DECOUPLING INEQUALITIES AND SOME MEAN-VALUE THEOREMS

JEAN BOURGAIN

Abstract. The purpose of this paper is to present some further applications of the general decoupling theory from [B-D1, 2] to certain diophantine issues. In particular, we consider mean value estimates relevant to the Bombieri-Iwaniec approach to exponential sums and arising in the work of Robert and Sargos [R-S]. Our main input is a new mean value theorem.

0. Summary

The aim of this Note is to illustrate how a version of the general decoupling inequality for hypersurfaces established in [B-D] permits to recover certain known mean-value theorems in number theory and establish some new ones. Easy applications in this direction were already pointed out in [B-D] and the material presented here is a further development. Our main emphasis will be on the method rather than the best exponents that can be obtained this way.

In the first section, we state a form of the main decoupling theorem from [B-D] to the situation of smooth hyper surfaces in $\mathbb{R}^n$ with non-degenerate (but not necessarily definite) second fundamental form (a detailed argument appears in [B-D2]). The motivation for this appears in Sections 2 and 3, which aims at proving decoupling inequalities for real analytic curves $\Gamma \subset \mathbb{R}^n$ not contained in a hyperplane. The assumption of real analyticity is purely for convenience (it suffices for the subsequent applications) and a similar result also holds in the smooth category. An $(n-1)$-fold convolution of $\Gamma$ leads indeed to a hypersurface $S \subset \mathbb{R}^n$ of non-vanishing curvature. The relevant statement is inequality (3.2) below with moment $q = 2(n+1)$, where we consider the multi-linear (i.e. $(n-1)$-linear) setting. The next step is to reformulate this inequality as a mean-value theorem for exponential sums stated as Theorem 1, which is a quite general and optimal result. Our first focus point are certain mean value inequalities arising in the Bombieri-Iwaniec
approach [B-I1, 2] to exponential sums and the subsequent developments of this technique (see [H] for the complete exposition). More specifically, Theorem 1 is relevant to the so-called ‘first spacing problem’ which is analytically captured by mean-value expressions of the type

$$N_8(\delta) = \int_0^1 \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(x_0 n + x_1 n^2 + x_2 \frac{1}{\delta} \left( \frac{n}{N} \right)^\frac{3}{2}) \right|^8 dx_0 dx_1 dx_2 \quad (0.1)$$

$$N_{10}(\delta, N\delta) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(x_0 n + x_1 n^2 + x_2 \frac{1}{\delta} \left( \frac{n}{N} \right)^\frac{3}{2} + x_3 \frac{1}{N\delta} \left( \frac{n}{N} \right)^\frac{1}{2}) \right|^{10} dx_0 dx_1 dx_2 dx_3 \quad (0.2)$$

and

$$N_{12}(\delta, N\delta) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(x_0 n + x_1 n^2 + x_2 \frac{1}{\delta} \left( \frac{n}{N} \right)^{3/2} + x_3 \frac{1}{N\delta} \left( \frac{n}{N} \right)^{1/2}) \right|^{12} dx_0 dx_1 dx_2 dx_3. \quad (0.3)$$

In the application, the most important range of $\delta$ is $\delta \sim \frac{1}{N^2}$. As a special case of a more general result, it was proven in [B-I2] that

$$N_8(\delta) \ll \delta N^{5+\varepsilon} + N^{4+\varepsilon} \quad (0.4)$$

and in [H-K] that

$$N_{10}(\delta, N\delta) \ll \delta N^{7+\varepsilon} + N^{5+\varepsilon}. \quad (0.5)$$

Our Fourier analytical approach gives a quite different treatment and unified approach to this problem. In particular, Theorem 9 in Section 4 below shows that in fact

$$N_{10}(\delta, N\delta) \ll N^{5+\varepsilon} \quad \text{for} \quad \delta < N^{-\frac{41}{12}}. \quad (0.6)$$

Since however the main contribution (at least in the treatment [H]) in the exponential sum problem

$$\sum_{m \sim M} e \left( TF \left( \frac{m}{M} \right) \right) \quad (0.7)$$

has $\delta = \frac{1}{N^2}$, the improvement (0.6) does not lead to new results on this matter.

Our next application are certain mean value results in the work of Robert and Sargos [R-S]. It is proven in [R-S] that

$$I_6(N^{-3}) = \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(n^2 x + N^{-3} n^4 y) \right|^6 dx dy \ll N^{3+\varepsilon} \quad (0.11)$$
\[ I_8(N^{-\frac{5}{2}}) = \int_0^1 \int_0^1 \left( \sum_{n \sim N} e(n^2x + N^{-\frac{5}{2}}n^4y) \right)^8 dx dy \ll N^{\frac{9}{2}+\varepsilon} \] (0.12)

\[ I_{10}(N^{-\frac{17}{8}}) = \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(n^2x + N^{-\frac{17}{8}}n^4y) \right|^{10} dx dy \ll N^{\frac{49}{8}+\varepsilon} \] (0.13)

Inequality (0.11) is the optimal statement for the 6th moment (a different proof using the decoupling theorem for curves appears in [B-D]). While (0.12), (0.13) are essentially sharp, they are not the optimal results for the 8th and 10th moment respectively. Since \( I_p(\lambda) \) is a decreasing function of \( \lambda \) for \( p \) an even integer, (0.13) obviously implies that

\[ I_{10} = \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(n^2x + n^4y) \right|^{10} dx dy \ll N^{\frac{42}{8}+\varepsilon}. \] (0.14)

In [R-S] an application of (0.14) to Weyl’s inequality is given, following a method initiated by Heath-Brown. In view of the present state of the art, the relevant statement is the bound

\[ |f_8(\alpha; N)| = \left| \sum_{1 \leq n \leq N} e(\alpha n^8) \right| \ll N^{1-3.2-8}(N^4q^{-1} + 1 + qN^{-4}) \frac{1}{\sqrt{q}} \] (0.15)

assuming \( |\alpha - \frac{a}{q}| \leq q^{-2}; q \geq 1, (a, q) = 1 \) (though the exponent \( \sigma(8) = 3.2^{-8} = 0.01171 \cdots \) is superseded by a recent result of Wooley, see Theorem 7.3 in [W2], which gives in particular \( \sigma(8) = \frac{1}{2 \times 8} = \frac{1}{16} = 0.01190 \cdots \)).

More recently, inequality (0.14) has been improved in [P] to

\[ I_{10} \ll N^{6+\varepsilon} \] (0.16)

using a different more arithmetical approach. As a consequence the first factor in the r.h.s. of (0.15) is replaced by \( N^{1-\frac{14}{15}2^{-8}} \), i.e. \( \sigma(8) = \frac{1}{15} = 0.0125 \cdots \).

In the final section of this paper, we establish the bounds

\[ I_8(N^{-\frac{5}{2}}) \ll N^{\frac{14}{15}+\varepsilon} \] (0.17)

\[ I_{10} \leq I_{10}(N^{-\frac{5}{2}}) \ll N^{\frac{17}{15}+\varepsilon} \] (0.18)

implying a corresponding improvement \( \sigma(8) = \frac{56}{15}2^{-8} = 0.0145 \cdots \) in Weyl’s inequality.
1. Decoupling inequality for smooth hypersurfaces with non-vanishing curvature

Let us start by recalling the main result from [B-D], which is the so-called $\ell^2$-decoupling theorem for the Fourier transform of distributions carried by hypersurfaces in $\mathbb{R}^n$ of positive curvature. This is a quite general harmonic analysis result with diverse applications, in particular to PDE’s and spectral theory (see [B-D] for some of these).

In order to formulate the result, we need some terminology. Let $S \subset \mathbb{R}^n$ be a compact smooth hypersurface of positive curvature and denote $S_\delta (\delta > 0$ a small parameter) a $\delta$-neighborhood of $S$. Decompose $S_\delta$ as a union of tangent $\sqrt{\delta} \times \cdots \times \sqrt{\delta} \times \delta$ boxes $\tau$ with bounded overlap. Denoting $B_R \subset \mathbb{R}^n$ a ball of radius $R$, the following inequality holds for functions $f$ s.t. $\text{supp} \hat{f} \subset S_\delta$

$$\|f\|_{L^p(B_{\frac{1}{2}})} \ll \delta^{-\varepsilon} \left( \sum_{\tau} \|f_{\tau}\|_{L^p(B_{\frac{1}{2}})}^2 \right)^{\frac{1}{2}} \text{ with } p = \frac{2(n+1)}{n-1} \quad (1.1)$$

and $f_{\tau} = (\hat{f}|_\tau)^\vee$ denoting the Fourier restriction of $f$ to the tile $\tau$.

By interpolation, (1.1) of course also holds for $2 \leq p \leq \frac{2(n+1)}{n-1}$ while for $\frac{2(n+1)}{n-1} \leq p \leq \infty$, the inequality becomes

$$\|f\|_{L^p(B_{\frac{1}{2}})} \ll \delta^{-\varepsilon} \left( \sum_{\tau} \|f_{\tau}\|_{L^p(B_{\frac{1}{2}})}^2 \right)^{\frac{1}{p}} \quad (1.2)$$

Next, let us relax the assumption on $S$, requiring $S$ to have non-degenerate (but not necessarily definite) second fundamental form. A statement such as (1.1) can not be valid any more. For instance, if $S \subset \mathbb{R}^3$ is a ruled surface, we may take $\text{supp} \hat{f}$ in a $\sqrt{\delta}$-neighborhood of a straight line segment with only the obvious decoupling available. This problem of curvature break-down for lower dimensional sections of $S$ can be bypassed by a suitable reformulation of the decoupling property. Assuming $S$ as above and $\text{supp} \hat{f} \subset S_\delta$, one has for $\frac{2(n+1)}{n-1} \leq p \leq \infty$

$$\|f\|_{L^p(B_{\frac{1}{2}})} \ll \delta^{-\varepsilon} \left( \sum_{\tau} \|f_{\tau}\|_{L^p(B_{\frac{1}{2}})}^p \right)^{\frac{1}{p}} \quad (1.3)$$
This statement is weaker than (1.2) but will perform equally well in what follows because in the applications below \( \text{supp} \hat{f} \) will be uniformly spread out over \( S \).

The proof of (1.3) requires a modification of the argument in [B-D] (for positive curvature). Details appear in [B-D2]. Our next goal is to derive from (1.3) a decoupling inequality for curves \( \Gamma \subset \mathbb{R}^n \) not lying in a hyperplane and which will imply our Theorem 1.

2. Construction of hypersurfaces from curves

Let \( \Gamma \subset \mathbb{R}^n \) be parametrized by \( \Phi : [0,1] \to \mathbb{R}^n : t \to (t, \varphi_1(t), \ldots, \varphi_{n-1}(t)) \) where we assume for simplicity that \( \varphi_1, \ldots, \varphi_{n-1} \) are real analytic and (importantly) that \( t, \varphi_1, \ldots, \varphi_{n-1} \) linearly independent. In particular, \( \Gamma \) does not lie in a hypersurface. Our assumption means non-vanishing of the Wronskian determinant

\[
W(\varphi_1'', \ldots, \varphi_{n-1}'') \neq 0. \tag{2.1}
\]

We build a hypersurface \( S \subset \mathbb{R}^n \) as \((n-1)\)-fold sum set

\[
S = \Gamma_1 + \cdots + \Gamma_{n-1} \tag{2.2}
\]

where \( \Gamma_j = \Phi(I_j) \) and \( I_1, \ldots, I_{n-1} \subset I \subset [0,1] \) are fixed consecutive disjoint subintervals. Hence \( S \) is parametrized by

\[
\begin{align*}
x_0 &= t_1 + \cdots + t_{n-1} \\
x_1 &= \phi_1(t_1) + \cdots + \phi_1(t_{n-1}) \\
& \quad \vdots \\
x_{n-1} &= \phi_{n-1}(t_1) + \cdots + \phi_{n-1}(t_{n-1})
\end{align*} \tag{2.3}
\]

with \( t_j \in I_j \). Our aim is to show that the second fundamental form of \( S \) is non-degenerate (but note that it may be indefinite).
Perturb \( t = (t_1, \ldots, t_{n-1}) \in I_1 \times \cdots \times I_{n-1} \) to \( (t_1 + s_1, \ldots, t_{n-1} + s_{n-1}) \), \(|s_j| = o(1)\), obtaining

\[
x_0 - t_1 - \cdots - t_{n-1} = x_0' = s_1 + \cdots + s_{n-1}
\]

\[
\begin{pmatrix}
  x_1 - \phi_1(t_1) - \cdots - \phi_1(t_{n-1}) \\
  \vdots \\
  x_{n-1} - \phi_{n-1}(t_1) - \cdots - \phi_{n-1}(t_{n-1})
\end{pmatrix} =
\begin{pmatrix}
  x_1' \\
  \vdots \\
  x_{n-1}'
\end{pmatrix} = D_1
\begin{pmatrix}
  s_1 \\
  \vdots \\
  s_{n-1}
\end{pmatrix} + \frac{1}{2} D_2 \begin{pmatrix}
  s_1^2 \\
  \vdots \\
  s_{n-1}^2
\end{pmatrix} + O(|s|^3)
\]

(2.4)

with

\[
D_1 = \begin{pmatrix}
  \phi'_1(t_1) \cdots \phi'_1(t_{n-1}) \\
  \vdots \\
  \phi'_{n-1}(t_1) \cdots \phi'_{n-1}(t_{n-1})
\end{pmatrix}
\quad \text{and} \quad
D_2 = \begin{pmatrix}
  \phi''_1(t_1) \cdots \phi''_1(t_{n-1}) \\
  \vdots \\
  \phi''_{n-1}(t_1) \cdots \phi''_{n-1}(t_{n-1})
\end{pmatrix}.
\]

(2.5)

The non-vanishing of \( \det D_1 \) can be derived from the non-vanishing of
\( W(\phi'_1, \ldots, \phi'_{n-1}) \) which is a consequence of our assumption (2.1).

Hence, since \( D_1 \) is invertible and denoting \( \xi = (1, \ldots, 1) \in \mathbb{R}^{n-1} \), the first equation in (2.4) gives

\[
x_0' = \langle D_1^{-1} \begin{pmatrix}
  x'_1 \\
  \vdots \\
  x'_{n-1}
\end{pmatrix}, \xi \rangle - \frac{1}{2} \sum_{j=1}^{n-1} s_j^2 \langle D_1^{-1} D_2 e_j, \xi \rangle + O(|s|^3)
\]

\[
= \langle \begin{pmatrix}
  x''_1 \\
  \vdots \\
  x''_{n-1}
\end{pmatrix}, \xi \rangle - \frac{1}{2} \sum_{j=1}^{n-1} (x''_j)^2 \langle D_1^{-1} D_2 e_j, \xi \rangle + O(|x''|^3)
\]

(2.6)

where

\[
\begin{pmatrix}
  x''_1 \\
  \vdots \\
  x''_{n-1}
\end{pmatrix} = D_1^{-1} \begin{pmatrix}
  x'_1 \\
  \vdots \\
  x'_{n-1}
\end{pmatrix}.
\]

(2.6)

From (2.6), it remains to ensure that

\[
\langle D_1^{-1} D_2 e_j, \xi \rangle \neq 0 \text{ for each } j = 1, \ldots, n-1.
\]

(2.7)

Take \( j = 1 \). Up to a multiplicative factor,
\begin{align*}
\langle D_2 e_1, (D_1^{-1})^* \xi \rangle &= \sum_{k=1}^{n-1} \phi_k''(t_1) \\
&= \sum_{k=1}^{n-1} \phi_k''(t_1) \cdot \phi_k'(t_{n-1})
\end{align*}

By the mean-value theorem, we obtain separated $t_1 < t'_2 < \cdots < t'_{n-1}$ such that

\begin{align*}
&= \sum_{k=1}^{n-1} (-1)^k \phi_k''(t_1) \cdot \phi_k'(t_{n-1}) \\
&= \phi_k''(t_1) \cdot \phi_k'(t_{n-1}) \cdot \phi_k'(t_1) \cdot \phi_k''(t_{n-1})
\end{align*}

and the non-vanishing can again be ensured by (2.1).

3. **Decoupling inequality for curves**

Next, we use (1.3) to derive a decoupling inequality for curves (a variant of this approach appears in [B-D2]).

Let $\Gamma_1, \ldots, \Gamma_{n-1} \subset \Gamma \subset \mathbb{R}^n$ be as in §2. Let $\delta > 0$ and denote by $\Gamma_j^\delta$ a $\delta$-neighborhood of $\Gamma_j$.

Assume $\text{supp } \hat{f}_j \subset \Gamma_j^\delta$.

Write with $x = (x_0, x_1, \ldots, x_{n-1}) \in \mathbb{R}^n$ and $\Phi$ as above

\begin{align*}
\prod_{j=1}^{n-1} \left[ \int_{I_j} \hat{f}_j(t_j) e(\langle \Phi(t_j), x \rangle) dt_j \right] &= \\
\int_{I_1} \cdots \int_{I_{n-1}} \left[ \prod_{j=1}^{n-1} \hat{f}_j(t_j) \right] e(\langle \Phi(t_1) + \cdots + \Phi(t_{n-1}), x \rangle) dt_1 \cdots dt_{n-1} &= \\
\int_{S} \left[ \prod_{j=1}^{n-1} \hat{f}_j(t_j) \right] e(\xi, x) \Omega(\xi) d\xi
\tag{3.1}
\end{align*}

with $\Omega$ some smooth density on $S$. 
Let \( p = \frac{2(n+1)}{n-1} \) and apply the decoupling inequality for \( S \) stated in (1.3) of Section 1. Observe that by the regularity of \( D_1 \) in (2.5), a partition of \( S \) in \( \sqrt{d} \)-caps \( \tau_\alpha \subset S \) is equivalent to a partition of \( I_1 \times \cdots \times I_{n-1} \) in \( \sqrt{d} \)-cubes. Hence, denoting by \( J \subset [0,1] \) \( \sqrt{d} \)-intervals, we obtain

\[
\| (3.1) \|_{L^p(B_1)} \ll \delta^{-\frac{n}{2(n+1)}} \left\{ \sum_{J_1, \ldots, J_{n-1}} \left\| \prod_{j=1}^{n-1} \left[ \int_{I_j} \hat{f}_j(t_j) e(t_j x_0 + f_1(t_j) x_1 + \cdots + f_{n-1}(t_j) x_{n-1}) \, dt_j \right] \right\|_{L^p(B_1)} \right\}^{1/p}.
\]

Next take \( N = \frac{1}{\delta} \) and discretize inequality (3.2) by setting \( t = \frac{k}{N}, k \in \{ \frac{N}{2}, \ldots, N \} \).

This leads to the following inequality for separated intervals \( U_1, \ldots, U_{n-1} \subset \{ \frac{N}{2}, \ldots, N \} \)

\[
\left\| \prod_{j=1}^{n-1} \left[ \sum_{k \in U_j} a_k e(k x_0 + N \varphi_1 \left( \frac{k}{N} \right) x_1 + \cdots + N \varphi_{n-1} \left( \frac{k}{N} \right) x_{n-1}) \right] \right\|_{L^p([0,1]^n)} \ll \]

\[
N^{-\frac{n}{2(n+1)} + \varepsilon} \left( \sum_{V_1, \ldots, V_{n-1}} \left\| \prod_{j=1}^{n-1} \left[ \sum_{k \in V_j} a_k e(\cdots) \right] \right\|_{L^p([0,1]^n)} \right)^{\frac{1}{p}}.
\]

with \( V \subset \{ \frac{N}{2}, \ldots, N \} \) running in a partition in \( \sqrt{N} \)-size intervals.

Note that the domain \([0,1]^n\) may always be replaced by a larger box \( \prod_{j=0}^{n-1} [0, K_j], K_j \geq 1 \). In particular, the function

\[ k x_0 + N \varphi_1 \left( \frac{k}{N} \right) x_1 + \cdots + N \varphi_{n-1} \left( \frac{k}{N} \right) x_{n-1} \]

in (3.3) may be replaced by

\[ k x_0 + N_1 \varphi_1 \left( \frac{k}{N} \right) x_1 + \cdots + N_{j-1} \varphi_{n-1} \left( \frac{k}{N} \right) x_{n-1} \]

where \( N_1, \ldots, N_{j-1} \geq N \).

Take \( \varphi_1(t) = t^2, N_1 = N^2, N_2 = \cdots = N_{j-1} = N \). We obtain

\[
\left\| \prod_{j=1}^{n-1} \left[ \sum_{k \in U_j} e(k x_0 + k^2 x_1 + N \varphi_2 \left( \frac{k}{N} \right) x_2 + \cdots + N \varphi_{n-1} \left( \frac{k}{N} \right) x_{n-1}) \right] \right\|_{L^p([0,1]^n)} \ll \]

\[
N^{-\frac{n}{2(n+1)} + \varepsilon} \left( \sum_{V_1, \ldots, V_{n-1}} \left\| \prod_{j=1}^{n-1} \left[ \sum_{k \in V_j} e(\cdots) \right] \right\|_{L^p([0,1]^n)} \right)^{\frac{1}{p}}.
\]
Our next task is to bound the individual summands in (3.4).

Write \( \bar{k} = (k_1, \ldots, k_{n-1}) \in V_1 \times \cdots \times V_{n-1} \) as \( \bar{k} = \bar{\ell} + \bar{m}, \ell_j \) the center of \( V_j \) and \( |m_j| < \sqrt{N}. \) Hence

\[
\sum_{j=1}^{n-1} \left( k_j x_0 + k_j^2 x_1 + N \phi_2 \left( \frac{k_j}{N} \right) x_2 + \cdots + N \phi_{n-1} \left( \frac{k_j}{N} \right) x_{n-1} \right) = \\
m_1 \left( x_0 + 2 \ell_1 x_1 + \phi_2' \left( \frac{\ell_1}{N} \right) x_2 + \cdots + \phi_{n-1}' \left( \frac{\ell_1}{N} \right) x_{n-1} \right) + \\
\vdots \\
+ m_{n-1} \left( x_0 + 2 \ell_{n-1} x_1 + \phi_2' \left( \frac{\ell_{n-1}}{N} \right) x_2 + \cdots + \phi_{n-1}' \left( \frac{\ell_{n-1}}{N} \right) x_{n-1} \right) + \\
(m_1^2 + \cdots + m_{n-1}^2) x_1 + \psi(\bar{m}, x) \\
\tag{3.5}
\]

where \( |\psi(\bar{m}, x)| < o(1) \) and \( |\partial_m \psi(\bar{m}, x)| < O(N^{-\frac{1}{2}}) \) since \( |\bar{m}| < \sqrt{N} \) and \( |x| < 1. \)

Thus \( \psi(\bar{m}, x) \) may be dismissed in (3.4) when evaluating

\[
\| \prod_{j=1}^{n-1} |\sum_{k \in V_j} e(\cdots)| \|_{L^p([0,1]^n)}. 
\]

Make an affine change of variables

\[
\begin{pmatrix}
  y_1 \\
  \vdots \\
  y_{n-1}
\end{pmatrix} = A \begin{pmatrix}
  x_0 \\
  x_2 \\
  \vdots \\
  x_{n-1}
\end{pmatrix}
\]

with

\[
A = \begin{bmatrix}
  1 & \phi_2' \left( \frac{\ell_1}{N} \right) & \cdots & \phi_{n-1}' \left( \frac{\ell_1}{N} \right) \\
  \vdots & \vdots & \ddots & \vdots \\
  1 & \phi_2' \left( \frac{\ell_{n-1}}{N} \right) & \cdots & \phi_{n-1}' \left( \frac{\ell_{n-1}}{N} \right)
\end{bmatrix}
\]

in the \( (x_0, x_2, \ldots, x_{n-1}) \) variables, noting that this linear coordinate change can be assumed regular provided \( W(\phi_2'', \ldots, \phi_{n-1}'') \neq 0 \) (which is implied by (2.1) for \( \phi_1(t) = t^2 \)).

Next, using periodicity, another coordinate shift leads to

\[
\begin{align*}
\| \prod_{j=1}^{n-1} & \sum_{k \in V_j} e(\cdots) \|_{L^p([0,1]^n)} \sim \\
\| \sum_{m_1, \ldots, m_{n-1} < \sqrt{N}} & e \left( m_1 y_1 + \cdots + m_{n-1} y_{n-1} + (m_1^2 + \cdots + m_{n-1}^2) x_1 \right) \|_{L^p_{y_1, y_2, \ldots, y_{n-1}}([0,1]^n)} \\
& \ll N^{\frac{1}{2} + \varepsilon} \\
\end{align*}
\]  

\tag{3.6}
by the Strichartz inequality on $\mathbb{T}^n$.

Summarizing, we proved the following multi-linear mean value theorem.

**Theorem 1.** Assume $n \geq 3$ and $\varphi_2, \ldots, \varphi_{n-1}$ satisfying

$$W(\varphi_2, \ldots, \varphi_{n-1}) \neq 0.$$ 

Let $U_1, \ldots, U_{n-1} \subset \left[\frac{N}{2}, N\right] \cap \mathbb{Z}$ be $O(N)$-separated intervals. Then

$$\left\| \prod_{j=1}^{n-1} \sum_{k \in U_j} e\left(k x_0 + k^2 x_1 + N \varphi_2\left(\frac{k}{N}\right) x_2 + \cdots + N \varphi_{n-1}\left(\frac{k}{N}\right) x_{n-1}\right) \right\|_{L^{2(n+1)}/(\mathbb{R}^n)} \leq N^{n+1/2} + \varepsilon. \quad (3.7)$$

**Remarks.**

(i) Theorem 1 remains valid (following the same argument) with coefficients $a_k$, $k \in U_j$ and r.h.s. replaced by $\prod_{j=1}^{n-1} (\sum_{k \in U_j} |a_k|^2)^{1/2}$.

(ii) Note that (3.7) is best possible. Indeed, restricting $|x_0| < \frac{1}{N}$, $|x_1| < \frac{1}{N^2}$, $|x_2| < \frac{1}{N^3}$, ..., $|x_{n-1}| < \frac{1}{N}$, one gets the contribution

$$N^{n-1 - \frac{n-1}{2(n+1)}(n+1)} = N^{\frac{n-1}{2}}.$$

(iii) Also, as we will see shortly, (3.7) is only valid in the above multi-linear form.

4. **Mean values estimates for the 8th and 10th moment**

Note that (3.7) is the multi-linear version of an estimate on

$$\left\| \sum_{k \sim N} e\left(k x_0 + k^2 x_1 + N \varphi_2\left(\frac{k}{N}\right) x_2 + \cdots + N \varphi_{n-1}\left(\frac{k}{N}\right) x_{n-1}\right) \right\|_{L^{2(n+1)}([0,1]^n)}.$$ 

(4.1)

Denoting $f_I = \sum_{k \in I} e\left(k x_0 + k^2 x_1 + N \varphi_2\left(\frac{k}{N}\right) x_2 + \cdots + N \varphi_{n-1}\left(\frac{k}{N}\right) x_{n-1}\right)$ for $I \subset [1, N]$ a subinterval, one adopts the following argument from [B-G].

Partition $[1, N]$ in intervals $I$ of size $N^{1-\tau}$ ($\tau > 0$ small). Fix a point $x$ and distinguish the following two scenarios. Either we can find $n - 1$ intervals $I_1, \ldots, I_{n-1}$ that are $O(N^{1-\tau})$-separated and such that

$$|f_{I_j}(x)| > N^{-2\tau} |f(x)| \quad \text{for} \quad 1 \leq j \leq n - 1 \quad (4.2)$$

or for some interval $I$, we have

$$|f_I(x)| > c |f(x)|. \quad (4.3)$$
The contribution of (4.2) is captured by the multi-linear estimate (3.7) and we obtain $N^{1+c}\tau$. For the (4.3)-contribution, bound by

$$\max |f_I| \leq \left[ \sum_I |f_I|^{2(n+1)} \right]^{\frac{1}{2(n+1)}}$$

contributing to

$$\left[ \sum_I \|f_I\|_{L^2(2(n+1))}^{2(n+1)} \right]^{\frac{1}{2(n+1)}}.$$

(4.4)

One may then repeat the process to each $f_I$. Note that after a coordinate change in $x_0, x_1$, we obtain exponential sums of the form

$$F(x) = \sum_{\ell \sim M} e\left( \ell x_0 + \ell^2 x_1 + N\varphi_2\left( \frac{k}{N} + \frac{\ell}{N} \right)x_2 + \cdots + N\varphi_{n-1}\left( \frac{k}{N} + \frac{\ell}{N} \right)x_{n-1} \right)$$

with $M = N^{1-\tau}$, $k \sim N$ fixed. Set for $j = 1, \ldots, n - 1$.

$$N\varphi_j\left( \frac{k}{N} + \frac{\ell}{N} \right) = M\psi_j\left( \frac{\ell}{M} \right)$$

(4.6)

with

$$\psi_j(t) = \frac{N}{M}\varphi\left( \frac{k}{N} + \frac{M}{N}t \right).$$

(4.7)

However the Wronskian condition $W(\psi''_2, \ldots, \psi''_{n-1}) > O(1)$ deteriorated. For $n = 3$ we will nevertheless be able to retrieve easily the expected bound, while for $n \geq 4$, the linear bounds turn out to be weaker than the multi-linear one.

Let $n = 3$. Then $\psi_2''(t) = \frac{M^2}{N^2}\varphi''\left( \frac{k}{N} + \frac{M}{N}t \right) = O\left( \frac{M^2}{N^2} \right)$ and replacing $\psi_2 = \frac{M^2}{N^2}\varphi_2, x'_2 = \frac{M^2}{N^2}x_2$, this leads to

$$\|F\|_{L^8_{x_0, x_1, x'_2 = O(1)}} \sim \left( \frac{N}{M} \right)^{\frac{1}{2}} \left\| \sum_{\ell \sim M} e\left( \ell x_0 + \ell^2 x_1 + M\varphi_2\left( \frac{\ell}{M} \right)x'_2 \right) \right\|_{L^8_{x_0, x_1, x'_2 = O(1)}}$$

$$\leq \left( \frac{N}{M} \right)^{\frac{1}{2}} \cdots \left\| \right\|_{L^8_{x_0, x_1, x'_2 = O(1)}}$$

$$< \left( \frac{N}{M} \right)^{\frac{1}{2}} M^{\frac{3}{4} + \varepsilon}$$

assuming the expected bound at scale $M$. The bound on (4.4) becomes then

$$\left( \frac{N}{M} \right)^{\frac{3}{8}} M^{\frac{3}{4} + \varepsilon} = \left( \frac{M}{N} \right)^{\frac{1}{8}} N^{\frac{3}{4} + \varepsilon}$$

and summing over dyadic $M < N$ we reproved the main result from [3-12].
Theorem 2. [B-I2].

Assume \( \varphi''' \neq 0 \). Then

\[
\left\| \sum_{k \sim N} e\left( kx_0 + k^2x_1 + N\varphi\left( \frac{k}{N} \right)x_2 \right) \right\|_8 \ll N^{\frac{1}{2} + \varepsilon}.
\] (4.8)

Note that in their application to \( \zeta(\frac{1}{2} + it) \), \( \varphi(t) = t^{3/2} \).

It is interesting to note that unlike in [B-I2], our derivation of (4.8) did not make use of Poisson summation (i.e. Process B).

The work of [B-I1] was extensively refined by Huxley and his collaborators, resulting in his book [H].

The present discussion is relevant to the so called ‘First Spacing Problem’; (4.8) indeed means that the system

\[
\begin{cases}
    k_1 + k_2 + k_3 + k_4 = k_5 + \cdots + k_8 \\
    k_1^2 + \cdots + k_4^2 = k_5^2 + \cdots + k_8^2 \\
    k_1^{3/2} + \cdots + k_4^{3/2} = k_5^{3/2} + \cdots + k_8^{3/2} + O(\sqrt{N}).
\end{cases}
\] (4.9)

has at most \( N^{4 + \varepsilon} \) solutions in integers \( k_1, \ldots, k_8 \sim N \) (the statement is clearly optimal).

Huxley considers the more elaborate problem in 10-variables

\[
\begin{cases}
    k_1 + k_5 = k_6 + \cdots + k_{10} \\
    k_1^2 + \cdots + k_5^2 = k_6^2 + \cdots + k_{10}^2 \\
    k_1^{3/2} + \cdots + k_5^{3/2} = k_6^{3/2} + \cdots + k_{10}^{3/2} + O(\delta N^{3/2}) \\
    k_1^{1/2} + \cdots + k_5^{1/2} = k_6^{1/2} + \cdots + k_{10}^{1/2} + O(\Delta N^{1/2})
\end{cases}
\] (4.10)

(see [H], §11) for which the number \( N_{10}(\delta, \Delta) \) of solutions is given by the 10th moment

\[
\left\| \sum_{k \sim N} e\left( kx_0 + k^2x_1 + \frac{1}{5} \left( \frac{k}{N} \right)^{3/2} x_2 + \frac{1}{\Delta} \left( \frac{k}{N} \right)^{1/2} x_3 \right) \right\|_{L^{10}_{x_0, x_1, x_2, x_3}}. \] (4.11)

In the applications to exponential sums, \( \Delta = \delta N, \frac{1}{N^{1/2}} < \delta < \frac{1}{N} \). In this setting, the following key inequality appears in [H-K].

Theorem 3. [H-K]. With \( \Delta = \delta N, \frac{1}{N^{1/2}} < \delta < \frac{1}{N} \), we have

\[ N_{10}(\delta, \delta N) \ll \delta N^{7+\varepsilon}. \] (4.12)
In what follows, we estimate (4.11) using Theorem 7 and will in particular retrieve (4.12) in a stronger form.

Start by observing that, as a consequence of (3.7), for $U_1, U_2, U_3$ and $\varphi_2, \varphi_3$ as in Theorem 1
\[
\int_0^1 \int_0^1 \int_0^1 \int_0^1 \left\{ \prod_{j=1}^3 \sum_{k \in U_j} e\left(kx_0 + k^2x_1 + \frac{1}{\delta} \varphi_2\left(\frac{k}{N}\right)x_2 + \frac{1}{\Delta} \varphi_3\left(\frac{k}{N}\right)x_3\right) \right\} dx_0 dx_1 dx_2 dx_3 \ll

\left[ \min(\delta N, N) + 1 \right] \left[ \min(\Delta N, N) + 1 \right] N^{5+\varepsilon} \tag{4.13}
\]

Using the scale reduction described in (4.1)-(4.7), we also need to evaluate the contributions of
\[
M \cdot (4.14) \tag{4.15}
\]
with
\[
(4.14) = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left\{ \prod_{j=1}^3 \sum_{\ell \in U_j} e\left(\ell x_0 + \ell^2 x_1 + \frac{1}{\delta} \varphi_2\left(\frac{k+\ell}{N}\right)x_2 + \frac{1}{\Delta} \varphi_3\left(\frac{k+\ell}{N}\right)x_3\right) \right\}
\]
where $k \in \left[ \frac{N}{2}, N \right]$, $I = [k, k+M]$ and $U_1, U_2, U_3$ are $\sim M$ separated subintervals of size $\sim M$ in $I$. By a change of variables in $x$, the phase function in (4.16) may be replaced by
\[
e\left(\ell x_0 + \ell^2 x_1 + \frac{M^3}{\delta N^3} \tilde{\varphi}_2\left(\frac{\ell}{M}\right)x_2 + \frac{M^4}{\Delta N^4} \tilde{\varphi}_3\left(\frac{\ell}{M}\right)x_3\right) \tag{4.16}
\]
where $\tilde{\varphi}_2(t)$ has leading monomial $t^3$ and $\tilde{\varphi}_3(t)$ leading monomial $t^4$. Hence $W(\tilde{\varphi}''_2, \tilde{\varphi}''_3) > c$ and (4.13) is applicable to (4.14) with $N, \delta, \Delta$ replaced by $M, \frac{\delta N^3}{M^2}, \frac{\Delta N^4}{M^4}$. Therefore
\[
(4.15) \ll \left[ 1 + \min \left( \frac{\delta N^3}{M^2}, M \right) \right] \left[ 1 + \min \left( \frac{\Delta N^4}{M^3}, M \right) \right] M^4 N^{1+\varepsilon} \tag{4.17}
\]
and (4.17) needs to be summed over dyadic $M < N$. One easily checks that the conclusion is as follows

**Theorem 4.** Assume $W(\varphi''_2, \varphi''_3) \neq 0$ and $\delta < \Delta$. Then
\[
\left\| \sum_{k \sim N} e\left(kx_0 + k^2x_1 + \frac{1}{\delta} \varphi_2\left(\frac{k}{N}\right)x_2 + \frac{1}{\Delta} \varphi_3\left(\frac{k}{N}\right)x_3\right) \right\|_{10}^{10} \ll

\left[ \delta \Delta^{3/4} N^7 + (\delta + \Delta) N^6 + N^5 \right] N^\varepsilon. \tag{4.18}
\]
In particular
\[
N_{10}(\delta, \Delta) < (4.18).
\]
Hence, we are retrieving Theorem 3.

**Remark.** We make the following comment on the role of the first term in the r.h.s. of (4.18), relevant to the Remark following Theorem 1.

Partition \([\frac{N}{2}^+, N]\) in intervals \(I = [n, n+M]\) of size \(M\). Obviously \(N_{10}(\delta, \Delta)\) is at least \(\frac{N}{M}\) times a lower bound on the number of solutions of

\[
\begin{align*}
\left\{ 
\begin{array}{l}
m_1 + \ldots + m5 = m6 + \ldots + m10 \\
m1^2 + \ldots + m5^2 = m6^2 + \ldots + m10^2 \\
\left( \frac{n+m1}{N} \right)^2 + \ldots + \left( \frac{n+m5}{N} \right)^2 = \left( \frac{n+m6}{N} \right)^2 + \ldots + \left( \frac{n+m10}{N} \right)^2 + O(\delta) \\
\left( \frac{n+m1}{N} \right)^2 + \ldots + \left( \frac{n+m5}{N} \right)^2 = \left( \frac{n+m6}{N} \right)^2 + \ldots + \left( \frac{n+m10}{N} \right)^2 + O(\Delta)
\end{array}
\right. \\
\end{align*}
\]

(4.19) \quad (4.20)

Since

\[
\left( \frac{n+m}{N} \right)^2 = \left( \frac{n}{N} \right)^2 + \frac{3}{2} \left( \frac{n}{N} \right) \frac{m}{N} + \frac{3}{2} \left( \frac{n}{N} \right) - \frac{7}{16} \left( \frac{n}{N} \right)^2 + \frac{3}{16} \left( \frac{n}{N} \right) - \frac{3}{4} \left( \frac{n}{N} \right)^3 + \ldots
\]

\[
\left( \frac{n+m}{N} \right)^2 = \left( \frac{n}{N} \right)^2 - \frac{1}{2} \left( \frac{n}{N} \right) \frac{m}{N} - \frac{1}{8} \left( \frac{n}{N} \right)^2 + \frac{3}{16} \left( \frac{n}{N} \right) - \frac{7}{4} \left( \frac{n}{N} \right)^3 - \frac{15}{128} \left( \frac{n}{N} \right)^4 + \ldots
\]

the equations (4.20) may be replaced by

\[
\begin{align*}
\left\{ \begin{array}{l}
\varphi \left( \frac{m}{N} \right) + \ldots - \varphi \left( \frac{m10}{N} \right) = O(\delta) \\
\psi \left( \frac{m}{N} \right) + \ldots - \psi \left( \frac{m10}{N} \right) = O(\Delta)
\end{array} \right. \\
\end{align*}
\]

with \(\varphi, \psi\) of the form \(\varphi(t) = a3t^3 + a4t^4 + \ldots\) and \(\psi(t) = b3t^3 + b4t^4 + \ldots\)

and where \(\begin{vmatrix} a3 & b3 \\ a4 & b4 \end{vmatrix} \neq 0\).

Assume \(\delta < \Delta\) and replace \(\psi\) by \(\psi_1 = \psi - \frac{b3}{a3} \varphi = c4t^4 + \ldots\)

Writing \(\frac{m}{N} = \frac{M}{N} \frac{n}{M}\), we obtain conditions of the form

\[
\begin{align*}
\begin{array}{l}
\tilde{\varphi} \left( \frac{m}{M^7} \right) + \ldots - \tilde{\varphi} \left( \frac{m10}{M^7} \right) = O \left( \frac{N^3}{M^7} \delta \right) \\
\tilde{\psi} \left( \frac{m}{M^7} \right) + \ldots - \tilde{\psi} \left( \frac{m10}{M^7} \right) = O \left( \frac{N^4}{M^4} \Delta \right)
\end{array}
\end{align*}
\]

(4.21)

where \(\tilde{\varphi} = t^3 + \ldots, \tilde{\psi}_1 = t^4 + \ldots\). Consider the system (4.19)+(4.21) with \(m_i \leq M\). Clearly the number of solutions is at least

\[
M^7 \min \left( 1, \frac{N^3}{M^3} \delta \right) \min \left( 1, \frac{N^4}{M^4} \Delta \right).
\]

Taking \(M = \Delta^{1/4}N\), we obtain \(N^7(\delta \Delta)\). The quantity is multiplied further with \(\frac{N}{M}\), leading to a lower bound \(\delta \Delta^{3/4} N^7\) for \(N_{10}(\delta, \Delta)\).
This shows that the first term in (4.18) (apart from the $N^\varepsilon$ factor) is also a lower bound.

In our applications, $\Delta$ tends to be much larger then $\delta$ which makes $\Delta N$ the leading term in (4.18). Next, we develop an argument to reduce the weight of $\Delta N$ by involving also some ideas and techniques from [II]. It is likely that our presentation can be improved at this point.

We will need the following variant of van der Corput’s exponential sum bound (cf. [Ko], Theorem 2.6).

Lemma 5. Assume $f$ a smooth function on $I = [\frac{N}{2}, N]$ and $f^{(3)} \sim \lambda_3$. Let \{V_j\} denote a partition of $I$ in intervals of size $D$. Then

$$\sum_j \left| \sum_{n \in V_j} e(f(n)) \right|^2 \lesssim \begin{cases} N + D^\frac{1}{2}\lambda_3^\frac{1}{2} + D^\frac{2}{3}\lambda_3^\frac{2}{3}N & \text{if } D > \lambda_3^{-\frac{2}{3}} \quad (4.22) \\ ND^\frac{1}{2} + D\lambda_3^{-\frac{2}{3}} & \text{if } D \leq \lambda_3^{-\frac{2}{3}} \quad (4.23) \end{cases}$$

We first proceed with a multi-linear bound considering instead of (4.13) 5-linear expressions with $U_j \subset [\frac{N}{2}, N]$ of size $\sim N$ and $\sim N$ separated $(1 \leq j \leq 5)$

$$\int \left\{ \prod_{j=1}^5 \left| \sum_{k \in U_j} e\left(kx_0 + k^2x_1 + \frac{1}{\delta}(\frac{k}{N})^{3/2}x_2 + \frac{1}{\Delta}(\frac{k}{N})^{\frac{1}{2}}x_3\right) \right|^2 \right\} dx_0dx_1dx_2dx_3. \quad (4.24)$$

This quantity will increase by increasing $\delta$ and we replace $\delta$ by a parameter $\delta_1 > \delta$ to be specified. An application of Hölder’s inequality permits then to bound (4.24) by (4.13) with $\delta$ replaced by $\delta_1$.

Assuming $\Delta < 1$, perform a decoupling at scale $N\Delta^\frac{1}{2}$ using (3.2). This gives an estimate on the l.h.s. of (4.13) by

$$\Delta^{-1} \int \prod_{j=1}^3 \left[ \sum_{V_j \subset U_j} \left| \sum_{k \in V_j} e(\cdots) \right|^{\frac{10}{3}} \right] dx \quad (4.25)$$

with $V_j \subset U_j$ a partition in $N\Delta^\frac{1}{2}$-intervals. Using again Hölder’s inequality, one may bound

$$\prod_{j=1}^3 \left[ \sum_{V_j \subset U_j} \left| \frac{10}{3} \right| \right] \leq \left( \sum_{V \subset [\frac{N}{2}, N]} \left| \frac{2}{3} \right| \right) \prod_{j=1}^2 \left( \sum_{V_j \subset U_j} \left| \frac{4}{3} \right| \right) + \cdots \quad (4.26)$$

where $\cdots$ refers to the pairs $U_2, U_3$ and $U_3, U_1$ instead of $U_1, U_2$. 
Specifying in (4.13), with $\delta$ replaced by $\delta_1$, a range

$$x_2 \sim X_2 < 1 \text{ assuming } X_2 \Delta > 100 \delta_1$$

(4.27)

an application of (4.22) to the first factor of (4.26) with $D = \Delta^{\frac{1}{2}} N, \lambda_3 \sim \frac{X_2}{\delta_1 N^2}$ gives the bound

$$N + \delta_1^\frac{1}{2} \Delta^{\frac{1}{2}} N^2 X_2^{-\frac{1}{2}} + \delta_1^{-\frac{1}{2}} \Delta^{\frac{3}{2}} N X_2^\frac{1}{2}.$$

(4.28)

We always assume

$$\Delta N > 100$$

(4.29)

(this condition remains clearly preserved at lower scales, cf. (4.17)).

Apply the bilinear estimate (Theorem 1 with $n = 3$) to the second factor of (4.26) considering the variables $x_0, x_1, x_2$ and restricting $x_2 \sim X_2$. By (4.24), this gives the contribution

$$(1 + \delta_1^{\frac{1}{2}} \Delta^{\frac{1}{2}} N X_2^{-\frac{1}{2}} + \delta_1^{-\frac{1}{2}} \Delta^{\frac{3}{2}} X_2^{\frac{1}{2}}) (X_2 + \delta_1 N)^{5+\epsilon}.$$

Assuming $X_2 > \delta_1 N$, which by (4.29) implies (4.27), gives the bound

$$(1 + \delta_1^{\frac{1}{2}} \Delta^{\frac{1}{2}} N + \delta_1^{-\frac{1}{2}} \Delta^{\frac{3}{2}}) N^{5+\epsilon}.$$

The contribution of $X_2 < \delta_1 N$ is estimated by

$$\Delta^{-1} \int_0^1 \int_0^1 \int_0^{\delta_1 N} \int_0^1 \left\{ \prod_{j=1}^3 \left[ \sum_{V_j \subset U_j} \right] \right\} dx_0 dx_1 dx_2 dx_3 \ll \Delta^{-1} \left( \frac{N}{\Delta^{1/2} N} \right)^3 (\Delta N) \delta_1 N (\Delta^{\frac{1}{2}} N)^5 \ll \Delta \delta_1 N^{7+\epsilon}$$

assuming $\delta_1 N < 1$. This gives

$$(1 + \delta_1^{\frac{1}{2}} \Delta^{\frac{1}{2}} + \delta_1^{\frac{3}{2}} \Delta^{\frac{3}{2}} N + \Delta \delta_1 N^2) N^{5+\epsilon}$$

(4.30)

and setting $\delta_1 = \Delta^{\frac{1}{2}} N^{-1}$, assuming $\delta_1 > \delta$ gives

$$((\Delta N)^\frac{1}{2} + \Delta^{3/2} N) N^{5+\epsilon}.$$

If $\delta_1 \leq \delta$, use (4.30) with $\delta_1 = \delta$.

Thus the multi-linear contribution in the 10th moment may be estimated by

$$\left( \delta^{\frac{1}{2}} \Delta^{\frac{1}{2}} N + \Delta \delta N^2 + (\Delta N)^{\frac{1}{2}} + \Delta^{3/2} N \right) N^{5+\epsilon}.$$

(4.31)
Next, consider the lower scale contributions

\[ \sum_{I \subseteq \left[ \frac{N}{2}, N \right]} \left\| \sum_{n \in I} e(\cdots) \right\|_{10}^{10}. \]  

(4.32)

Fixing \( M < N \) and replacing \( \delta \), resp. \( \Delta \), by \( \frac{N^3}{M^2} \delta \), resp. \( \frac{N^4}{M^2} \Delta \), we obtain the bound

\[ (\delta^{\frac{1}{2}} \Delta^{\frac{1}{4}} N + (\Delta N)^{\frac{1}{2}}) N^{5+\varepsilon} + (\Delta \Delta N + \Delta^{3/2}) \frac{N^{7+\varepsilon}}{M} \]  

for the multi-linear contribution at scale \( M \).

On the other hand, we can also make a crude estimate using the \( L^8 \)-norm, leading to the contribution

\[ \frac{N}{M} M^2 \left( 1 + \delta \frac{N^3}{M^2} \right) M^{4+\varepsilon} \ll N M^{5+\varepsilon} + \delta N^4 M^{3+\varepsilon} \]  

(4.34)

and \( (\Delta N)^{\frac{1}{2}} N^{5+\varepsilon} \) for \( M < (\Delta N)^{\frac{1}{10}} N^{4/5} \) and \( \delta < N^{-7/5}(\Delta N)^{\frac{1}{5}} \).

Hence we get

**Lemma 6.** For \( N \Delta > 1 \) and \( \delta < N^{-7/5}(\Delta N)^{\frac{1}{5}} \)

\[ N_{10}(\delta, \Delta) \ll (\delta^{\frac{1}{2}} \Delta^{\frac{1}{4}} N + (\Delta N)^{\frac{1}{2}}) N^{5+\varepsilon} + (\Delta \Delta N + \Delta^{3/2}) (\Delta N)^{\frac{1}{10}} N^{\frac{11}{5}+\varepsilon}. \]  

(4.35)

Next, recall Lemma 11, 3.3 in [H].

**Lemma 7.** Assume \( \frac{1}{N} > \delta > \frac{1}{N^2} \) and \( \frac{1}{N} < \Delta < \delta N \). Letting

\[ 2 \leq T \leq \frac{1}{\sqrt{\delta N}} \]

be a parameter, the following inequality holds

\[ N_{10}(\delta, \Delta) \lesssim \frac{1}{T} N_{10}(T^2 \delta, T \Delta) + N_{10}(\delta, CT \delta). \]  

(4.36)

Combining Theorem 4, Lemmas 6 and 7 (applied with \( T = \Delta N \)) gives

**Lemma 8.** Assume \( \frac{1}{N} > \delta > \frac{1}{N^2}, \frac{1}{N} < \Delta < \delta N \) and \( \Delta \sqrt{\delta N}^{3/2} < 1 \). Then

\[ N_{10}(\delta, \Delta) < (1 + \delta (\Delta N)^{\frac{3}{2}} N^{\frac{5}{2}}) \left( 1 + \delta^{\frac{1}{2}} (\Delta N)^{\frac{3}{2}} N^{\frac{5}{2}} \right) N^{5+\varepsilon} + (\delta (\Delta N)^{\frac{14}{5}} + (\Delta N)^{\frac{14}{5}} + \Delta^{2}) \frac{N^{\frac{11}{5}+\varepsilon}}{M}. \]  

(4.37)

Setting \( \Delta = \delta N \) leads to the following strengthening of Theorem 3

**Theorem 9.** For \( N^{-\frac{11}{5}} \geq \delta \geq N^{-2} \), we have \( N_{10}(\delta, \delta N) \ll N^{5+\varepsilon} \).
Note that in view of the Remark following Theorem 4, the conclusion of Theorem 9 fails for $\delta > N^{-11/7}$.

5. ON AN INEQUALITY OF ROBERT AND SARGOS

In [R-S] established the inequity

$$I_{10} = \int_0^1 \int_0^1 \left( \sum_{n \sim N} e(n^2 x + n^4 y) \right)^{10} dxdy \ll N^{49/8 + \varepsilon}$$

(5.1)

which they applied to obtain new estimates on Weyl sums. An improvement of (5.1) appears in [P], who obtains

$$I_{10} \ll N^{6+\varepsilon}.$$  

(5.2)

Using our methods, we present a further improvement.

**Theorem 10.**

$$I_{10} \ll N^{17/3 + \varepsilon}$$  

(5.3)

The corresponding improvement in Weyl’s inequality following Heath-Brown’s method was recorded in the Introduction.

Note that bounding $I_{10}$ is tantamount to estimating the number of integral solutions $n_i \sim N$ ($1 \leq 1 \leq 10$) of the system

$$\begin{cases}
n_1^2 + n_2^2 + n_3^2 + n_4^2 + n_5^2 = n_6^2 + n_7^2 + n_8^2 + n_9^2 + n_{10}^2 \\
n_1^4 + n_2^4 + n_3^4 + n_4^4 + n_5^4 = n_6^4 + n_7^4 + n_8^4 + n_9^4 + n_{10}^4.
\end{cases}$$  

(5.4)

The problem is not shift invariant and therefore as it stands not captured by a Vinogradov mean value theorem of the usual kind. Following Wooley’s approach for $(n, n^3)$ (see [W]), knowledge of the (conjectural) optimal VMVT for $k = 4$ (which would involve the 20th moment) and interpolation with the 6th moment would at the best deliver $I_{10} \ll N^{41/7}$, inferior to (5.3).

A crude summary of our argument. As in [R-S], we need to consider the more general expressions

$$I_p(\lambda) = \int_0^1 \int_0^1 \left| \sum_{n \sim N} e(n^2 x + \lambda n^4 y) \right|^p dxdy$$  

(5.5)

with $p \geq 6$ and $0 < \lambda \leq 1$. A first step is an application of the decoupling theorem from [B-D] for planar curves similarly as in [B-D], Theorem 2.18 (where an extension of the result $I_6(N^{-3}) \ll N^{3+\varepsilon}$ from [R-S] is established). At this stage, one gets shorter sums, of length $M$ say, i.e. $n \in [n_0, n_0 + M]$
with \( n_0 \) ranging in \([N, \frac{N}{2}]\). Exploiting \( n_0 \) as an additional variable leads then to mean value expressions of the form
\[
\int_0^1 \int_0^1 \int_0^1 \int_0^1 \left| \sum_{m \sim M} e(xm + ym^2 + \lambda Nz^3 + \lambda w^4) \right|^p dx dy dz dw \quad (5.6)
\]
to which Theorem 1 is applicable. In the above \( \lambda \) plays the role of a parameter, nothing that \( I_p(\lambda) \) decreases with \( \lambda \) for \( p \) an even integer.

5.1. Preliminary decoupling.

Denote \( S = \sum_{n \sim N} e(n^2 x + \lambda n^4 y) \) and \( S_I = \sum_{n \in I} e(n^2 x + \lambda n^4 y) \) for \( I \subset [\frac{N}{2}, N] \) an interval. Assuming
\[
\lambda N^4 > \frac{N^2}{M^2}, \quad \text{i.e.} \quad \lambda N^2 M^2 > 1 \quad (5.7)
\]
the decoupling theorem for curves gives for \( p \geq 6 \)
\[
\|S\|_p \ll N^\varepsilon \left( \frac{N}{M} \right)^{\frac{3}{2} - \frac{3}{p}} \left( \sum_I \|S_I\|_2^2 \right)^{\frac{3}{2}} \quad (5.8)
\]
with \( \{I\} \) a partition of \([\frac{N}{2}, N]\) in \( M \)-intervals. Hence
\[
I_p(\lambda) \ll N^\varepsilon \left( \frac{N}{M} \right)^{p-3} \left( \frac{1}{N} \sum_{n \sim N} \int_0^1 \int_0^1 |S_{[n,n+M]}(x,y)|^p dx dy \right) \quad (5.9)
\]
where
\[
|S_{[n,n+M]}(x,y)| = \left| \sum_{m \sim M} e((2nx + 4\lambda n^3 y)m + (x + 6\lambda n^2 y)m^2 + 4\lambda nm^3 y + \lambda n^4 y) \right|. \quad (5.10)
\]

5.2. Distributional considerations.

In view of (5.9), (5.10) and exploiting the additional average over \( n \), it is natural to analyze the distribution induced by the map
\[
\varphi : [0,1] \times [0,1] \times \{n \sim N\} \to \mathbb{T} \times \mathbb{T} \times [0, 4N] \times [0, 1]
\]
\[(x, y, n) \mapsto (2nx + 4\lambda n^3 y, x + 6\lambda n^2 y, 4ny, y) = (x', y', z', \omega'). \quad (5.11)
\]
For the time, restrict \( y \) to \([\frac{1}{2}, 1]\) and denote \( \mu \) the (normalized) image measure of \( \varphi \). A translation \( x \mapsto x - 2\lambda n^2 y \) (mod 1) clearly permits to replace \( \varphi \) by the map
\[
(x, y, n) \mapsto (2nx, x + 4\lambda n^2 y, 4ny, y)
\]
and we need to analyze the distribution of $\mu$ at scale $\frac{1}{M} \times \frac{1}{M^2} \times \frac{1}{M^3} \times \frac{1}{M^4}$.

Hence, let $k, \ell \in \mathbb{Z}, |k| \lesssim M, |\ell| \lesssim M^2$ and $\xi, \eta \in \mathbb{R}, |\xi| \lesssim \lambda M^3, |\eta| < \lambda M^4$.

The Fourier transform $\hat{\mu}$ of $\mu$ is given by

$$
\hat{\mu}(k, \ell, \xi, \eta) = \frac{1}{N} \sum_{n \sim N} \int \int dx dy \, e(2nk + (x + 4\lambda n^2 y)\ell + 4ny\xi + y\eta)
$$

implying

$$
|\hat{\mu}(k, \ell, \xi, \eta)| \ll \frac{1}{N} \sum_{n \sim N} 1_{[2nk+\ell=0]} 1_{[|4\lambda n^2 \ell + 4n\xi + \eta| \ll \lambda N^3]}.
$$

It follows from the restrictions on $\xi, \eta$ that

$$
|\ell| \ll \frac{N^3}{\lambda N^2} + \frac{M^3}{N} = \frac{M^4}{N^2} < \frac{1}{\lambda N^{2-\varepsilon}} + \frac{M^3}{N}
$$

and

$$
|k| \ll \frac{1}{\lambda N^{3-\varepsilon}} + \frac{M^3}{N^2}.
$$

Assume further

$$
\lambda > N^{-3+\varepsilon} \text{ and } M < N^{\frac{3}{2}-\varepsilon}
$$

(5.12)

as to ensure $k = \ell = 0$.

Hence $\mu \ll N^\varepsilon \pi_{x', y'}[\mu]$. Returning to (5.9), we may therefore bound

$$
\frac{1}{N} \sum_{n \sim N} \int_0^1 \int_0^1 dx dy \sum_{m \sim M} e((2nx + 4\lambda n^3 y)m + (x + 6\lambda n^2 y)m^2 + 4\lambda nym^3 + \lambda m^4 y)^p
$$

by

$$
\frac{1}{N^{1-\varepsilon}} \sum_{n \sim N} \int_0^1 \int_0^1 \int_0^1 dx' dy' dy \left( \sum_{m \sim M} e(x'm + y'm^2 + 4\lambda nym^3 + \lambda m^4 y) \right)^p
$$

(5.13)

Since $[\frac{1}{2}, 1] \times \{ n \sim N \} \rightarrow [0, 4N] \times [0, 1] : (y, n) \mapsto (4ny, y)$ induces a measure bounded by the uniform measure at scale $1 \times dw'$, it follows that at scale $\frac{1}{M} \times \frac{1}{M^2} \times \frac{1}{M^3} \times \frac{1}{M^4}$, $\mu$ may be majorized by uniform measure up to a factor $N^\varepsilon(1 + \lambda M^3)$. Hence (5.13) may be bounded by

$$
N^\varepsilon(1 + \lambda M^3) \int_0^1 \int_0^1 \int_0^1 \int_0^1 \left| \sum_{m \sim M} e(x'm + y'm^2 + \lambda Nz'm^3 + \lambda w'm^4) \right|^p \, dx' dy' dz' dw'.
$$

(5.14)
One may do better. Assume $\lambda M^3 > 100$ and shift in (5.13) the $y$-variable by $o(\frac{1}{\lambda M^3})$, i.e. replace $y$ by $y + \frac{z}{\lambda M^3}$, $z = o(1)$. One obtains
\[
\frac{1}{N^{1-\varepsilon}} \sum_{n \sim N} \int_0^1 \int_0^1 \int_0^1 \int_0^1 dx'dy'dyz \left| \sum_{m \sim M} e(x'm + y'm^2 + 4\lambda n(y + \frac{z}{\lambda M^3}) m^3 + \lambda m^4 y) \right|^p.
\]
Assuming
\[
\lambda M^4 < N \tag{5.15}
\]
we note that for fixed $\frac{1}{2} \leq y \leq 1$, the map $(n, z) \mapsto n(y + \frac{z}{\lambda M^3})$ induces a normalized measure essentially bounded by $\frac{1}{N} I[0,2N]$. Consequently, under the condition (5.15), (5.13) is bounded by
\[
N^\varepsilon \int_0^1 \int_0^1 \int_0^1 \int_0^1 \sum_{m \sim M} e(x'm + y'm^2 + \lambda N z'm^3 + \lambda w'm^4 \big| |dx'dy'dz'dw'. \tag{5.16}
\]
Taking $M < N^\frac{2}{5}$, (5.15) will hold for $\lambda < N^{-5/3}$.

5.3. Application of mean value theorems.

Use the 8th moment bound (5.12), or equivalently, Theorem 2 in the paper, we get
\[
\max_{|\omega| \leq 1} \int \int \sum_{m \sim M} e(x'm + y'm^2 + \frac{z'}{M} m^3) |^8 \ll M^{4+\varepsilon} \tag{5.17}
\]
Application of (5.17) to (5.16) with fixed $w'$ and $p = 8$ implies then
\[
\int_0^1 \int_\frac{1}{2}^1 \sum_{n \sim N} e(n^2 x + \lambda n^4 y) |^8 dx dy \ll N^\varepsilon \left( \frac{N}{M} \right)^5 M^4 \left( 1 + \frac{1}{\lambda NM^2} \right) \ll N^{4+\frac{1}{4}+\varepsilon} + \frac{N^{2+\varepsilon}}{\lambda} \tag{5.18}
\]
Taking $M = N^{\frac{2}{5}-\varepsilon}$ and $N^{-3+\varepsilon} < \lambda < N^{-2}$.

Braking up the range $y \in [0,1]$ in sub-intervals $[\frac{1}{2}\sigma, \sigma]$, $\sigma = 2^{-s}$ a change of variables and replacement of $\lambda$ by $N^{-\frac{2}{5}\sigma}$ in (5.18) gives
\[
\int_0^1 \int_{N^{-\frac{2}{5}}}^1 \sum_{n \sim N} e(n^2 x + N^{-\frac{2}{5}} n^4 y) |^8 dx dy \ll N^{4+\frac{1}{5}+\varepsilon}.
\]
The remaining range is simply bounded by
\[
N^{-\frac{2}{5}} I_5 (N^{-3}) \leq N^{\frac{2}{5}} I_6 (N^{-3}) \ll N^{\frac{13}{5}+\varepsilon}.
\]
Hence we establish 0.17.
Theorem 11.

\[ I_8 \leq I_8(N^{-\frac{7}{3}}) \ll N^{\frac{14}{3} + \varepsilon} \quad (5.19) \]

Next, one may consider the 10th moment. Setting \( p = 10 \) in \( (5.9) \) implies with \( M = N^{\frac{7}{3} - \varepsilon}, N^{-\frac{4}{3}} < \lambda < N^{-\frac{7}{3}} \)

\[
\int_0^1 \int_0^1 \sum_{n \sim N} e(n^2x + \lambda n^4y)^{10} \, dx \, dy \ll \\
N^\varepsilon \left( \frac{N}{M} \right)^7 \int_0^1 \int_0^1 \int_0^1 \int_0^1 \sum_{m \sim M} e(mx + m^2y + \lambda Nm^3z + \lambda m^4w)^{10} \, dx \, dy \, dz \, dw.
\]

(5.20)

Apply Theorem 4 with \( \varphi_2(t) = t^3, \varphi_3(t) = t^4 \) and \( \delta = \lambda^{-1}N^{-1}M^{-3}, \Delta = \lambda^{-1}M^{-4} \)

This gives the bound

\[
N^\varepsilon \left( \frac{N}{M} \right)^7 \{ \delta \Delta^{3/4} M^7 + (\delta + \Delta) M^6 + M^5 \} \ll N^\varepsilon (N^2 \lambda^{-7/4} + N^{14/3} \lambda^{-1} + N^{47/3}) \ll N^{4+\varepsilon} \lambda^{-1}
\]

(5.21)

for \( \lambda \) as above.

Thus

\[
\int_0^1 \int_{N^{-1}}^1 \sum_{n \sim N} e(n^2x + N^{-\frac{7}{3}}n^4y) \, dx \, dy \ll N^{\frac{47}{3} + \varepsilon}.
\]

(5.22)

The remaining range may be captured using \( (5.19) \), i.e.

\[
\int_0^1 \int_0^{N^{-2/3}} \sum_{n \sim N} e(n^2x + N^{-\frac{7}{3}}n^4y)^{10} \, dx \, dy \ll N^{-\frac{7}{3}} N^2 I_8(N^{-\frac{7}{3}}) \ll N^{\frac{17}{3} + \varepsilon}.
\]

Hence we establish Theorem 11.

REFERENCES

[B-D] J. Bourgain, C. Demeter, The proof of the \( l^2 \)-decoupling conjecture, arXiv: 1405335.

[B-D2] J. Bourgain, C. Demeter, \( \ell^p \) decouplings for hypersurfaces with nonzero Gaussian curvature, in preparation.

[B-G] J. Bourgain, L. Guth, Bounds on oscillatory integral operators based on multilinear estimates, GAFA 21 (2011), no 6, 1239-1295.

[B-I1] E. Bombieri, H. Iwaniec, On the order of \( \zeta(\frac{1}{2} + it) \), Ann. Scuola Norm. Sup. Pisa Cl. Sci (4) 13 (1986), 449–472.

[B-I2] E. Bombieri, H. Iwaniec, Some mean value theorems for exponential sums, Ann. Scuola Norm. Sup. Pisa Cl. Sci (4) 13(1986), 473–486.
[H] M.N. Huxley, *Ares, Lattice Points and Exponential Sums*, LMS monographs, 13 (1996).

[H4] M.N. Huxley, *Exponential sums and the Riemann zeta function, IV*, Proc. London Math. Soc. (3) 66 (1993), 1–40.

[H5] M.N. Huxley, *Exponential sums and the Riemann zeta function, V*, Proc. London Math. Soc (3) 90 (2005), 1–41.

[H-K] M.N. Huxley, G. Kolesnik, *Exponential sums and the Riemann zeta function III*, Proc. London Math. Soc. (3) 62 (1991), 449–468.

[P] S. Parsell, *A note on Weyl’s inequality for eight powers*, (preprint)

[R-S] O. Robert, P. Sargos, *Un théorème de moyenne pour les sommes d’exponentielles. Application à l’inégalité de Weil*, Publ. Inst. math. (Beograd) N.S. 67 (2000), 14–30.

[W] T. Wooley, *Mean value estimates for odd cubic Weyl sums*, arXiv 1401.7152v1, (2014).

[W2] T. Wooley, *Translation invariance, exponential sums and Waring’s problem*, arXiv:1404.3508v1, (2014).

School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540