C:MeanLargeVals

GUTH-MAYNARD SETUP

1. Introduction and Notation

I will follow the notation of Guth-Maynard. Let $\mathcal{W}$ be a set of real numbers t with $|t-t'|\ge 1$ for $t,t'\in\mathcal{W}$ and $\mathcal{M}$ denote the card $W\times N$ matrix with general entries

$$M_{t,n} = w(n/N)n^{it} \quad (t\in\mathcal{W}, N<n\le 2N) \tag{1.1}$$ eq:one1

with

$$w\in\mathcal{C}^\infty(\mathbb{R}),\ \operatorname{supp} w\in[1,2],\ w(x)=1\,(x\in[6/5,9/5]). \tag{1.2}$$ eq:one2

Guth and Maynard employ the bizarre notation $\mathcal{M}_{\mathcal{W}}$, but no N! Also, $w(n/N)$ will have a peak when $\frac{3}{2}-\frac{3}{10}\le\frac{n}{N}\le\frac{3}{2}+\frac{3}{10}$. In other word when $\frac{n}{2N}$ is close to $\frac{3}{4}$. How this relates to *sigma* being near to $\frac{3}{4}$ is not explained.

Then

$$D_N(t)=\sum_{n=N+1}^{2N} b_n w(n/N)n^{it} \tag{1.3}$$ eq:one3

is the entry corresponding to t in the column vector

$$\mathcal{M}\mathbf{b}. \tag{1.4}$$ eq:one4

By the way I usually think of vectors as being row vectors, so I may get confused at some point! Please bear with me.

Ultimately we may be concerned with

$$D_N(s)=\sum_{n=N+1}^{2N} b_n w(n/N)n^{-s}.$$

In practice we can obtain bounds first when $\sigma=0$ and t is replaced by $-t$, and then take complex conjugates and finally apply Gallagher's partial integration method. They also assume that $|b_n|\ge 1$, although in practice one needs to weaken that to $|b_n|\ll n^\varepsilon$. Since they don't mind losing a T^ε in the final conclusion that can be done easily by renormalising the b_n. They also use the sloppy notation $T^o(1)$ for T^ε although that would imply that the quantity it represents satisfies

$$T^{-\varepsilon}<T^{o(1)}<T^\varepsilon$$

whereas I am betting they really only mean the second inequality.

A basic observation is that

$$\sum_{t\in\mathcal{W}} |D_N(t)|^2 = \sum_{m=N+1}^{2N} \overline{b}_m w(m/N) \sum_{n=N+1}^{2N} b_n w(n/N) \sum_{t\in\mathcal{W}} (n/m)^{it} = \mathbf{b}^*(\mathcal{M}^* \times \mathcal{M})\mathbf{b}. \quad (1.5)$$

eq:biform

2. Some Linear Algebra

At this point let me remind you of some basic linear algebra. Let $A = [a_{mn}]$ be an $M \times N$ matrix with complex entries. Then A determines a linear map $\mathbf{x} \mapsto \mathbf{y} = A\mathbf{x}$ from $\mathbb{C}^N$ to $\mathbb{C}^M$. The norm of A, as a linear operator, is

$$\|A\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$$

where $\|\mathbf{x}\| = \big(\sum |x_n|^2\big)^{1/2}$ denotes the usual Euclidean norm. By homogeneity we may write instead

$$\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

We are particularly interested in the smallest λ such that

$$\Big|\sum_{m,n} a_{mn} x_n y_m\Big| \le \lambda \Big(\sum_n |x_n|^2\Big)^{1/2} \Big(\sum_m |y_m|^2\Big)^{1/2}$$

holds for all non-zero $\mathbf{x}, \mathbf{y}$.

T:BilinDuality

Theorem 2.1 (Duality). *Let $A = [a_{mn}]$ be a fixed $M \times N$ matrix. The following three assertions concerning the positive number τ are equivalent.*

1. For any $\mathbf{x} \in \mathbb{C}^N$,

$$\sum_{m=1}^{M} \Big|\sum_{n=1}^{N} a_{mn} x_n\Big|^2 \le \tau^2 \sum_{n=1}^{N} |x_n|^2.$$

2. For any $\mathbf{x} \in \mathbb{C}^N$ and any $\mathbf{y} \in \mathbb{C}^M$,

$$\Big|\sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn} x_n y_m\Big| \le \tau \Big(\sum_{n=1}^{N} |x_n|^2\Big)^{1/2} \Big(\sum_{m=1}^{M} |y_m|^2\Big)^{1/2}.$$

3. For any $\mathbf{y} \in \mathbb{C}^M$,

$$\sum_{n=1}^{N} \Big|\sum_{m=1}^{M} a_{mn} y_m\Big|^2 \le \tau^2 \sum_{m=1}^{M} |y_m|^2.$$

In terms of linear maps and inner products, these inequalities assert that

1. $$\|A\mathbf{x}\| \le \tau\|\mathbf{x}\|,$$
2. $$|(A\mathbf{x},\mathbf{y})| \le \tau\|\mathbf{x}\|\|\mathbf{y}\|,$$
3. $$\|A^*\mathbf{y}\| \le \tau\|\mathbf{y}\|$$

where A^* is the *adjoint* of A. That is, $A^* = (\overline{A})^{\mathrm{T}}$ is the $N \times M$ matrix $A^* = [\overline{a_{nm}}]$. In terms of inner products, A^* is characterized by the property that $(A\mathbf{x},\mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$ for all $\mathbf{x}$ and $\mathbf{y}$. Since 1. and 3. are equivalent, we deduce that

$$\|A\| = \|A^*\|.$$

The matrix $\mathcal{H} = \mathcal{A}^* \times \mathcal{A}$ is square, $N \times N$, and Hermitian. For the time being let $\mathcal{H}$ be an arbitrary $N \times N$ Hermitian matrix $\mathcal{H} = (h_{mn})$. Its eigenvalues λ_k are real and non-negative - here multiple values are counted multiply. [Note that if $\mathbf{v}$ is an eigenvector associated with the eigenvalue λ, then

$$(\mathbf{v}^*\mathcal{A}^*) \times (\mathcal{A}\mathbf{v}) = \|\mathcal{A}\mathbf{v}\| \ge 0$$

and

$$(\mathbf{v}^*\mathcal{A}^*) \times (\mathcal{A}\mathbf{v}) = \mathbf{v}^*(\mathcal{A}^* \times \mathcal{A})\mathbf{v} = \mathbf{v}^*\mathbf{v}\lambda$$

so $\lambda \ge 0$.

Also there is an orthogonal basis of eigenvectors $\mathbf{e}_k$ for $\mathbb{C}^N$ and we can suppose $\|\mathbf{e}_k\| = 1$. Thus given $\mathbf{b}$ there are $\mathbf{c}$ so that

$$\mathbf{b} = \sum_k c_k\mathbf{e}_k$$

and then

$$\mathbf{b}^*\mathcal{H}\mathbf{b} = \sum_j \overline{c}_j e_j \cdot \sum_k c_k\lambda_k\mathbf{e}_k = \sum_j |c_j|^2\lambda_j$$

Thus

$$\mathbf{b}^*\mathcal{H}\mathbf{b} \le \sum_j |c_j|^2 \max_k \lambda_k = \|\mathbf{b}\|^2 \max_k \lambda_k$$

so we are particularly interested in the largest eigenvalue.

The characteristic polynomial $P(x)$ of $\mathcal{H}$ is given by

$$P(x) = \det(x\mathbf{I} - \mathcal{H}) = \prod_k (x - \lambda_k).$$

Let $\mathrm{tr}(\mathcal{B})$ of any square matrix $\mathcal{B}$ denote the sum of the diagonal elements. It is relatively easy to show that

$$P(x) = x^N - \mathrm{tr}(\mathcal{H})x^{N-1} + \cdots$$

by induction on N. Expand $\det(x\mathbf{I} - \mathcal{H})$ along the top row. Then the second and subsequent terms are of the form $-h_{1,n}\det\big(\mathcal{H}_n(x)\big)$

where the minor $\mathcal{H}_n(x)$ is $(N-1)\times(N-1)$ and has no x in the first column, so $\det\big(\mathcal{H}_n(x)\big)$ is of degree at most $N-2$ in x. Applying the inductive hypothesis to the first term $(x-h_{11})\det(\mathcal{H}_1(x)$ gives the desired conclusion.

It follows that

$$\sum_{k=1}^{N}\lambda_k = \operatorname{tr}\mathcal{H}.$$

3. Singular Values

In §4 Guth and Maynard discuss singular values $s_j(\mathcal{M})$ of $\mathcal{M}$ and associated matrices. I think that is a bit pointless since ultimately these things are evaluated in terms of the eigenvalues and Lemma 4.2 is really just a trivial inequality in which eigenvalues and traces are substituted.

In general the singular values are defined for $M\times N$ matrices, with M not necessarily equal to N, in terms of extremal values on subspaces of $\mathbb{C}^n$. From our point of view we could define

$$s_k(\mathcal{M})^2 = \lambda_k(\mathcal{M}^*\times\mathcal{M})$$

and take the positive sign. They then order the eigenvalues so that

$$\lambda_1\ge\lambda_2\ge\ldots\ge 0$$

and so

$$s_1(\mathcal{M})\ge s_2(\mathcal{M})\ge\ldots\ge 0.$$

4. Reduction and Fourier Analysis

Their Lemma 4.1 is a bit misleading I think, and is nothing more than what follows from the observation that

$$\sum_{t\in\mathcal{W}}|D_N(t)|^2\le\|\mathbf{b}\|^2\lambda_1(\mathcal{M}^*\mathcal{M}) = \|\mathbf{b}\|^2 s_1(\mathcal{M})^2,$$

Lemma 4.1 (Lemma 4.1)**.** *If we suppose that $|D_N(t)|\ge V$ for $t\in\mathcal{W}$, then*

$$V^2\operatorname{card}W\le\|b\|^2 s_1(\mathcal{M})^2$$

and so if also $\|b\|^2\le N$ and $V=N^\sigma$, then

$$\operatorname{card}W\le N^{1-2\sigma}s_1(\mathcal{M})^2.$$

lem:four2 **Lemma 4.2** (Lemma 4.2)**.** *Let $\mathcal{A}$ be an $M\times N$ matrix with entries in $\mathbb{C}$ and define $\mathcal{H}=\mathcal{A}\mathcal{A}^*$. Then with the above notation*

$$s_1(\mathcal{A})=\lambda_1(\mathcal{H})^{\frac12}\le 2^{1/6}\left(\operatorname{tr}\left(\mathcal{H}^3\right)-m^{-2}\operatorname{tr}(\mathcal{H})^3\right)^{1/6}+6^{1/4}\big(m^{-1}\operatorname{tr}(\mathcal{H})\big)^{1/2}$$

and

$$\lambda_1(\mathcal{H}) \ll \left(\operatorname{tr}\left(\mathcal{H}^3\right) - m^{-2}\operatorname{tr}(\mathcal{H})^3\right)^{1/3} + \left(m^{-1}\operatorname{tr}(\mathcal{H})\right)$$

The constants $2^{1/6}$ and $6^{1/4}$ are what arise fairly naturally in the argument, but are not very important and are both replaced by 2 in Guth-Maynard. It would be good to understand precisely in what circumstance this gives significant bounds for $\lambda_1(\mathcal{A})$. Clearly if there is one very dominant eigenvalue it is useless. The right hand side will be close to

$$2^{1/6}\lambda_1^{1/2}$$

for large m. On the other hand if

$$m^{-1}\operatorname{tr}\left(\mathcal{H}^3\right) - m^{-3}\operatorname{tr}(\mathcal{H})^3 \le m^{-1}\Delta^3$$

and $m^{-1}\operatorname{tr}(\mathcal{H}) \le \Delta$, then

$$\lambda_1 \ll \Delta.$$

It would not matter that Δ is large as long as it improves on previous results. The motivation is that generally in the kind of situation we are studying $\operatorname{tr}\left(\mathcal{H}^3\right)$ is close to $m^{-2}\operatorname{tr}(\mathcal{H})^3$ and then we can take advantage of the m^{-1} in the other term.

Before embarking on the proof we should observe that if λ is an eigenvalue of $\mathcal{H}$, and $\mathbf{v}$ a corresponding eigenvector, then

$$\mathcal{H}^r\mathbf{v} = \lambda^r\mathbf{v}$$

so λ^r is an eigenvalue of $\mathcal{H}^r$. Moreover an orthogonal system of eigenvectors for $\mathcal{H}$ will also be an orthogonal system of eigenvectors for $\mathcal{H}^r$, so the multiplicities of the eigenvalues remains the same. Hence if $\lambda_1, \dots, \lambda_m$ are the eigenvectors of $\mathcal{H}$ counted with multiplicity, then so are $\lambda_1^r, \dots, \lambda_M^r$ for $\mathcal{H}^r$. Moreover

$$\operatorname{tr}(\mathcal{H}^r) = \sum_{k=1}^{m} \lambda_k^r.$$

Proof. This inequality is equivalent to

$$\lambda_1^{1/2} \le 2^{1/6}\left(\sum_{k=1}^{m}\lambda_k^3 - m^{-2}\left(\sum_{k=1}^{m}\lambda_k\right)^3\right)^{1/6} + 6^{1/4}\left(m^{-1}\sum_{k=1}^{m}\lambda_k\right)^{1/2} \tag{4.1}$$

eq:one6

where we now assume only that λ_1 is the largest of the λ_k and all the λ_k are non-negative. By Hölders' inequality

$$\left(\sum_{k=2}^{m}\lambda_k\right)^3 \le (m-1)^2\sum_{k=2}^{m}\lambda_k^3 \le m^2\sum_{k=2}^{m}\lambda_k^3$$

and so

$$\lambda_1^3 \le \sum_{k=1}^m \lambda_k^3 - m^{-2}\left(\sum_{k=2}^m \lambda_k\right)^3.$$

We also have

$$\begin{aligned}\left(\sum_{k=1}^m \lambda_k - \lambda_1\right)^3 &= \left(\sum_{k=1}^m \lambda_k\right)^3 - 3\left(\sum_{k=1}^m \lambda_k\right)^2\lambda_1 + 3\left(\sum_{k=1}^m \lambda_k\right)\lambda_1^2 - \lambda_1^3 \\ &\ge \left(\sum_{k=1}^m \lambda_k\right)^3 - 3\left(\sum_{k=1}^m \lambda_k\right)^2\lambda_1.\end{aligned}$$

Thus

$$\lambda_1^3 \le \sum_{k=1}^m \lambda_k^3 - m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^3 + 3m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^2\lambda_1.$$

For any non-negative numbers X and Y we have

$$(X+Y)^{1/6} \le \big(2\max(X,Y)\big)^{1/6}.$$

Thus if

$$\sum_{k=1}^m \lambda_k^3 - m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^3 \ge 3m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^2\lambda_1.$$

then we have (4.1) at once. If on the contrary we have

$$\sum_{k=1}^m \lambda_k^3 - m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^3 < 3m^{-2}\left(\sum_{k=1}^m \lambda_k\right)^2\lambda_1,$$

then

$$\lambda_1^{1/2} \le (6m^{-2})^{1/6}\left(\sum_{k=1}^m \lambda_k\right)^{1/3}\lambda_1^{1/6}$$

and so

$$\lambda_1^{1/2} \le (6m^{-2})^{1/4}\left(\sum_{k=1}^m \lambda_k\right)^{1/2}$$

Hence

$$\lambda_1^{1/2} \le 6^{1/4}\left(m^{-1}\sum_{k=1}^m \lambda_k\right)^{1/2}$$

□

Up until page 6 the object of the analysis is $\mathcal{M}^*\mathcal{M}$. This then disappears until it makes a brief appearance in the third displayed formula on page 10. Then at the bottom of page 10 there is a switch to $\mathcal{M}\mathcal{M}^*$ without any explanation. I think the point is that the largest eigenvalue for either is the same, although one would expect that the matrices are a different size and so the total number of eigenvalues in each case is different. This way round has the double benefit that it is easier to do the fourier analysis and to get a bound for the outer dimension which is now card W.

The next step in the setup is a fourier analysis of the traces. Recall

$$\mathcal{M} = \big(w(n/N)n^{it}\big) \quad (t \in \mathcal{W}, N < n \le 2N)$$

where

$$w \in \mathcal{C}^\infty(\mathbb{R}),\ \operatorname{supp} w \in [1,2],\ w(x) = 1\,(x \in [6/5, 9/5]).$$

Then the general entry in

$$\mathcal{M}^*\mathcal{M}$$

is

$$\sum_{t \in \mathcal{W}} w(m/N)w(n/N)(m/n)n^{it}$$

and that in

$$\mathcal{M}\mathcal{M}^*$$

ia

$$\sum_{n=N+1}^{2N} w(n/N)^2 n^{it-it'}.$$

It is the latter which motivates the study of

$$h_t(u) = w(u)^2 u^{it}$$

and its Fourier transform

$$\hat{h}_t(v) = \int_{\mathbb{R}} h_t(u)e(-uv)du$$

lem:four3 **Lemma 4.3** (Lemma 4.3)**.** *Suppose j is a positive integer. Then*

$$\hat{h}_t(v) \ll_j (1+|t|)^j |v|^{-j}$$

and

$$\hat{h}_t(v) \ll_j (1+|v|)^j |t|^{-j}$$

Proof. We have

$$\hat{h}_t(v) = \int_{\mathbb{R}} w(u)^2 u^{it} e(-uv)du.$$

The stated bounds follow by integrating by parts with either

$$(w(u)^2u^{it})e(-uv)$$

or

$$u^{it}\big(w(u)^2e(-uv)\big).$$

□

lem:four4 **Lemma 4.4** (Lemma 4.4). *Suppose that* $\mathcal{W} \in \mathbb{R}$ *and there is an* $\varepsilon > 0$ *such that* $\operatorname{card}\mathcal{W} \ll_\varepsilon N^\varepsilon$. *Then*

$$\operatorname{tr}(\mathcal{M}\mathcal{M}^*) = N(\operatorname{card}\mathcal{W})\int_{\mathbb{R}} w(u)^2du + O(N^{-100}).$$

It might be more appropriate where one sees an exponent like 100 here and below to add to the hypothesis "and suppose that A is a fixed real number ≥ 1" and rather than "(N^{-100})" write "$O_A(N^{-A})$". This is certainly true and relieves any burden of checking that the result matches later needs.

Note that

$$\int_{\mathbb{R}} w(u)^2du \asymp 1$$

so as long as $\operatorname{card}\mathcal{W} \geq 1$ we have

$$\operatorname{tr}(\mathcal{M}\mathcal{M}^*) \asymp N(\operatorname{card}\mathcal{W}).$$

Proof. Recall that the general entry in $\mathcal{M}\mathcal{M}^*$ is

$$\sum_{n=N+1}^{2N} w(n/N)^2n^{it-it'} \text{ and } h_t(u) = w(u)^2u^{it}.$$

Thus

$$\operatorname{tr}(\mathcal{M}\mathcal{M}^*) = \sum_{t\in\mathcal{W}}\sum_n w(n/N)^2 = \operatorname{card}(\mathcal{W})\sum_n h_0(n/N).$$

Let $g(u) = h_0(\frac{u}{N})$, so that

$$\hat{g}(t) = \int_{\mathbb{R}} g(u)e(-ut)du = N\hat{h}_0(tn)$$

Then by the Poisson summation formula

$$\operatorname{tr}(\mathcal{M}\mathcal{M}^*) = \operatorname{card}(\mathcal{W})\sum_n N\hat{h}_0(nN).$$

By Lemma lem:four3 4.3 the sum here is

$$N\int_{\mathbb{R}} w(u)^2du + O\left(\sum_{n\neq 0}\frac{N}{N^{101}|n|^{101}}\right).$$

□

Of course we have also to deal with the cubic trace.

lem:four5 **Lemma 4.5** (Lemma 4.5)**.** *Suppose that $\mathcal{W}$ is T^ε separated. Then*

$$\operatorname{tr}\left((MM^*)^3\right) = N^3(\operatorname{card}\mathcal{W})\left(\int_{\mathbb{R}} w(u)^2du\right)^3 + \sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} I_{\mathbf{m}} + O_\varepsilon(T^{-100})$$

where

$$I_{\mathbf{m}} = N^3 \sum_{\mathbf{t}\in\mathcal{W}^3} \hat{h}_{t_1-t_2}(m_1N)\hat{h}_{t_2-t_3}(m_2N)\hat{h}_{t_3-t_1}(m_3N).$$

Proof. This imitates what we did for the linear trace, but with more factors. Thus $\operatorname{tr}\left((MM^*)^3\right)$

$$= \sum_{\mathbf{n}\in\mathbb{Z}^3}\sum_{\mathbf{t}\in\mathcal{W}^3} w\left(\frac{n_1}{N}\right)^2 w\left(\frac{n_2}{N}\right)^2 w\left(\frac{n_3}{N}\right)^2 n_1^{i(t_1-t_2)}n_2^{i(t_2-t_3)}n_3^{i(t_3-t_1)}$$

$$= \sum_{\mathbf{t}\in\mathcal{W}^3}\sum_{\mathbf{n}\in\mathbb{Z}^3} h_{t_1-t_2}\left(\frac{n_1}{N}\right)h_{t_2-t_3}\left(\frac{n_2}{N}\right)h_{t_3-t_1}\left(\frac{n_3}{N}\right).$$

We apply Poisson summation to each sum over n_j, $j = 1, 2, 3$. Thus $\operatorname{tr}\left((MM^*)^3\right)$

$$= N^3 \sum_{\mathbf{t}\in\mathcal{W}^3}\sum_{\mathbf{m}\in\mathbb{Z}^3} \hat{h}_{t_1-t_2}(m_1N)\hat{h}_{t_2-t_3}(m_2N)\hat{h}_{t_3-t_1}(m_3N)$$

$$= \sum_{\mathbf{m}\in\mathbb{Z}^3} I_{\mathbf{m}}.$$

We separate out the term $\mathbf{m} = \mathbf{0}$. Then we can estimate terms with $t_q - t_r \neq 0$ for some $q \neq r$, since $\mathcal{W}$ is T^ε separated, by using Lemma 4.3 with, say, $j = \lceil 200/\epsilon\rceil$ to give the required result. □

We can now put the last few lemmas together to get a bound for the largest eigenvalue of $\mathcal{M}\mathcal{M}^*$. Thus by Lemma 4.4

$$\operatorname{tr}(\mathcal{M}\mathcal{M}^*) = N\Xi \operatorname{card}\mathcal{W} + O(N^{-100})$$

where

$$\Xi = \int_{\mathbb{R}} w(u)^2du.$$

and by Lemma 4.5

$$\operatorname{tr}\left((\mathcal{M}\mathcal{M}^*)^3\right) = N^3\Xi^3 \operatorname{card}\mathcal{W} + \Theta + O(T^{-100})$$

where

$$\Theta = \sum_{\mathbf{m}\in\mathbb{Z}\setminus\{\mathbf{0}\}} I_{\mathbf{m}}.$$

Thus

$$\operatorname{card}\mathcal{W}^{-2}\operatorname{tr}(\mathcal{M}\mathcal{M}^*)^3 = N^3\Xi^3 \operatorname{card}\mathcal{W} + O(N^{-98})$$

Note that $\mathcal{M}\mathcal{M}^*$ is a card $W \times$ card W matrix, and we want to apply Lemma lem:four2 4.2. Thus

$$\operatorname{tr}\left((\mathcal{M}\mathcal{M}^*)^3\right) - \operatorname{card}\mathcal{W}^{-2}\operatorname{tr}(\mathcal{M}\mathcal{M}^*)^3 = \Theta + O(T^{-100} + N^{-98})$$

and so by Lemma lem:four2 4.2 we have

$$\lambda_1(\mathcal{M}\mathcal{M}^*) \ll \Theta^{1/3} + N.$$

Therefore

$$\sum_{t\in\mathcal{W}} |D_N(t)|^2 \ll (N + \Theta^{1/3})\|\mathbf{b}\|^2$$

We have just established that

Theorem 4.6 (Proposition 4.6)**.** *Suppose that $\mathcal{W}$ is T^ε separated and $|D_N(t)| > V$ for each $t \in \mathcal{W}$. Then*

$$\operatorname{card} W \ll_\varepsilon \left(N + \Big(\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} I_{\mathbf{m}}\Big)^{1/3}\right)\|\mathbf{b}\|^2 V^{-2}$$

Thus we have reduced the investigation to bounding

$$\sum_{\mathbf{m}\in\mathbb{Z}^3\setminus\{\mathbf{0}\}} I_{\mathbf{m}} = S_1 + S_2 + S_3.$$

where S_j denotes the sum over $\mathbf{m}$ with exactly j of the m_j being non-zero. S_3 is the main focus of their paper. S_2 can be dealt with by using Heath-Brown's theorem and S_1 is essentially trivial.

5. The sum S_1

To deal with S_1 note that

$$\begin{aligned} S_1 &= I_{0,0,1} + I_{0,1,0} + I_{1,0,0} \\ &= N^3 \sum_{\mathbf{t}\in\mathcal{W}^3} \sum_{m\in\mathbb{Z}\setminus\{0\}} \big(\hat{h}_{t_1-t_2}(mN)\hat{h}_{t_2-t_3}(0)\hat{h}_{t_3-t_1}(0) \\ &\quad + \hat{h}_{t_1-t_2}(0)\hat{h}_{t_2-t_3}(mN)\hat{h}_{t_3-t_1}(0) + \hat{h}_{t_1-t_2}(0)\hat{h}_{t_2-t_3}(0)\hat{h}_{t_3-t_1}(mN)\big). \end{aligned}$$

By permuting the $\mathbf{t}$ in the second general term, $t_2 \to t_1$, $t_3 \to t_2$, $t_1 \to t_3$, and similarly for the third general term we find that

$$S_1 = 3N^3 \sum_{\mathbf{t}\in\mathcal{W}^3} \sum_{m\in\mathbb{Z}\setminus\{0\}} \hat{h}_{t_1-t_2}(mN)\hat{h}_{t_2-t_3}(0)\hat{h}_{t_3-t_1}(0).$$

The assumption in Proposition 3.1, which is presumably the "main proposition", is that $T = N^{6/5}$. At this stage one only needs an assumption of the kind

$$N^{\kappa_1} \ll T \ll N^{\kappa_2}$$

for some positive constants κ_j with $\kappa_1 \le \kappa_2$.

Suppose first that t_1, t_2, t_3 are distinct and use $'$ to denote this. By Lemma 4.3 for fixed k and j with $k \ge 2j \ge 200(1 + 1/\varepsilon)$ we have

$$\hat{h}_{t_1-t_2}(mN)\hat{h}_{t_2-t_3}(0)\hat{h}_{t_3-t_1}(0) \ll \frac{1 + |t_1 - t_2|^j}{|m|^j N^j (1 + |t_2 - t_3|^k)(1 + |t_3 - t_1|^k)}$$

Summing over m we obtain

$$S_1' \ll N^{3-j} \sum_{\mathbf{t} \in \mathcal{W}^3}{}' \frac{1 + |t_3 - t_1|^j + |t_2 - t_3|^j}{(1 + |t_2 - t_3|^k)(1 + |t_3 - t_1|^k)}$$
$$\ll \sum_{\mathbf{t} \in \mathcal{W}^3}{}' \frac{N^{3-j}}{1 + |t_2 - t_3|^j} + \frac{N^{3-j}}{1 + |t_3 - t_1|^j}$$

Since the t_l are T^ε separated the contribution from the terms with t_1, t_2, t_3 distinct is

$$\ll N^{-100} T^{-98}.$$

Suppose $t_1 = t_2 \ne t_3$. Then

$$\hat{h}_{t_1-t_2}(mN)\hat{h}_{t_2-t_3}(0)\hat{h}_{t_3-t_1}(0) = \hat{h}_0(mN)|\hat{h}_{t_1-t_3}(0)|^2 \ll \frac{1}{|m|^j N^j (1 + |t_1 - t_3|^{2k})}$$

Summing over all such $\mathbf{t}$ and $m \ne 0$ gives a bound

$$\ll N^{-100} T^{-399}$$

We likewise obtain the same bound for terms with $t_1 = t_3 \ne t_2$ and $t_2 = t_3 \ne t_1$.

Finally the terms with $t_1 = t_2 = t_3$, and $m \ne 0$ satisfy

$$\hat{h}_0(mN)\hat{h}_0(0)\hat{h}_0(0) \ll |m|^{-j} N^{-j}$$

and such terms collectively contribute

$$\ll T N^{-j} \ll T^{-100}$$

provided that $N > T^\varepsilon$, which is surely ensured by their implicit assumptions when ε is sufficiently small. Thus we have established.

Theorem 5.1 (Proposition 5.1)**.** *We have*

$$S_1 \ll_\varepsilon T^{-10}.$$