{d(k)} \int\limits_{-\frac{1}{2Q} \leq \theta^v \leq \frac{1}{2Q}} e^{-2\pi i Qn_k \cdot \theta}\, \widehat{U}^Q(\theta)\, d\theta \tag{9.3}$$

and

$$\sum |U(Qn_k)|^2 = Q^{d(k)} \int\limits_{-\frac{1}{2Q} \leq \theta^v \leq \frac{1}{2Q}} |\widehat{U}^Q(\theta)|^2\, d\theta. \tag{9.4}$$

For U_1 and U_2 two functions on $\mathbb{Z}^{d(k)}$, we define their convolution as

$$U_1 * U_2\,(Qm) \;=\; \sum_{nQ} U_1(Qm - Qn)\, U_2(Qn)\,.$$

Then

(9.5) $\widehat{U_1 * U_2} \;=\; \widehat{U}_1 \cdot \widehat{U}_2.$

Next, we consider the Fourier transform of $\frac{1}{\lambda}\,\hat{\chi}\,\left(\frac{Qn_k}{\lambda}\right)$, (that is, the Fourier transform on the group $Q\mathbb{Z}^{d(k)}$). Recall that we always take $\lambda = 2^{j(k-\epsilon)}$.

By definition

$$\begin{aligned}\frac{1}{\lambda}\,\hat{\chi}\left(\frac{Qn_k}{\lambda}\right) &= \frac{1}{\lambda}\int \chi(\xi)\, e^{2\pi i Qn_k\,\cdot\,\frac{\xi}{\lambda}}\, d\xi \\ &= \int \chi(\lambda\xi)\, e^{-2\pi i Qn_k\,\cdot\,\xi}\, d\xi \\ = Q^{d(k)} \int\limits_{-\frac{1}{2Q}\,\le\,\theta^v\,\le\,\frac{1}{2Q}} & e^{-2\pi i Qn_k\,\cdot\,\theta}\;\cdot\;\frac{\chi(\lambda\xi)}{Q^{d(k)}}\; d\xi\,.\end{aligned}$$

So then the Fourier transform of $\frac{1}{\lambda}\,\hat{\chi}\,\left(\frac{Qn_k}{\lambda}\right) \;=\; \frac{\chi(\lambda\xi)}{Q^{d(k)}}$.

Thus we see from (9.5) that

$$\mathcal{N}_j\, F(Qm, Qm_k) \;=\; Q^{d(k)} \int\limits_{-\frac{1}{2Q}\,\le\,\theta^v\,\le\,\frac{1}{2Q}} e^{-2\pi i Qm_k\,\cdot\,\theta}\, \frac{1}{Q^{d(k)}}\, \chi(\lambda\theta)\, S_j'^{\theta}\, \hat{F}(\cdot,\theta)\,(Qm)\, d\theta\,.$$

Here $S_j'^{\theta}$ acts on functions V defined on $Q\Gamma'$ by

$$S_j'^{\theta}\, V\,(Qm) \;=\; \frac{1}{2^{jD'}} \sum_{nQ} \psi\,(2^{-j} \,\circ\, (mQ)^{-1} \,\circ\, nQ))\, e^{-2\pi i\theta\,\cdot\,(P_k'(mQ,nQ) - R((mQ)^{-1}\,\circ\, nQ))}\; V(nQ)\,.$$

This is just a variant of (7.1$*$) defined on $Q\Gamma'$ instead of Γ'.

We will prove the following lemma.

Lemma 9.1: *For $2^{-jk} \le |\theta| \le 2^{-j(k-\epsilon)}$,*

$$\|\; S_j'^{\theta}\;\| \le\; \frac{C}{Q^{d'}}\, \frac{1}{(2^{jk}|\theta|)^{1/2}}\,.$$

Let us assume Lemma 9.1 for the moment. Then using Plancherel's Theorem, we see

$$\sum_{Qm_k} |\mathcal{N}_j\, F(Q\mu)\, -\, \mathcal{N}_j^0\, F(Q\mu)|^2$$
$$\leq\, \frac{C}{Q^{d(k)}} \int |\chi(2^{j(k-\epsilon)}(\theta))\, -\, \chi(2^{jk}\theta)|\,^2\, \cdot\, |S_j'^{\theta}\, \hat{F}(\cdot\,,\theta)\,(Q\mu)|^2\, d\theta.$$

So if we use Lemma 9.1, we see

$$\sum_{Q\mu} |\mathcal{N}_j\, F(Q\mu)\, -\, \mathcal{N}_j^0\, F(Q\mu)|^2$$
$$\leq\, \frac{C}{Q^{d(k)}Q^{2d'}} \int\limits_{\frac{1}{2^{j(k-\epsilon)}} > |\theta| > \frac{1}{2^{jk}}} \left(\frac{1}{(\theta 2^{jk})^{1/2}}\right)^2 \cdot \sum_{Qm} |\hat{F}(Qm,\theta)|^2\, d\theta\,.$$

Thus

$$\sum_j \sum_{Q\mu} |\mathcal{N}_j\, F(Q\mu)\, -\, \mathcal{N}_j^0\, F(Q\mu)|^2$$
$$\leq\, \frac{C}{Q^{d(k)}Q^{2d'}} \int\limits_{|\theta| \leq \frac{1}{2Q}} \sum_{mQ} |\hat{F}(mQ,\theta)|^2\, d\theta\, \leq\, \frac{C}{Q^{2d(k)}Q^{2d'}}\, \parallel F \parallel^2_{\ell^2(Q\Gamma)}\,.$$

We have used the fact that $\sum\limits_j \dfrac{1}{|\theta|2^{jk}}$ is uniformly bounded in θ, if the summation in j is restricted to the range $|\theta|2^{jk} \geq 1$, (when $j \geq 1$). Thus since

$$\sup_j\, |\mathcal{N}_j(F)\, - \mathcal{N}_j^0(F)|^2\, \leq\, \sum_j |\mathcal{N}_j(F)\, - \mathcal{N}_j^0(F)|^2\,,$$

we have proved

$$(9.6)\quad \parallel\, \sup_j |\,(\mathcal{N}_j\, -\, \mathcal{N}_j^0)\, F\, \|_{\ell^2(Q\Gamma)} \leq\, C\, Q^{-d}\, \parallel F \parallel_{\ell^2(Q\Gamma)}\,.$$

If we take into account that $d = d' + d(k)$. Together with (9.1), (9.6) gives (8.6).

We turn to the proof of Lemma 9.1.

To estimate the norm of $S_j'^{\theta}$, we will need to discuss the oscillatory term

$$e^{-2\pi i\theta\,\cdot\,(P'_k(Qm,Qn)\,-\,R((Qm)^{-1}\,\cdot\,nQ))}\,.$$

Let

$$r(m,n)\;=\;P'_k(m,n)\;-\;R(m^{-1}\,\circ\,n)\,.$$

From Propositions 3.7 and 4.2 we see that

$$\begin{aligned} r(m,n)\,\cdot\,\theta\;=\;&\sum_{e,f}\sum_{v}A^v_{e,f}\,m^e_{k-1}\,n^f_1\,\theta^v\\ -\sum_{e,f}\sum_{v}&B^v_{e,f}\,(n^e_{k-1}\;-\;m^e_{k-1})\,(n^f_1\;-\;m^f_1)\,\theta^v\\ &+\text{ terms involving only } m\\ &+\text{ terms involving only } n\\ &+\text{ terms not involving } m_{k-1}\text{ or }n_{k-1}\,. \end{aligned}$$

(The terms involving only m or only n will not affect the norm of S'^{θ}_j.)

Here, the $A^v_{e,f}$ and $B^v_{e,f}$ are integers and the matrix $(A^v_{e,f})$ has a left inverse, as was discussed in Section 5.

Thus we can write

$$\begin{aligned} r(m,n)\,\cdot\,\theta\;=\;&\sum_{e,f}\phi^1_{e,f}\,m^e_{k-1}\,n^f_1\;+\;\sum_{e,f}\phi^2_{e,f}\,n^e_{k-1}\,m^f_1\\ &+\;\text{ terms depending only on } m\\ &+\;\text{ terms depending only on } n\\ &+\;\text{ terms not involving } m_{k-1}\text{ or }n_{k-1} \end{aligned}\tag{9.7}$$

where

$$\phi^1_{e,f}\;=\;\sum_{v}(-A^v_{e,f}\;+\;B^v_{e,f})\,\theta^v$$

and

$$\phi^2_{e,f}\;=\;\sum_{v}B^v_{e,f}\,\theta^v\,.$$

We will need the following lemma.

Lemma 9.2: *We have*

$$\theta^v \;=\; \sum_{e,f} D^1_{e,f}\,\phi^1_{e,f} \;+\; \sum_{e,f} D^2_{e,f}\,\phi^2_{e,f}$$

where the $D^i_{e,f} = D^i_{e,f}(v)$, $i = 1, 2$ *are rational numbers.*

Let us assume Lemma 9.2 and complete the proof of Lemma 9.1. Recall that $2^{-jk} \le |\theta| \le 2^{-j(k-\epsilon)}$. Notice that at least one of the $\phi^i_{e,f}$ satisfies $|\phi^i_{e,f}| \ge \delta|\theta|$ for some uniform $\delta > 0$. Let us suppose $|\phi^1_{1,1}| > \delta|\theta|$.

Let $K(Qn, Qn')$ denote the kernel of $(S'^{\theta}_j)^*\, S^{\theta}_j$. Then

$$|K(Qn,Qn')| \le \sum_{m_1} \cdots \sum_{m_{k-2}} \sum_{\substack{m^e_{k-1}\\ e\ge 2}} 2^{-2jD'} \Big| \sum_{m^1_{k-1}} \psi\left(2^{-j}((Qm)^{-1}\circ Qn)\right)\psi(2^{-j}((Qm)^{-1}\circ Qn'))$$

$$e^{2\pi i m^1_{k-1}\{[(n^1_1-(n^1_1)')\,Q^2\phi^1_{1,1} + \sum_{e=2}^{d(1)}(n^e_1-(n^e_1)')Q^2\,\phi^1_{1,e}]\}}\Big|.$$

Consideration of the arguments of ψ in the above, shows $K(Qn, Qn')$ is supported in $\rho(An, Qn') < C\,2^j$.

We wish to replace the sum in (9.7) by a sum of integrals, replacing the variable m'_{k-1} by a continuous variable t below. In doing this, since $|n_1 - n'_1| \le c2^j$ the error due to $K(Qn, Qn')$ will be at most

$$C\,2^{-jD'}(2^j\,\max|\phi^i_{e,f}|) \;\le\; C\,2^{-jD'}\,2^j\,2^{-j(k-\epsilon)} \;\le\; C\,2^{-jD'}\,2^{-j/2}\,,\ \text{if } \epsilon < \frac{1}{2}\,,$$

since $k \ge 2$.

Thus the error in

$$\sum_{Qn} |K(Qn,Qn')| \;+\; \sum_{Qn'} |K(Qn,Qn')|$$

is at most $\le C\,2^{-j/4}Q^{-2d'}$. So the error gives a contribution to the norm of S'^{θ}_j which can be subsummed in the right hand side of the inequality in Lemma 9.1.

Returning to (9.7), note that the number of terms to be summed in $m_1, \dots m_{k-2}, m_{k-1}^2, \dots m_{k-1}^{d(k-1)}$ is at most $C\,\frac{2^{jD'}}{Q^{d'}}\,\cdot\,2^{-(k-1)j}Q$.

In the integral on m_{k-1}^1, we put $t\,=\,2^{-j(k-1)}\,Q\,m_{k-1}^1$. The integral becomes

$$\frac{2^{j(k-1)}}{Q}\int \tilde{\psi}(t)\,e^{\{2\pi i 2^{j(k-1)}\,Qt\{(n_1^1-n_1^{1\prime})\,Q^2\,\phi_{1,1}^1\,+\,\sum_{e\,=\,2}^{d(1)}\,(n_1^e\,-\,n_1^{e\prime})\,\phi_{1,e}^1\}}\,dt\,.$$

Here $\tilde{\psi}(t)$ depends on $m_1, m_2\,\dots\,m_{k-1}^2\,\dots\,m_{k-1}^{d(k-1)}$, but its derivatives are all uniformly bounded.

Thus, the integral is at most

$$C\,\frac{2^{j(k-1)}}{Q}\,\left\{\frac{1}{1\,+\,|2^{j(k-1)}\,Q\,[(n_1^1\,-\,n_1^{1\prime})\,\phi_{1,1}^1\,+\,L]|^2}\right\}$$

where L is a linear combination of the $n_1^e - n_1^{e\prime}$ with $e \geq 2$. Summing over $m_1, -m_{k-2}^2\,\dots\,m_{k-1}^{d(k-1)}$, we see

$$|K(Qn, Qn')|\,\leq\,\frac{C}{Q^{d'}2^{jD'}}\,\cdot\,\left\{\frac{1}{[1\,+\,|2^{j(k-1)}\,Q(n_1^1\,-\,n_1^{1\prime})\,\phi_{1,1}^1\,+\,L]|^2}\right\}\,.$$

Then summing over n_1^1, we see

$$\sum_{n_1^1}|K(Qn,\,Qn')|\,\leq\,\frac{C}{Q^{d'}2^{jD'}}\,\cdot\,\frac{1}{2^{j(k-1)}|Q|\,|\phi_{1,1}^1|}$$

Now summing over the rest of the n-variables, we see

$$\begin{aligned}\sum_{n}|K(Qn,\,Qn')| \;&\leq\; \tfrac{C}{Q^{2d'}\,2^{jk}|\phi_{1,1}^1|}\\ &\leq\; \tfrac{C}{Q^{2d'}\,2^{jk}|\theta|}\,.\end{aligned}$$

which completes the proof of Lemma 9.1.

Let us now consider Lemma 9.2. Since the matrix $A_{e,f}^h$ has a left inverse, there is a $d(k)\times d(k)$ sub-matrix that is invertible. Thus, Lemma 9.2 will be a consequence of an observation about $n\times n$ matrices, where $n = d(k)$. Suppose $A = (a_{j,k})$ is an invertible $n\times n$ matrix and $B = (b_{j,k})$ is any $n\times n$ matrix. Set

$$U_j^1\,=\,\sum_k\,(a_{j,k}\,+\,b_{j,k})\,v_k \tag{9.8}$$

and

$$U_j^2 \;=\; \sum_k b_{j,k}\, v_k. \tag{9.9}$$

The equations (9.8) and (9.9) define 2^n linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$, by mapping

$$(v_1, \ldots v_n) \;\longrightarrow\; \left(U_1^{i(1)}, \ldots U_n^{i(n)}\right)$$

with choices $i(j)$ fixed to be 1 or 2. Then to prove Lemma 9.2, it suffices to show (if A is non-singular) at least one of these 2^n transformations is non-singular. Since A is non-singular, it is easy to see that we may assume A is the identity.

Then we are in fact reduced to the following assertion.

Lemma 9.3: *Let $F = (f_{j,k})$ be any $n \times n$ matrix. Let $\bar{F}$ be any of the 2^n matrices*

$$\bar{F} \;=\; F \;+\; \begin{pmatrix} \delta(1) & 0 & \cdots & 0 \\ 0 & \delta(2) & \cdots & 0 \\ 0 & 0 & \cdots & \delta(n) \end{pmatrix}$$

where $\delta(j)$ is either 0 or 1. Then at least one of the 2^n matrices $\bar{F}$ is non-singular.

In the above, the matirx F is B and A is the identity matrix.

For $n = 1$ the assertion is obvious. We prove the general case by induction on n.

Suppose we know the lemma to be true for $n \times n$ matrices, and let $F = (f_{j,k})$ be an $(n+1) \times (n+1)$ matrix. Put

$$E^1 \;=\; \begin{pmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,n+1} \\ f_{2,1} & & & \\ \vdots & & E & \\ f_{n+1,1} & & & \end{pmatrix}$$

and

$$E^2 \;=\; \begin{pmatrix} f_{1,1}\,+\,1, & f_{1,2} & \dots & f_{1,n+1} \\ f_{2,1} & & & \\ \vdots & & E & \\ f_{n+1,1} & & & \end{pmatrix}$$

where E is any of the 2^n matrices of the size $n \times n$ arising from the matrix

$$f_{j,k} \;\; 2 \le j \le n+1\,, \;\; 2 \le k \le n+1\,.$$

By induction hypothesis, one of the matrices E is non-singular. Now if $\det E^1 \;=\; 0$ and $\det E^2 \;=\; 0$, then by expanding on minors of the first row, we would find $\det E \;=\; 0$ for all the appropriate 2^n choices of the matrices E, contradicting the induction hypothesis.

The statement concerning the rationality of the $D^i_{e,f}$ follows because they are the coefficients of matrices inverse to matrices formed from the $A^v_{e,f}$ and $B^v_{e,f}$.

This finishes the proof of Lemma 9.2 and then also of the estimate 8.6.

§10. The estimate for H^σ.

In this section, we will obtain the estimate 8.7. We will continue with the notation $\sigma = (s_1, \dots s_{k-1}\; s_k) \;=\; (s, s_k)$ and $\tau \;=\; (t_1, \dots t_{k-1} t_k) \;=\; (t, t_k)$. Then we recall for $0 \;\le\; s^v_u \;\le\; Q-1$

$$H^\sigma(Q\mu) \;=\; \sum_{q \in \mathcal{F}} \sum_{\substack{\ell \\ 1 \le \ell^v \le Q \\ (\ell, q) \,=\, 1}} \sum_{\substack{\tau \\ 0 \le t^v_u \le Q-1}} e^{2\pi i [r(s,t) \,+\, (t_k - s_k)] \,\cdot\, \frac{\ell}{q}} \; f^{(\ell, q)}\,(Q\mu \,\cdot\, \tau)$$

where as in Section 9,

$$r(s,t) \;=\; P'(s,t) - R(s^{-1} \,\circ\, t)\,.$$

So we can write

$$H^\sigma(Q\mu) \;=\; \sum_{q \in \mathcal{F}} \sum_{\substack{\ell \\ (\ell, q) = 1}} e^{-2\pi i s_k \,\cdot\, \frac{a}{q}} \; V(\ell, q, s)\,,$$

where

$$V(\ell,q,s) \;=\; \sum_{\tau} e^{2\pi i[t_k + r(s,t)]\cdot \frac{\ell}{q}}\; f^{(\ell,q)}\,(Q\mu \,\cdot\, \tau)\,.$$

So

$$\sum_{s_k} |H^\sigma(Q\mu)|^2 \;=\; Q^{d(k)} \sum_{\ell,q} |V(\ell,q,s)|^2$$

since

$$\sum_{\substack{s_k \\ 0\,\le\, s_k^v \,\le\, Q-1}} e^{-2\pi i s_k \,\cdot\, \left(\frac{\ell}{q} - \frac{\ell'}{q'}\right)}$$

$$= \begin{cases} Q^{d_k} & \text{if} \qquad \frac{\ell}{q} \;=\; \frac{\ell'}{q'} \\ 0 & \text{otherwise} \quad . \end{cases}$$

That is

$$\sum_{s_k} |\, H^\sigma(Q\mu)\,|^2 \;=\; Q^{d(k)} \sum_{(\ell,q)} |\sum_{\tau=(t,t_k)} e^{2\pi i(t_k + r(s,t))\,\cdot\, \frac{\ell}{q}}\, f^{(\ell,q)}\,(Q\mu \,\cdot\, \tau)\,|^2 \tag{10.1}$$

$$\le\; Q^{2d(k)} \sum_{\ell,q} \sum_{t_k} |\sum_{t} e^{2\pi i\, r(s,t)\,\cdot\, \frac{\ell}{q}}\; f^{(\ell,q)}\,(Q\mu \,\cdot\, \tau)\,|^2\,.$$

For a function F define on $\mathbb{Z}^{d'}/q$, and $s \in \mathbb{Z}^{d'}/q$, let us put

$$T_{\ell,q}\, F(s) \;=\; \sum_{\substack{t\,\in\,\mathbb{Z}^{d'}/q \\ 0\,\le\, t_u^v \,\le\, q-1}} e^{2\pi i\, r(s,t)\,\cdot\, \frac{\ell}{q}}\, F(t)\,.$$

We will prove below the following lemma.

Lemma 10.1:

$$\|\; T_{\ell,q}\, F \;\|^2_{\ell^2(\mathbb{Z}^{d'}/q)} \le\; C\, q^{2d'-1} \;\|\; F \;\|^2_{\ell^2(\mathbb{Z}^{d'}/q)}\;\;.$$

Let us assume Lemma 10.1 and show how it gives (8.7). Note that the factor $e^{2\pi i\, r(s,t)\,\cdot\, \frac{\ell}{q}}$ has period q in the s and t variables. So applying Lemma 10.1, the estimate 10.1 becomes

$$\sum_{\substack{\sigma \\ 0\,\le\, s_u^v \,\le\, Q-1}} |H^\sigma(Q\mu)|^2 \;\le\; Q^{2d(k)} \sum_{\ell,q} \left(\frac{Q}{q}\right)^{2d'} \; q^{2d'-1} \sum_{\substack{\tau \\ 0\,\le\, t_u^v \,\le\, Q-1}} |f^{(\ell,q)}\,(Q\mu \,\cdot\, \tau)|^2\,.$$

Next we sum on $Q\mu$. Since $q \approx 2^\kappa$ and the functions $f^{(\ell,q)}$ are orthogonal, we obtain 8.7.

We turn to the proof of Lemma 10.1. One difficulty is that we only know $(\ell_1, \dots \ell_{d(k)}, q) = 1$ and not that $(\ell_v, q) =$ for any v. In order to remedy that situation we first reduce matters to the case that q is a power of a prime. (In that case, we would know $(\ell_v, q) = 1$ for some v.)

To this end, we would like to prove a multiplicative formula for $T_{\ell,q}$. That is, if $(\ell, q_1 q_2) = 1$ and $(q_1, q_2) = 1$, we would like to say

$$T_{\ell,q_1 q_2} \;=\; T_{\ell_1,q_1} \circ T_{\ell_2,q_2}$$

with $(\ell_1, q_1) = (\ell_2, q_2) = 1$. In order to achieve such a formula, we must consider a slightly generalization of $T_{\ell,q}$. Let

$$\Omega(s,t) \;=\; (\Omega^1(s,t), \dots\; \Omega^{d(k)}\,(s,t))$$

where each $\Omega^k(s,t)$ is a polynomial. Moreover, assume for $1 \le v \le d(k)$

$$\begin{aligned}
\Omega^v(s,t) \;=\; & - \sum_{e,f} A^v_{e,f}\, s^e_{k-1}\, t^f_1 \\
& + \sum_{e,f} B^v_{e,f}\, s^e_{k-1}\, t^f_1 \\
& + \sum_{e,f} B^v_{e,f}\, t^e_{k-1}\, s^f_1 \\
& + p_1(s) + p_2(t) \\
& + p_3(s_1, \dots s_{k-2}, t_1, \dots\; t_{k-2})\,.
\end{aligned}$$

with $A^v_{e,f}$ and $B^v_{e,f}$ as in the expressions for $P^{v'}_k(s,t)$ and $R(s^{-1} \circ t)$. Here we allow more general p_1, p_2, p_3, as opposed to the particular ones that arise after (9.6).

Set in the more general case

$$T_{\ell,q}\, F(s) \;=\; \sum_t e^{2\pi i\, \Omega(s,t)\,\cdot\, \frac{\ell}{q}}\, F(t)\,.$$

We will prove Lemma 10.1 by showing

$$\| \; T_{\ell,q}\, F \; \|^2 \le\; C q^{2d'-1} \;\; \| \; F \; \|^2$$

where C does not depend on the particular choice of the polynomials p_1, p_2 and p_3.

We need the following lemma.

Lemma 10.2: *Let $q = q_1 q_2$ with $(q_1, q_2) = 1$. Also, suppose $(\ell, q) = 1$. Then*

$$\| T_{\ell,q} \|_{\ell^2\left(\mathbb{Z}^{d'}/q\right)} \leq \| T'_{\ell_1,q_1} \|_{\ell^2\left(\mathbb{Z}^{d'}/q_1\right)} \| T'_{\ell_2,q_2} \|_{\ell^2\left(\mathbb{Z}^{d'}/q_2\right)}$$

with $\ell_1 = \ell q_2$ and $\ell_2 = \ell q_1$.

The polynomials p_1, p_2 and p_3 arising in T'_{ℓ_1,q_1} and T'_{ℓ_2,q_2} will in general differ from those in $T_{\ell,q}$.

Proof of Lemma 10.2:

Consider the mapping from

$$\mathbb{Z}^{d'}/q_1 \times \mathbb{Z}^{d'}/q_2 \longrightarrow \mathbb{Z}^{d'}/q_1 q_2$$

that sends $(s'^{\nu}_{\mu},\, s''^{\nu}_{\mu})$ into s^{ν}_{μ} with

$$s^{\nu}_{\mu} = q_2\, s'^{\nu}_{\mu} + q_1 s''^{\nu}_{\mu} \,(\mathrm{mod}\, q_1 q_2),,$$

$$1 \leq \mu \leq k-1\,,\, 1 \leq \nu \leq d(\mu)\,,\;\; 0 \leq s'^{\nu}_{\nu} \leq q_2 - 1\,, 0 \leq s''^{\nu}_{\mu} \leq q_1 - 1\,.$$

This map also sends

$$(t'^{\nu}_{\mu},\, t''^{\nu}_{\mu}) \text{ into } t^{\nu}_{\mu}$$

with

$$t^{\nu}_{\mu} = q_2\, t'^{\nu}_{\mu} + q_1\, t''^{\nu}_{\mu}\,.$$

By the Chinese remainder theorem, this map is one-to-one and onto.

Then we may write

$$T_{\ell,q}\, F(s'q_1 + s''q_2) = \left(T'_{\ell_1,q_1} \circ \left(T'_{\ell_2,q_2}\right) F\right)\, (s'q' + s''q_2)\,,$$

where T'_{ℓ_2,q_2} acts only on the t'' variable, and T'_{ℓ_1,q_1} acts on the t' variables. Then Lemma 10.2 follows.

In view of Lemma 10.2 it suffices to show that for all but a finite set of primes, p,

(10.2) $\| T'_{(\ell,p^r)} \|^2 \leq p^{(2d'-1)r}$

and for all primes

(10.3) $\| T'_{(\ell,p^r)} \| \leq C(p)\, p^{(2d'-1)r}\,.$

To prove (10.2) and (10.3) we will employ Lemma 9.2. As in Lemma 9.2, put

$$\phi^1_{e,f} \;=\; \sum_v (-A^v_{e,g} \;+\; B^v_{e,f})\,\theta^v$$

and

$$\phi^2_{e,f} \;=\; \sum_v B^v_{e,f}\,\theta^v\,.$$

Then by Lemma 9.2, we may write

$$Q'\theta^v \;=\; \sum_{e,f} D'^1_{e,f}\,\phi^1_{e,f} \;+\; D'^2_{e,f}\,\phi^2_{e,f} \tag{10.4}$$

where Q', $D'^1_{e,f}$ and $D'^2_{e,f}$ are integers with a bound depending only on the coefficients of the polynomials arising in the group multiplication.

Now since $A^v_{e,f}$ and $B^v_{e,f}$ are integral, each of the $\varphi^i_{e,f}$ equal $\frac{b^k_{e,f}}{p^r}$ for some integers $b^i_{e,f}$.

There are two cases: either p divides all the $b^i_{e,f}$, or p does not divide at least one of them. Let us begin with the second case and assume p does not devide $b^1_{1,1}$. Let $K(t,t')$ denote the kernel of $T'^*_{\ell,q}\; T'_{\ell,q}$. Then

$$K(t,t') \;=\; \sum_{s_1,\cdots s_{k-2}} \;\sum_{s^2_{k-1}\cdots s^{d(k-1)}_{k-1}}$$

$$\sum_{s^1_{k-1}} \exp 2\pi i s^1_{k-1} \left\{ \frac{b^1_{1,1}}{p^r}\,(t^1_1 - t'^1_1) \;+\; \sum_{e\,\geq\,2} \frac{b^1_{1,e}}{p^r}\,(t^e_1 \;-\; t'^e_1) \right\}\,.$$

Since $(b^1_{1,1},\, p) \;=\; 1$, $b^1_{1,1}$ has an inverse $\bmod p^r$. Hence, for $0 \;\leq\; t'^1_1 \;\leq\; p^r - 1$, there is only one value of t'_1 such that

$$b^1_{1,1}(t^1_1 \;-\; t'^1_1) \;+\; \sum_{e\,\geq\,2} b^1_{1,e}(t^e_1 \;-\; t'^e_1) \;\equiv\; 0(p^r)\,.$$

Thus the sum on s^1_{k-1} is non-zero for at most one value of t'_1, $0 \;\leq\; t'_1 \;\leq\; p^r - 1$. Thus

$$\sum_t |K(t,t')| \;\leq\; p^{r(2d'-1)}\,.$$

Similarly,

$$\sum_{t'} |K(t,t')| \leq p^{r(2d'-1)}.$$

This shows

$$\| T'_{\ell,p^r} \|^2 \leq p^{r(2d'-1)}$$

in this case.

In the first case, if p divides all the integers $b^i_{e,f}$, then p must necessarily divide Q', because it does not divide $\ell_1, \ell_2, \ldots \ell_{d(k)}$. Thus if m_0 is the largest integer so that p^{m_0} divides Q', a similar argument shows

$$\| T'_{\ell,p^r} \| \leq p^{m_0/2} p^{r(d'-\frac{1}{2})}.$$

This finishes the proof of Lemma 10.2 and hence of (8.7).

§11. The estimate for M_j^2

The purpose of this section is to prove the estimater (7.4). To obtain this it suffice to show that for $j^\gamma \leq 2^\kappa \leq 2^{\epsilon j}$

(11.1) $$\| \mathcal{M}_{\kappa,j} f \|_{\ell^2} \leq C 2^{-\kappa/2} \| f \|_{\ell^2}$$

uniformly in j.

In fact, since $|M_j^2(f)| \leq \sum_{j^\gamma \leq 2^\kappa} |\mathcal{M}_{\kappa,j}(f)|$, then (11.1) would give $\| M_j^2 \| \leq c j^{-\gamma/2}$, which is (7.4). So because $\sup_j |M_j^2(f)|^2 \leq \sum_{j=1}^{\infty} |M_j^2(f)|^2$, the convergence of $\sum_{j\geq 1} j^{-\gamma}$ would yield the desired control of $\sup_j |M_j^2(f)|$.

As before,

$$\chi_\lambda \left(\theta - \frac{\ell}{q}\right) \chi_{c2^{2\kappa}} \left(\theta - \frac{\ell}{q}\right) = \chi_\lambda \left(\theta - \frac{\ell}{q}\right),$$

because $c2^{2\kappa} \leq \lambda = 2^{j(k-\epsilon)}$. Thus, if $2^\kappa \leq q < 2^{\kappa+1}$,

$$\mathcal{M}_j^{(\ell,q)} f(\alpha) = \int\limits_{T^{d(k)}} e^{-2\pi i\, a_k \cdot \theta}\, \chi_\lambda \left(\theta - \frac{\ell}{q}\right) S_j^\theta\, \hat{f}^{(\ell,q)}(\cdot, 0)\,(a)\, d\theta\,,$$

where $f^{(\ell,q)}$ is defined by the relation

$$\hat{f}^{(\ell,q)}(\theta) = \chi_{c2^{2\kappa}} \left(\theta - \frac{\ell}{q}\right) \hat{f}(\theta)\,.$$

For any f, not only are the supports of the $\hat{f}^{(\ell,q)}$ disjoint, but also the supports of the $\widehat{\mathcal{M}_j^{(\ell,q)} f^{(\ell,q)}}$ are disjoint. So to prove the estimate (11.1), it suffices to prove that for each (ℓ, q)

(11.2) $\| \mathcal{M}_j^{(\ell,q)} f \|_{\ell^2} \leq Cq^{-1/2} \| f \|_{\ell^2}$.

To obtain this estimate we use the arguments in Sections 7-10, but in a simpler setup. This is because it suffices to make the key estimates for each q, with $2^\kappa \leq q \leq 2^{\kappa+1}$, instead of the corresponding estimate for the collection of q's (whose product is Q).

We proceed throughout with the sub-group $q\Gamma$ in place of $Q\Gamma$. We begin, as in Section 8, by replacing $\mathcal{M}_j^{(\ell,q)}$ by $\mathcal{M}_j'^{(\ell,q)}$, giving an error (see (8.4))

$$\| \Delta_j^{(\ell,q)} \| \leq cq^k 2^{-\eta j}\,, \text{ with } \eta = 1/k - \epsilon\,.$$

Suppose we could prove that

(11.3) $\| \mathcal{M}_j'^{(\ell,q)} \| \leq cq^{1/2}$.

Then we would have

$$\| \mathcal{M}_j^{\ell,q} \| \leq cq^{-1/2}\,, \text{ because } q^k 2^{-\eta j} \leq cq^{-1/2}$$

whenever $q \leq 2^{\epsilon j}$ and $\eta = 1/k - \epsilon$, in view of the fact that $\epsilon \leq \frac{1}{k(k+3/2)}$. Thus would have established (11.2) and therfore (11.1).

To prove (11.3) we factor the operator into the corresponding tensor product, where the operator $\mathcal{N}_j$ is defined as in ((8.5*)), but now with Q replaced by q. The main simplification occurs in that we need only observe that uniformly in j.

$$\| \mathcal{N}_j \| \leq cq^{-d}\,.$$

This is a simple ℓ^2 estimate, as opposed to the more difficult maximal estimate (8.6).

Also the estimate (8.7) is replaced by the parallel estimate with q instead of Q, with the same proof as in Section 10. This then yields (11.3) and therefore (11.1).

§12. The error term

We shall now prove (7.5).

Since $E_j(f)(\cdot, a_k) = \int\limits_{\mathcal{E}_j} e^{-2\pi i a_k \theta}\, B(\theta)\, S_j^\theta\, \hat{f}(\cdot,\theta)\, d\theta$ and $0 \le B(\theta) \le 1$, it suffices by Plancherel's theorem to prove that uniformly in j,

(12.1) $\| S_j^\theta \| \le c 2^{-\eta j}$, for $\theta \in \mathcal{E}_j$,

and η a (small) positive number.

Recall from (7.1*), that

$$S_j^\theta(F)(m) = \sum_{n\in\mathbb{Z}^{d'}} \psi_j(n)\, e^{-2\pi i[P_k(m,n)+R(n)]\cdot\theta}\, F(m\circ n)$$

$$= \sum_{n\in\mathbb{Z}^{d'}} \psi_j(m^{-1}\circ n)\, e^{2\pi i r(m,n)\cdot\theta}\, f(n)\,,$$

where $r(m,n)$ is given by equation (9.7). Here the coefficients of the principal terms of $r(m,n)$ (the monomials $m_{k-1}^e n_1 f$ or $n_{k-1}^e m_1^f$) are respectively designated by $\varphi_{e,f}^1$ and $\varphi_{e,f}^2$. Now in (10.4) we have inverted the relation between θ and the φ's and have obtained

(12.2) $Q'\theta^v = \sum\limits_{e,f} D'^1_{e,f}\, \varphi^1_{e,f} + D'^2_{e,f}\, \varphi^2_{e,f}), \quad 1 \le v \le d(k).$

where Q', $D'^1_{e,f} = D'^1_{e,f}(v)$, and $D'^2_{e,f} = D'^2_{e,f}(v)$ are fixed integers.

Now let $N = 2d(1)\cdot d(k-1)$, the number of different indices (i,e,f) for $D'^i_{e,f}$. Set $D = \sum |D^i_{e,f}|$.

Our claim is that of $\theta \in \mathcal{E}_j$ then for at least one index (i,e,f) of the N possible choices, there are integers $a^i_{e,f}$, $q^i_{e,f}$ with $(a^i_{a,f}, q^i_{e,f}) = 1$, $\dfrac{2^{\epsilon j/N}}{Q'} \le q^i_{e,f} \le \lambda D$, (recall that $\lambda = 2^{j(k-\epsilon)}$), so that

(12.3) $|\varphi^i_{e,f} - \frac{a^i_{e,f}}{q^i_{e,f}}| \leq \frac{1}{q^i_{e,f}\lambda D}$.

In fact, such an approximation exists for *all* (i,e,f) by Dirichlet's principle, without however the assurance $2^{\epsilon j/N}/Q' \leq q^i_{e,f}$. Suppose the assertion fails, because the reverse inequality $q^i_{e,f} \leq 2^{\epsilon j/N}/Q'$ holds for all choices. Then picking the last common multiple of Q' and all the $q^i_{e,f}$, we could find a q, with $q \leq 2^{\epsilon j}$, and $\ell_1, \dots \ell_{d(k)}$ so that $(q_1\ell_1, \dots, \ell_{d(k)}) = 1$, and because of (12.2) and (12.3) we would have

$$|\theta^v - \ell_v/q| \leq 1/\lambda\,, \text{ for } 1 \leq v \leq d(k)\,.$$

This means $\chi_\lambda(\theta - \ell/q) = 1$, and thus $B(\theta) = 0$, that is $\theta \notin \mathcal{E}_j$. This is a contradiction, and so (12.3) must hold for at least one (i,e,f), which for simplicity we take to be the triple $(1,1,1)$.

Once we have (12.3) our desired estimate falls in the framework of known estimates for opertor Weyl-sums; see [SW2], Proposition 5. However we can also prove (12.1) directly as follows. It suffices to estimate the norm of the operator $(S^\theta_j)^* S^\theta_j$. Its kernel $K(m,n)$ is given by

$$(12.4) \qquad \sum_{a \in \mathbb{Z}^{d'}} \psi_j(a^{-1} \circ n)\, \bar{\psi}_j(a_a^{-1} \circ m)\, e^{2\pi i (r(a,n) - r(a,m))\cdot\theta},$$

and it suffices to see that

$$(12.5) \qquad \sum_{n \in \mathbb{Z}^{d'}} |K(m,n)| \leq c 2^{-2\eta j}.$$

Recalling the equation (9.7) we see that

$$(r(a,n) - r(a,m)) \cdot \theta = a^1_{k-1}\,[\varphi^1_{11}(n^1_1 - m^1_j) + \sum_{\ell=2}^{d(k)} \varphi^1_{1,\ell}(n^1_1 - m^1_1)]$$

$$+ \text{ terms that do not involve } a^1_{k-1}\,.$$

In the sum (12.4) we sum first in the a^1_{k-1} variable, and then in all the other a variables. Schematically,

$$|K(m,n)| \leq \sum_{\substack{\text{other } a \\ \text{variables}}} |\sum_{a^1_{k-1}}|\,.$$

We see that the inner sum is majorized by

$$c2^{-jD'}\,2^{-jD'}\,\min\,[2^{j(k-1)}\,,\,\{\varphi^1_{11}(n^1_1-m^1_1)\,+\,\sum_{\ell=1}^{d(k)}\varphi^1_{1,\ell}\,(n^\ell_1-n^\ell_1)\}^{-1}]$$

where $\{\,\cdot\,\}$ denotes the fractional part.

Next, we sum this in n^1_1 and simplify the notation by writing q for $q^1_{1,1}$. We follow the usual argument (as in [M0], §2 in Chapter 3) where we break the range of n^1_1 into essentially $2^j/q$ blocks of length q, and use (12.3) for φ^1_{11}. This gives an estimate

$$c(2^{-jD'}\,2^{-jD'}\,2^{j(k-1)}\,\cdot\,2^j/q\,+\,2^{-jD'}\,2^{-jD'}\,(\log q)\,2^j/q)\,.$$

We must still sum over all the remaining a variables (introducing a factor $2^{jD'}\,2^{-j(k-1)}$) and all the remaining n variables, (introducing a further factor of $2^{jD'}\,2^{-j}$). This then shows that $\sum_n|K(m,n)|\,\le\,c/q\,\le\,c2^{-2\eta j}$, with $\eta\,=\,\epsilon/2N$, since $q\,\ge\,2^{\epsilon j/N}/Q'$. Thus (12.5) is proved and (12.1) is established.

References

[AO] G.I. Arkhipov and K. Oskolkov, "A trigonometrical series and its application," translation in *Math.*, USSR - Sb., **62**, (1989), 145-155.

[BO1] J. Bourgain, "On the maximal ergodic theorem for certain subsets of the integers," *Isreal J. Math.*, **61**, (1988), 39-72.

[BO2] J. Bourgain, "On the pointwise ergodic theorem on L^p for arithmetic sets," *Israel J. Math.*, **61**, (1988), 73-84.

[BO3] J. Borugain, "Pointwise ergodic theorems for arithmetic sets," *Inst. Hautes Etudes Sci. Publ. Math.*, **69**, (1989), 5-45.

[CG] L. Corwin and F.P. Greenleaf, *Representations of nilpotent Lie groups and their applications, Part 1: Basic theory and examples*, Cambridge University Press, (1990).

[CNSW] M. Christ, A. Nagel, E.M. Stein, and S. Wainger, "Singular and maximal Radon tranforms: analysis and geometry," *Annals of Math.*, **150**, (1999), 489-577.

[IMSW] A.D. Ionescu, A. Magyar, E.M. Stein, and S. Wainger, "Discrete Radon transforms on nilpotent groups," manuscript.

[IW] A.D. Ionescu and S. Wainger, "L^p estimates for discrete singular Radon transforms," (to appear in *Jour. of A.M.S.*).

[M] A. Magyar, "Discrete maximal functions and ergodic theorems related to polynomials," (to appear).

[MO] H. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, *CBMS*, Regional Conference Series in Mathematics, **84**, 1994.

[RS] F. Ricci and E.M. Stein, "Harmonic analysis on nilpotent groups and singular integrals II," *J. Funct. Analysis*, **78**, (1988), 56-84.

[SW1] E.M. Stein and S. Wainger, "Discrete analogues of singular Radon transforms," *Bull. Amer. Math. Soc.*, **23**, (1990), 537-544.

[SW2] E.M. Stein and S. Wainger, "Discrete analogues in harmonic analysis, 1.," *Amer. J. Math.*, **121**, (1999), 1291-1336.

Akos Magyar
Department of Mathematics
University of Georgia
Athens, GA 30602
email: magyar@math.uga.edu

Elias M. Stein
Department of Mathematics
Princeton University
Princeton, NJ 08544, USA
email: stein@math.princeton.edu

Stephen Wainger
Department of Mathematics
University of Wisconsin
Madison, WI 53706, USA
email: wainger@math.wisc.edu

This work was supported by &

March 8, 2006:gpp.