RECENT PROGRESS ON THE ESTIMATION OF WEYL SUMS

KEVIN FORD

1. Introduction

The classical *Weyl sums* take the form

$$S(\boldsymbol{\alpha}; P) = \sum_{n \le P} e(\alpha_k n^k + \cdots + \alpha_1 n), \quad e(z) = e^{2\pi i z}, \tag{1.1}$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)$ is a point in $\mathbb{R}^k$. Bounds for Weyl sums have numerous applications, particularly in number theory. For example, there are well-known applications to Waring's problem ([6], [5], [17], [31], [32], [34], [35]), the distribution of fractional parts of polynomials ([1], [21], [37]), the theory of the Riemann zeta function ([4], [14], [8], [11], [18], [25], [28], [30]), as well as a recent application to arithmetic progressions in integer sequences ([19], [20]).

Most frequently in applications, one requires upper bounds on $|S(\boldsymbol{\alpha}; P)|$ in terms of rational approximations of the numbers α_i and/or upper bounds for mean values of $|S(\boldsymbol{\alpha}; P)|^n$. In this note we concentrate on recent work (the past 10–15 years) on such bounds, focusing on the quality of the bounds as k becomes large. We omit discussion of the extensive literature on bounding Weyl sums with k small (≤ 20), as well as bounds for multiple exponential sums

$$\sum_{(x_1, \dots, x_n) \in R} e(f(x_1, \dots, x_n)),$$

where f is a polynomial and R is a subset of $\mathbb{R}^n$ (see [3]).

Perhaps the most fundamental quantity is a mean value over all of the coefficients α_i which is known as *Vinogradov's integral* or *Vinogradov's mean value.* It is defined by

$$J_{s,k}(P) = \int_{[0,1]^k} |S(\boldsymbol{\alpha}; P)|^{2s} \, d\boldsymbol{\alpha}. \tag{1.2}$$

Recently proved bounds for $J_{s,k}(P)$ will be discussed in section 2. Bounds for the single variable mean value

$$I_{s,f}(P) = \int_0^1 \left| \sum_{n \le P} e(\alpha f(n)) \right|^{2s} d\alpha, \tag{1.3}$$

where f is a polynomial with integer coefficients, are examined in section 3. Non-averaged upper bounds for $|S(\boldsymbol{\alpha}; P)|$ are described in section 4.

It is often convenient in applications to number theory problems to consider the more general sums

$$S(\boldsymbol{\alpha}; T) = \sum_{n \in T} e(\alpha_k n^n + \cdots + \alpha_1 n), \tag{1.4}$$

Research supported in part by National Science Foundation grant DMS-0070618.

where T is a finite subset of integers $\le P$. There is much literature on bounds for such sums with $T = \mathbb{Z} \cap (P/2, P]$ or $T = \{p \le P : p \text{ prime}\}$ (see e.g. [24]). We do not discuss these here. An extremely fruitful and important line of investigation in recent years, which we will discuss in this note, has been the development of bounds of $S(\boldsymbol{\alpha}; T)$ when $T = \mathscr{A}(P, R)$, the set of integers $\le P$ composed only of prime factors $\le R$ (so-called "smooth" or "friable" numbers). See especially [32], [35] and [43]. See [22] for an account of the distribution of smooth numbers. Section 5 provides a summary of bounds for these so-called *smooth Weyl sums*. Finally, in section 6, we indicate how the bounds in the first 5 sections may be applied to Waring's problem, the fractional parts of polynomials, and to the Riemann zeta function.

Throughout, $\lfloor x \rfloor$ denotes the greatest integer $\le x$, $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x and $\|x\| = \min(\{x\}, 1 - \{x\})$ is the distance from x to the nearest integer. Constants implies by the Landau $O-$ and Vinogradov $\ll -$ symbols do not depend on any parameter unless specified with subscripts. Also, $c_1, c_2, \dots$ will denote certain absolute positive constants.

2. Vinogradov's integral

Bounds for the mean value (1.2) were first considered by Vinogradov in the 1930's. Crucial to estimations is the observation that $J_{s,k}(P)$ is equal to the number of solutions of the simultaneous Diophantine equations

$$x_1^j + \cdots + x_s^j = y_1^j + \cdots + y_s^j \quad (1 \le j \le k); \quad 1 \le x_i, y_i \le P. \tag{2.1}$$

More generally, for $\mathbf{h} = (h_1, \dots, h_k)$, let $J_{s,k}(P; \mathbf{h})$ be the number of solutions of

$$\sum_{i=1}^{s} (x_i^j - y_i^j) = h_j \quad (1 \le j \le k); \quad 1 \le x_i, y_i \le P.$$

In particular,

$$\begin{aligned} J_{s,k}(P; \mathbf{h}) &= \int_{[0,1]^k} |S(\boldsymbol{\alpha}; P)|^{2s} e(-\alpha_1 h_1 - \cdots - \alpha_k h_k) d\boldsymbol{\alpha} \\ &\le J_{s,k}(P; (0, \dots, 0)) = J_{s,k}(P). \end{aligned} \tag{2.2}$$

Hence, writing $Q = \lfloor P \rfloor$, we obtain

$$Q^{2s} = \sum_{\substack{\mathbf{h} \\ |h_j| \le s(Q^j - 1)}} J_{s,k}(P; \mathbf{h}) \le \sum_{\mathbf{h}} J_{s,k}(P) \le (2s)^k Q^{k(k+1)/2} J_{s,k}(P). \tag{2.3}$$

Counting only the solutions of (2.1) with $x_i = y_i$ for each i gives a trivial lower bound $J_{s,k}(P) \ge Q^s$. Therefore

$$J_{s,k}(P) \ge \max\left((2s)^{-k} \lfloor P \rfloor^{2s - \frac{1}{2}k(k+1)}, \lfloor P \rfloor^s \right). \tag{2.4}$$

It is conjectured that the right side of 2.4 is the true order of magnitude of $J_{s,k}(P)$, at least up to a factor P^{ε}. Upper bounds take the form of

$$J_{s,k}(P) \le D(s,k) P^{2s - \frac{1}{2}k(k+1) + \eta}, \tag{2.5}$$

where $\eta \ge 0$ and $D(s,k)$ is independent of P. For convenience, we shall *define* $\eta(s,k)$ to be the infimum of values of η for which (2.5) holds for some constant $D(s,k)$ (where $D(s,k)$ may depend on η also).

For example, Stechkin [13] in 1975 proved the all-purpose bound

$$\eta(rk,k) \le \frac{1}{2}k^2(1-1/k)^r. \tag{2.6}$$

Bounding $J_{s,k}(P)$ is usually accomplished in an iterative way, by relating $J_{s,k}(P)$ to $J_{s',k}(P')$ where $s' < s$ and $P' < P$. For instance, the p-adic method in its basic form gives

$$J_{s,k}(P) \ll_{k,s} P^{2s/k+\frac{1}{2}(3k-5)} J_{s-k,k}(P^{1-1/k}).$$

Iterating this inequality gives (2.6). There have been several improvements of (2.6). First are improvements in the exponent when s is small ($s \le \frac{1}{2}k(k+1)$). For such s it is conjectured that $J_{s,k}(P) \ll_\varepsilon P^{s+\varepsilon}$ for every $\varepsilon > 0$, so we define $\delta(k,s)$ to be the infimum of numbers δ for which

$$J_{s,k}(P) \ll_{s,k,\delta} P^{s+\delta}. \tag{2.7}$$

Table 1 lists upper bounds for $\delta(k,s)$.

$\delta(k,s)$	Reference
$O(s^{3/2}/k)$	Archipov and Karatsuba [2], 1978
$O(s^2/k^2)$	Tyrina [12], 1987
$O(sk^{3/2}\exp\{-c_1k^3/s^2\})$	Wooley [38], 1994

TABLE 1. Upper bounds on $J_{s,k}(P)$ for small s

In particular, for $s \le c_2k^{3/2}/\sqrt{\log k}$, Wooley's result implies $\delta(k,s) \ll 1/k$, which is very close to the conjectured bound. Wooley's bound is based on iteration of the following result (Lemma 4.2 of [38]) which relates bound for $J_{s,k}(P)$ in terms of bounds for quantities $J_{s',k}(Q)$ with $s' < s$ and $s' > s$.

Lemma 2.1. *Suppose* $3 \le r \le k$, $3 \le t < 2\lfloor k/2 \rfloor$, $r+t \ge k$ *and put* $u = \lfloor 1/(1-\frac{1}{2}t/\lfloor k/2\rfloor)+1 \rfloor$. *Suppose* $1 < M \le P^{1/r}$ *and put* $H = PM^{-r}$ *and* $Q = PM^{-1}$. *Then*

$$J_{s+t,k}(P) \ll_{k,s} M^{2s+\frac{1}{2}(r+t-k+1)(r+t-k)} \left[P^t J_{s,k}(Q) + (PH)^{t/2} J_{u,k}(Q)^{s/u} \right].$$

The most substantial improvement of Stechkin's bound throughout the range $1 \le s \ll k^2 \log k$ is due to Wooley [36]. He bounds $J_{s+k,k}(P)$ in terms of $J_{s,k}(Q)$ using an iterative method which we now sketch. Suppose $0 \le d \le k-1$ and T is a positive integer. We say the k-tuple of polynomials $\mathbf{\Psi} = (\Psi_1, \ldots, \Psi_k) \in \mathbb{Z}[x]^k$ is of type (d,T) if Ψ_j is identically zero for $j \le d$, and when $j > d$, Ψ_j has degree $j-d$ with leading coefficient $\frac{j!}{(j-d)!}T$.

Fix k and suppose $1 \le r \le k$. If $\mathbf{\Psi} = (\Psi_1, \ldots \Psi_k)$ is a system of polynomials, let $K_s(P,Q;\mathbf{\Psi};q)$ be the number of solutions of the simultaneous equations

$$\sum_{i=1}^{k}(\Psi_j(z_i) - \Psi_j(w_i)) + q^j \sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \qquad (1 \le j \le k),$$
$$1 \le z_i, w_i \le P; \quad 1 \le x_i, y_i \le Q.$$

Here the inequalities on the variables z_i, w_i, x_i, y_i hold for every i. For prime p, let $L_s(P,Q;\mathbf{\Psi};p,q)$ be the number of solutions of

$$\sum_{i=1}^{k}(\Psi_j(z_i)-\Psi_j(w_i))+(pq)^j\sum_{i=1}^{s}(u_i^j-v_i^j)=0 \qquad (1\le j\le k),$$
$$1\le z_i,w_i\le P;\ z_i\equiv w_i \pmod{p^k};\ 1\le u_i,v_i\le Q.$$

Below are the main iteration lemmas (essentially Lemmas 3.1 and 4.1 of [36]; see also Lemmas 3.2 and 3.3 of [18]).

Lemma 2.2. *Suppose $P^{1/2k}\le M\le P^{1/k}$, $M\ll Q\le P$ and $\mathbf{\Psi}$ is of type (d,T). For some prime in $[M,2M]$ and a system of polynomials $\mathbf{\Phi}$ of type (d,T) we have*

$$K_s(P,Q;\mathbf{\Psi};q)\ll_{k,s} M^{2s+\frac12(r^2-r+d^2-d)}L_s(P,\tfrac{Q}{p};\mathbf{\Phi};p,q).$$

Lemma 2.3. *Suppose that $s\ge d$, $d\le k-2$, $q\ge 1$, p is a prime and $\mathbf{\Phi}$ is a system of polynomials of type (d,T). Then there is a system of polynomials $\mathbf{\Upsilon}$ of type $(d+1,T')$ with $T\le T'\le PT$ such that*

$$L_s(P;Q;\mathbf{\Phi};p,q)\le(2P)^k\max\big[k^kJ_{s,k}(Q),2p^{-k^2}\left\{J_{s,k}(Q)K_s(P,Q;\mathbf{\Upsilon};pq)\right\}^{1/2}\big].$$

Starting with $\mathbf{\Psi}=(x,x^2,\dots,x^k)$ and iterating these two lemmas produces the bounds (see [17], inequality (5.2))

(2.8)
$$\eta(s,k)\le\left(\frac12+O\left(\frac{1}{\sqrt{\log k}}\right)\right)k^2e^{1/2-2s/k^2} \qquad (1\le s\le k^2(\log k-\log\log k)).$$

A slight improvement (though important in the application to the Riemann zeta function) was given by the author ([18], Theorem 3):

$$\eta(s,k)\le\left(\frac38+o(1)\right)k^2e^{1/2-2s/k^2}. \tag{2.9}$$

For even larger s it is possible to prove an asymptotic formula

$$J_{s,k}(P)=C(k,s)P^{2s-\frac12k(k+1)}+O(P^{2s-\frac12k(k+1)-c_3} \tag{2.10}$$

using the circle method. This was first accomplished by Vinogradov in the 1930's. In 1957, Hua [24] proved that (2.10) holds for $s\ge(3+o(1))k^2\log k$. This was not improved until Wooley [36] used (2.8) to show that (2.10) holds with $s\ge(5/3+o(1))k^2\log k$. Later Wooley [42] showed that (2.10) holds with $s\ge(1+o(1))k^2\log k$.

3. Single variable mean values

By orthogonality, $I_{s,f}(P)$ is the number of solutions of the equation

$$f(x_1)+\cdots+f(x_s)=f(y_1)+\cdots+f(y_s) \qquad (1\le x_i,y_i\le P).$$

Suppose that

$$f(x)=a_1x+\cdots+a_kx^k,\quad a_k\ne0.$$

By an argument similar to the one leading to (2.4), we have

$$I_{s,f}(P)\gg_{s,f}\max(P^s,P^{2s-k}).$$

It is conjectured that the right side represents the true order of magnitude of $I_{s,f}(P)$, at least up to a factor P^ε.

It is possible to use bounds for $J_{s,k}(P)$ to bound $I_{s,f}(P)$ in a simple way. Recalling the definition of $J_{s,k}(P;\mathbf{h})$ from section 2, we find that

$$\begin{aligned} I_{s,f}(P) &= \sum_{h_1,\dots,h_{k-1}} J_{s,k}(P;(h_1,h_2,\dots,h_{k-1},(a_1h_1+\cdots+a_{k-1}h_{k-1})/a_k) \\ &\le \sum_{h_1,\dots,h_{k-1}} J_{s,k}(P) \\ &\ll_s P^{k(k-1)/2} J_{s,k}(P) \\ &\ll_{s,\varepsilon} P^{2s-k+\eta(s,k)+\varepsilon}. \end{aligned} \tag{3.1}$$

Bombieri ([15], Theorem 5) proved a more general inequality which implies that

$$I_{s,f}(P) \ll_s \frac{P^{2s} J_{s,k}(P)}{J_{s,k-1}(P)}. \tag{3.2}$$

However, applying the lower bound (2.4) to the denominator $J_{s,k-1}(P)$, we recover (3.1).

Until recently, (3.1) was the best known method for bounding $I_{s,f}(P)$ when k is large. Using Wooley's bounds (2.8) produces non-trivial upper bounds for $I_{s,f}(P)$ only when $s \gg k^2 \log k$.

In 1995, the author [17] developed a more sophisticated method of bounding $I_{s,f}(P)$, which produces the bounds

$$I_{s,f}(P) \ll_{s,f,\varepsilon} P^{2s-k+\frac{1}{m}\eta(s-m(m-1)/2,k)+\varepsilon}, \tag{3.3}$$

where $1 \le m \le k$. This gives non-trivial upper bounds when $s \gg k^2$. The fundamental idea is to transition from $I_{s,f}(P)$ to $J_{s,k}(P)$ in a multi-step procedure, adding one equation at each step. Specifically, let $I_{s,m}(P;q,b)$ be the number of solutions of

$$\begin{aligned} &\sum_{i=1}^{s}(f(qx_i+b)-f(qy_i+b)) = 0 \\ &\sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \qquad (1 \le j \le m-1) \end{aligned}$$

with $0 \le x_i, y_i \le P/q$ for each i. Suppose $k \ge 3$, $1 \le m \le k-1$, $q \approx P^{1-1/m}$, $0 \le b < q$ and $(f^{(j)}(b),q) = 1$ for $1 \le j \le k-1$. Then for some prime $p \approx P^{1/(m(m+1))}$ and $0 \le a < p$ we have (Lemma 4.1 of [17])

$$I_{s,m}(P;q,b) \ll_{s,f} p^{2s-2m+\frac{3}{2}m(m+1)} I_{s-m,m+1}(P;pq,b+aq).$$

Iterating this expression leads to (3.3). However, in light of (2.4), one can never prove the conjectured bound for $I_{s,f}(P)$ using (3.3).

For larger s one can prove via the circle method an asymptotic formula

$$I_{s,f}(P) = (C(f,s)+o(1))P^{2s-k} \qquad (P \to \infty). \tag{3.4}$$

See for example the proof of Theorem 5.4 of Vaughan [33]. Hua's bound [23] for $J_{s,k}(P)$ implies that (3.4) holds for $s \ge (2+o(1))k^2 \log k$. Wooley's bound (2.8) gives (3.4) holds for $s \ge (1+o(1))k^2 \log k$, and (3.3) implies (3.4) for $s \ge (1/2+o(1))k^2 \log k$.

4. NON-AVERAGED UPPER BOUNDS

The fundamental idea behind upper bounds on $|S(\boldsymbol{\alpha};P)|$ is that if $|S(\boldsymbol{\alpha};P)|$ is large at a point $\boldsymbol{\alpha}$, then it also large on a set of points $\boldsymbol{\alpha}\in[0,1]^k$ of large measure. Upper bounds on $J_{s,k}(P)$ then imply upper bounds on $|S(\boldsymbol{\alpha};P)|$. Given an $\boldsymbol{\alpha}$ with $|S(\boldsymbol{\alpha};P)|$ large, Taylor's theorem implies that $|S(\boldsymbol{\beta};P)|$ will be large when $\boldsymbol{\beta}$ is "close" to $\boldsymbol{\alpha}$ (in some small box in $\mathbb{R}^k$ centered at $\boldsymbol{\alpha}$). For example, Montgomery [27] proves that if $|\alpha_j-\beta_j|\le\frac{1}{2kP^j}$ for $1\le j\le k$, then

$$\frac{|S(\boldsymbol{\alpha};P)|}{1+\pi}<|S(\boldsymbol{\beta};P)|<(1+\pi)|S(\boldsymbol{\alpha};P)|. \tag{4.1}$$

The other method used to produce large Weyl sums from a given one is to shift the variable of summation. Trivially

$$\left|\sum_{n\le P}e(\alpha_k n^k+\cdots)-\sum_{N<n\le P+N}e(\alpha_k n^k+\cdots)\right|\le 2N. \tag{4.2}$$

Making the substitution $n=h+N$, we see that the second sum is $S(\boldsymbol{\beta};P)$, where

$$\beta_j=\sum_{i=j}^{k}\binom{i}{j}\alpha_i N^{i-j}\quad(1\le j\le k). \tag{4.3}$$

If α_k is not too close to a rational number with small denominator, $\boldsymbol{\beta}$ will not be too close to $\boldsymbol{\alpha}$. Combined, these two ideas produce many small boxes in $[0,1]^k$ on which the associated Weyl sums are about the same magnitude.

The most standard estimate, due to Vinogradov, depends on a rational approximation of a single α_j: Suppose $|\alpha_j-p/q|\le 1/q^2$, where p and q are integers, $\gcd(p,q)=1$ and $P\le q\le P^{j-1}$. Then we have (see Chapter 5 of [33])

$$|S(\boldsymbol{\alpha};P)|\ll_{k,s}\left(J_{s,k-1}(2P)P^{k(k-1)/2-1}\right)^{1/2s}\log P, \tag{4.4}$$

which holds for any $s\ge1$. In practice one usually uses this result with $j=k$ because (i) the term $\alpha_k n^k$ is the most oscillating part of the exponential in (1.1) and (ii) the larger possible range of q. Inequality (4.4) is non-trivial only if $J_{s,k-1}(2P)=o(P^{2s-\frac12 k(k-1)+1})$, i.e. when $\eta(s,k)<1$. Let $\rho(k)$ be the supremum of all numbers ρ such that if $|\alpha_k-p/q|\le 1/q^2$ with $\gcd(p,q)=1$ and $P\le q\le P^{k-1}$, one has

$$|S(\boldsymbol{\alpha};P)|\ll P^{1-\rho}.$$

Then (4.4) implies that

$$\rho(k)\le\min_{s\ge1}\frac{1-\eta(s,k-1)}{2s}. \tag{4.5}$$

In particular, (2.8) implies that $\rho(k)^{-1}\le(2+o(1))k^2\log k$. this improves a result of Hua [23], who proved that $\rho(k)^{-1}\le(4+o(1))k^2\log k$. A further improvement was given by Wooley [41], who proved that $\rho(k)^{-1}\le(3/2+o(1))k^2\log k$. The key idea was to use a result of Bombieri ([15], Corollary 1 to Theorem 8), which states that for $P^r\le q\le P^{k-r}$,

$$|S(\boldsymbol{\alpha};P)|\ll_\varepsilon P^{1-\mu+\varepsilon},\quad \mu=\frac{r-\eta(s,k-1)}{2rs}. \tag{4.6}$$

Bombieri [15] has interesting discussions on different approaches for estimating $|S(\boldsymbol{\alpha};P)|$.

Bounds for $|S(\boldsymbol{\alpha};P)|$ which depend on rational approximations of more than one of the numbers α_j are more complicated to state. We defer discussion of them until section 6, where we concentrate on the estimates related to the Riemann zeta function.

5. Smooth Weyl sums

The general smooth Weyl sum is defined by

$$S(\boldsymbol{\alpha};P,R) = \sum_{n\in\mathscr{A}(P,R)} e(\alpha_1 n + \cdots + \alpha_k n^k).$$

In some applications, it suffices to consider the simpler sum

$$f_k(\alpha;P,R) = \sum_{n\in\mathscr{A}(P,R)} e(\alpha n^k).$$

The advantages of using $f_k(\alpha;P,R)$ are (i) n^k is multiplicative, and (ii) the numbers in $\mathscr{A}(P,R)$ have closely spaced divisors (if $n\in\mathscr{A}(P,R)$ and $1\le T\le n$ then n has a divisor in $[T,TR]$). Using these facts, Vaughan [Va1] showed that one can imitate the iterative method used to bound Vinogradov's integral $J_{s,k}(P)$ to bound the mean value

$$S_s(P,R) = \int_0^1 |f_k(\alpha;P,R)|^{2s}\,d\alpha. \tag{5.1}$$

In particular, Vaughan showed that for each $s\ge 1$, for every $\varepsilon>0$ and for $\eta>0$ depending on ε,k,s, that

$$S_s(P,P^\eta) \ll_{k,s,\varepsilon} P^{2s-k+k(1-1/k)^s+\varepsilon}. \tag{5.2}$$

This can be considered an analog of Stechkin's bound for $J_{s,k}(P)$, and is greatly superior to known bounds for $I_s(f;P)$ (see section 3). Indeed, the bounds come very close to the conjectured upper bounds $S_s(P,P^\eta)\ll_{k,s,\varepsilon} P^{2s-k+\varepsilon}$ when s has order $k\log k$. Vaughan also exploited the properties of $\mathscr{A}(P,R)$ to prove for α not too close to a rational number with small denominator that

$$|f_k(\alpha;P,P^\eta)| \ll_{k,\eta} P^{1-\nu(k)}, \quad \nu(k)^{-1} = (4+o(1))k\log k. \tag{5.3}$$

Again this is vastly superior to the known upper bounds for $|S(\boldsymbol{\alpha};P)|$ given in section 4, e.g. (4.5).

Wooley developed a more sophisticated iterative scheme, similar to the method he used to prove (2.8), which improved both (5.2) and (5.3). In [35], [37] he proves that for each k,s and for η depending on k,s,

$$S_s(P,P^\eta) \ll_{s,k,\eta} P^{2s-k+ke^{1-2s/k}}, \tag{5.4}$$

the significance being that the exponent of P approaches $2s-k$ twice as fast as the exponent of P in (5.2). One consequence of this mean value is (5.3) with $\nu(k)^{-1}=(2+o(1))k\log k$. This last bound was improved further by Wooley [40], so that (5.3) holds with $\nu(k)^{-1}=(1+o(1))k\log k$.

Improvements to the upper bounds for $S_s(P,P^\eta)$ when $s\ll\sqrt{k}$ have been given by both Vaughan and Wooley. In this range of s (indeed for $s\le k$) it is expected that $S_s(P,P^\eta)\ll_{s,k,\varepsilon} P^{s+\varepsilon}$, and the estimates of Vaughan and Wooley for small s take the form

$$S_s(P,P^\eta) \ll P^{s+\gamma(s,k)} \tag{5.5}$$

where $\gamma(s,k)$ is small. Vaughan [32] proved (5.5) with

$$\gamma(s,k) \approx e^{-C\log^2(k/s^2)}.$$

Wooley improved this a little bit in 1994 [38], then in 1995 [39] greatly improved the bound, showing (5.5) holds with

$$\gamma(s,k) = \frac{4\sqrt{k}}{es} \exp\left\{ -\frac{4k}{e^2 s^2} \right\} \qquad (2 < s \le (2/e)\sqrt{k}).$$

The techniques used are similar to those used by Wooley to prove his upper bound for $\delta(k,s)$ given in Table 1.

In 1997, Wooley [43] extended the method for bounding $S_s(P,P^\eta)$ to bound the more general smooth Weyl sums

$$f_{\mathbf{k}}(\boldsymbol{\alpha}; P, P^\eta) = \sum_{n\in\mathscr{A}(P,P^\eta)} e\left(\sum_{j=1}^t \alpha_j n^{k_j} \right),$$
$$\mathbf{k} = (k_1, k_2, \ldots, k_t), \quad k_1 > k_2 > \cdots > k_t.$$

and the associated mean values

$$S_{s,\mathbf{k}}(P,P^\eta) = \int_{[0,1]^t} |f_{\mathbf{k}}(\boldsymbol{\alpha}; P, P^\eta)|^{2s} \, d\alpha_1 \cdots d\alpha_t.$$

Define $\Delta(s,\mathbf{k})$ to be the infimum of numbers Δ such that for some $\eta > 0$ we have

$$S_{s,\mathbf{k}}(P,P^\eta) \ll_{s,\mathbf{k},\eta} P^{2s-(k_1+\cdots+k_t)+\Delta}.$$

Then we have (Theorem 2 of [43])
(5.6)

$$\Delta(s,\mathbf{k}) \le \begin{cases} tk_1 e^{2-2s/(tk_1)} & 1 \le s \le s_0 := \frac{1}{2}tk_1(\log(tk_1) - 2\log\log k_1) \\ e^3 \log^2 k_1 (1-1/k_1)^{(s-s_0)/t} & s > s_0. \end{cases}$$

These mean value bounds interpolate between Wooley's bounds for $S_s(P,P^\eta)$ given in section 5 and his bounds for $J_{s,k}(P,P^\eta)$ given in section 2. Likewise, Wooley's bounds for $|f_{\mathbf{k}}(\boldsymbol{\alpha}; P, P^\eta)|$ interpolate between known bounds for $|S(\boldsymbol{\alpha}; P)|$ and bounds for $|f_k(\alpha; P, P^\eta)|$ (see [43], Theorems 4 and 5).

6. Some applications

Here we briefly indicate some of the applications of the estimates from the previous sections. Fuller expositions may be found in the original papers, as well as the monographs [24], [33] (for Waring's problem and the circle method), [1] (for problems on fractional parts of polynomials), [30], [25], [14] (for applications to the Riemann zeta function).

I. **Waring's problem**. Let $R_{s,k}(n)$ be the number of solutions in non-negative integers $x_1, \ldots, x_s$ of the equation $x_1^k + \cdots + x_s^k = n$. The main quantity one wants to bound in Waring's problem is $G(k)$, the smallest number s so that $R_{s,k}(n) > 0$ for all sufficiently large n. The existence of $G(k)$ for all k was established by Hilbert in 1909. In the 1920's, Hardy and Littlewood developed the *circle method* to attack this and related additive problems. The circle method is based on the identity

$$R_{s,k}(n) = \int_0^1 g(\alpha)^s e(-n\alpha)\, d\alpha$$

where

$$g(\alpha) = \sum_{m \le n^{1/k}} e(\alpha m^k).$$

One then breaks $[0, 1]$ into many subintervals, applying different estimates for $g(\alpha)$ on each subinterval. In this way Hardy and Littlewood proved upper bounds on $G(k)$ roughly of order 2^k. The machinery of the circle method was improved many times by Vinogradov, the most significant innovation being the introduction of the mean value (1.2). Vinogradov showed in the 1930's that $G(k) \ll k \log k$. Table 2 summarizes more recent progress on bounding $G(k)$ for large k. Vinogradov's 1959 result depended on his asymptotic formula (2.10). Wooley's bound depended on his bounds (5.4).

Upper bound on $G(k)$	Reference
$3k(\log k + O(1))$	Vinogradov [6], 1947
$2k(\log k + O(\log \log k))$	Vinogradov [5], 1959
$k(\log k + O(\log \log k))$	Wooley [35], 1992

TABLE 2. Progression of upper bounds on $G(k)$.

Hardy and Littlewood also established an asymptotic formula for $R_{s,k}(n)$ as $n \to \infty$, provided s is large enough. Let $\widetilde{G}(k)$ be the smallest value of s for which the Hardy-Littlewood asymptotic formula holds. Table 3 gives recent progress on bounding $\widetilde{G}(k)$. Wooley's bound was a direct consequence of his improved bounds (2.8) on $J_{s,k}(P)$. The author's 1995 result is a consequence of his improvements to the bounds for $I_s(P; f)$ (3.3).

Upper bound on $\widetilde{G}(k)$	Reference
$(4 + o(1))k^2 \log k$	Hua [23], 1949
$(2 + o(1))k^2 \log k$	Wooley [36], 1992
$(1 + o(1))k^2 \log k$	Ford [17], 1995

TABLE 3. Progression of upper bounds on $\widetilde{G}(k)$.

II. **Fractional parts of polynomials**. The main idea in this topic, due to H. Weyl, is that the fractional parts $\{a_i\}$ of a sequence $a_1, a_2, \ldots$ are uniformly distributed in the interval $[0, 1]$ provided that

$$\sum_{m=1}^{N} e(h a_m) = o(N)$$

for every $h \geq 1$. Of particular interest is the distribution of integers n for which $\|f(n)\|$ is small, where $f(x) \in \mathbb{R}[x]$. Define $\tau(f)$ to be the supremum of all numbers τ for which

$$\min_{n \leq N} \|f(n)\| \ll_\tau N^{-\tau}.$$

The polynomial αn^k has received the most attention, and Table 4 details recent progress on bounding $\tau(\alpha n^k)$.

Upper bound on $\tau(\alpha n^k)^{-1}$	Reference
$(4 + o(1))k \log k$	Baker [1], 1986
$(2 + o(1))k \log k$	Wooley [37], 1993
$(1 + o(1))k \log k$	Wooley [40], 1995

TABLE 4. Progression of upper bounds on $\tau(\alpha n^k)^{-1}$.

For an arbitrary polynomial $f(n) = \alpha_1 x + \cdots + \alpha_k x^k$, the bounds are considerably worse. Let $\tau^*(k)$ be the supremum of $\tau(f)$ over all polynomials f of degree k. Baker [1] proved that $\tau^*(k)^{-1} \leq (8 + o(1))k^2 \log k$ and this was improved by Wooley [36] to $\tau^*(k)^{-1} \leq (4 + o(1))k^2 \log k$ as a consequence of (2.8).

One may also ask for *unlocalized* bounds. Let $\sigma(f)$ be the supremum of all numbers σ for which $\|f(n)\| \leq n^{-\sigma}$ has infinitely many solutions (they may be widely separated). Heath-Brown [21] showed that $\sigma(\alpha n^k)^{-1} \leq 14.425k$ and Wooley [37] use the theory of smooth Weyl sums to improve this to $\sigma(\alpha n^k)^{-1} \leq 9.063k$.

In [43], Wooley uses bounds for general mean values (5.6) to bound $\sigma(f)$ and $\tau(f)$ when $f(n) = \alpha_1 n^{k_1} + \cdots + \alpha_t n^{k_t}$ and $1 \leq k_t < \cdots < k_1$. Wooley proved that ([43], Theorems 6 and 7)

$$\begin{aligned} \tau(f) &\geq \max_{2s \geq k_1+1} \frac{1 - (t-1)\Delta(s, \mathbf{k})}{2st + 1 + \Delta(s, \mathbf{k})} \\ &\geq \frac{1}{t^2 k_1 (\log(k_1 t^3) + O(\log \log k_1))} \end{aligned}$$

and that ([43], Theorem 8, Corollary 8.1)

$$\begin{aligned} \sigma(f) &\geq \max_{s \geq 1} \frac{k_1 - 2\Delta(s, \mathbf{k})}{4s^2} \\ &\geq \frac{1}{t^2 k_1 (\log t + \log \log t + 7)^2}. \end{aligned}$$

III. **The Riemann zeta function**. The location of the non-trivial zeros of the Riemann zeta function $\zeta(s)$ is critically important to numerous questions in prime number theory. It is conjectured (the Riemann Hypothesis) that they all lie on the line $\Re s = \frac{1}{2}$, but what we know is far weaker and progress has been very slow. Table 5 summarizes progress of the zero-free region. For each line in Table 5, $\zeta(s) \neq 0$ for $\Re s \geq 1 - \beta(t)$, and the letter c stands for an absolute constant, which may be different on each line.

$\beta(t)$	Reference
$\frac{c}{\log \lvert t\rvert}$	de la Vallée Poussin [16], 1899
$\frac{c \log\log \lvert t\rvert}{\log \lvert t\rvert}$	Littlewood [26], 1922
$\frac{c(\varepsilon)}{(\log \lvert t\rvert)^{3/4+\varepsilon}}$	Tchudakoff [29], 1938
$\frac{c}{(\log \lvert t\rvert)^{2/3}(\log\log \lvert t\rvert)^{1/3}}$	Korobov [11], Vinogradov [4], 1958

TABLE 5. Zero-free regions for $\zeta(s)$.

The zero-free region depends on upper bounds for $|\zeta(s)|$ for s inside the critical strip $0 < \Re s < 1$, and these bounds ultimately depend on exponential sum bounds of the type

$$S(N,t) = \sum_{N<n\le 2N} n^{-it} = \sum_{N<n\le 2N} e^{-it\log n}. \tag{6.1}$$

Tchudakoff, Korobov and Vinogradov apply the Weyl shifting technique and approximate $\log(1+x)$ by a Taylor series of order k, reducing the problem to bounding a Weyl sum for which bounds on $J_{s,k}(P)$ (with explicit constants $D(s,k)$) come into play. It is not difficult to work out the constant $D(s,k)$, even in Wooley's method. One gets for $s \ll k^2$ that $D(s,k) = k^{c_3 k^3}$ works (see e.g. [18], Theorem 3).

The Weyl sum estimates of Korobov and Vinogradov ([9], [8], [7], [10], [11], [4]) depend on rational approximations of many of the coefficients α_i of the Weyl sums which arise. Assuming conditions on several of the coefficients α_j, the methods produce bounds of the form

$$|S(\boldsymbol{\alpha}; P)| \ll P^{1-c_3/k^2},$$

improving on (4.5). A particularly nice estimate was given by Vinogradov [4], and depends on using (4.2) with $N = ab$ and summing on both a and b. Incorporating a small refinement due to Heath-Brown (see [30], pp. 134–137 and [18], Lemma 5.1), one obtains the following.

Lemma 6.1. *Suppose $k \ge 2$, $s \ge 2$ and $1 < M \le \sqrt{N}$. Then*

$$S(N,t) \le N\left[(5s)^k M^{k(k+1)/2-4s}(J_{s,k}(M))^2 W_1\cdots W_k\right]^{\frac{1}{2s^2}} + 2M^2 + \frac{tM^{2k+2}}{kN^k},$$

where, for $1 \le j \le k$,

$$\begin{aligned} W_j &= \max_{N<n\le 2N} \left|\{|d| < sM^j - 1 : \|d\tfrac{t}{2\pi j n^j}\| < \tfrac{1}{2sM^j}\}\right| \\ &\le \min\left(2sM^j, 4 + \frac{stM^j}{\pi j N^j} + \frac{4\pi j(2N)^j}{stM^j}\right). \end{aligned}$$

This lemma can be used to prove an upper bound for $\zeta(s)$ near the line $\Re s = 1$. In 1967, Richert [28] proved that

$$|\zeta(\sigma+it)| \ll |t|^{B(1-\sigma)^{3/2}} \log^{2/3}|t| \quad (|t| \ge 3) \tag{6.2}$$

with $B = 100$. Recently, the author proved (6.2) with $B = 4.45$ [18]. The main innovation was incorporating bounds for smooth Weyl sums into the argument. For a set $\mathscr{B}$ of positive integers, let $J_{s,g,h}(\mathscr{B})$ be the number of solutions of the system

$$\sum_{i=1}^{s}(x_i^j - y_i^j) = 0 \qquad (h \le j \le g), x_i, y_i \in \mathscr{B}.$$

Lemma 6.2. *Suppose k, r and s are integers ≥ 2, and h and g are integers satisfying $1 \le h \le g \le k$. Let N be a positive integer, and M_1, M_2 be real numbers with $1 \le M_i \le N$. Let $\mathscr{B}$ be a nonempty subset of the positive integers $\le M_2$. Then*

$$(6.3) \quad S(N,t) \le 2M_1M_2 + \frac{t(M_1M_2)^{k+1}}{kN^k} + N\left(\frac{M_2}{|\mathscr{B}|}\right)^{\frac{1}{r}} \left((5r)^k M_2^{-2s} \lfloor M_1 \rfloor^{-2r+k(k+1)/2} J_{r,k}(M_1) J_{s,g,h}(\mathscr{B}) W_h \cdots W_g\right)^{\frac{1}{2rs}},$$

where

$$W_j = \max_{N<n\le 2N} |\{|d| < sM_2^j - 1 : \|d\tfrac{t}{2\pi j n^j}\| < \tfrac{1}{2sM_1^j}\}| \le \min\left(2sM_2^j, \frac{2sM_2^j}{r\lfloor M_1\rfloor^j} + \frac{stM_2^j}{\pi j N^j} + \frac{4\pi j(2N)^j}{rt\lfloor M_1\rfloor^j} + 2\right).$$

One applies this with g and h relatively close together (so that the parameter t in (5.6) is small; about $0.06k$), $\mathscr{B} = \{n \le M_2 : p|n \implies P^{\eta/2} < p \le P^{\eta}\}$ and then uses a version of (5.6) with explicit constants. The reason for using this $\mathscr{B}$ is that the number in $\mathscr{A}(P, P^{\eta})$ might have a large number of prime factors, and this leads to very large constants in (5.6).

References

1. R. C. Baker, *Diophantine inequalities*, Clarendon Press, Oxford, 1986.
2. Г. И. Архипов and А. А. Карацуба, *Новая оценка интеграла И. М. Виноградова*, Известия Ак. Наук СССР, серия матем. **42** (1978), no. 4, 751–762, English translation: G. I. Arhipov and A. A. Karacuba, *A new estimate of an integral of I. M. Vinogradov* Math. USSR Izvestiya **13** (1979), 52–62.
3. Г. И. Архипов, А. А. Карацуба, and В. Н. Чубариков, *Теория кратных тригонометрических сумм*, Наука, Москва, 1987.
4. И. М. Виноградов, *Новая оценка функции $\zeta(1+it)$*, Известия Ак. Наук СССР, серия матем. **22** (1958), 161–164.
5. ———, *К вопросу о верхней границе для $G(n)$*, Известия Ак. Наук СССР, серия матем. **23** (1959), 637–642.
6. ———, *Метод тригонометрических сымм в теопии чисел*, Наука, Москва, 1980, English translation of 1947 edition, revised and annotated by K. F. Roth and Anne Davenport; Interscience Publishers, London and New York; 1954.
7. Н. М. Коробов, *Новые теоретикочисловые оценки*, Доклады Ак. Наук СССР **119** (1958), 433–434.
8. ———, *О нулях функции $\zeta(s)$*, Доклады Ак. Наук СССР **118** (1958), 431–432.
9. ———, *Об оценке рационал ных тригонометрических сумм*, Доклады Ак. Наук СССР **118** (1958), 231–232.
10. ———, *Оценки сумм Вейля и распределение простых чисел*, Доклады Ак. Наук СССР **123** (1958), no. 1, 28–31.
11. ———, *Оценки тригонометрических сумм и их приложения*, Успехи матем. наук **13** (1958), no. 4 (82), 185–192.

12. О. В. Тырина, *Новая оценка тригонометрического интеграла И. М. Виноградова*, Известия Ак. Наук СССР, серия матем. **51** (1987), no. 2, 363–378, 447, English translation: O. V. Tyrina, *A new estimate for a trigonometric integral of I. M. Vinogradov*, Math. USSR-Izv. **30** (1988), no. 2, 337–351.
13. С. Б. Стечкин, *О средних значениях модуля тригонометрической сыммы*, Тр. Матем. ин-та им. В. А. Стеклова АН СССР **134** (1975), 283–309, 411, English translation: S. B. Stechkin, *Mean values of the modulus of a trigonometric sum*, Trudy Mat. Inst. Steklov. 134 (1975).
14. С. М. Воронин and А. А. Карацуба, *Дзета-функция Римана*, Москва , 1994, Translated from the Russian by Neal Koblitz; Walter de Gruyter & Co., Berlin, 1992.
15. E. Bombieri, *On Vinogradov's mean value theorem and Weyl sums*, Automorphic forms and analytic number theory (Montreal, PQ, 1989), Univ. Montréal, Montreal, QC, 1990, pp. 7–24.
16. C.-J. de la Vallée Poussin, *Sur la fonction $\zeta(s)$ de Riemann et le nombres des nombres premiers inférieurs à une limite donnée*, Mém. Couronnés et Autres Mém. Publ. Acad. Roy. Sci. des lettres Beaux-Arts Belg. **59** (1899–1900), 1–74.
17. K. B. Ford, *New estimates for mean values of Weyl sums*, Internat. Math. Res. Notices (1995), 155–171.
18. ———, *Vinogradov's integral and bounds for the Riemann zeta function*, Proc. London Math. Soc. (2002).
19. W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), no. 3, 529–551.
20. ———, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.
21. D. R. Heath-Brown, *The fractional part of αn^k*, Mathematika **35** (1988), no. 1, 28–37.
22. A. J. Hildebrand and G. Tenenbaum, *The distribution of numbers without large prime factors*, J. Théor. Nombres Bordeaux **5** (1993), 411–484.
23. L.-K. Hua, *An improvement of Vinogradov's mean-value theorem and several applications*, Quart. J. Math., Oxford Ser. **20** (1949), 48–61.
24. ———, *Additive theory of prime numbers*, American Mathematical Society, Providence, R.I., 1965, Translated from the 1957 Chinese edition.
25. Aleksandar Ivić, *The Riemann zeta-function*, John Wiley & Sons Inc., New York, 1985.
26. J. E. Littlewood, *Researches in the theory of the Riemann ζ-function*, Proc. London Math. Soc. (2) **20** (1922), XXII–XXVIII, Records, Feb. 10, 1921.
27. H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, American Mathematical Society, Providence, R.I., 1994.
28. H.-E. Richert, *Zur Abschätzung der Riemannschen Zetafunktion in der Nähe der Vertikalen* $\sigma = 1$, Math. Ann. **169** (1967), 97–101.
29. N. Tchudakoff, *On the functions $\zeta(s)$ and $\pi(x)$*, Comptes Rendus (Doklady) l'Académie des Sciences de l'URSS, n.s. **21** (1938), 421–422.
30. E. C. Titchmarsh, *The theory of the Riemann zeta-function*, second ed., The Clarendon Press Oxford University Press, New York, 1986, Edited and with a preface by D. R. Heath-Brown.
31. R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), no. 1-2, 1–71.
32. ———, *A new iterative method in Waring's problem. II*, J. London Math. Soc. (2) **39** (1989), no. 2, 219–230.
33. ———, *The Hardy-Littlewood method*, second ed., Cambridge University Press, Cambridge, 1997.
34. R. C. Vaughan and T. D. Wooley, *Waring's problem: a survey*, Proceedings of the Millennial Conference on Number Theory, Urbana, Illinois, 2000 (2002).
35. T. D. Wooley, *Large improvements in Waring's problem*, Ann. of Math. (2) **135** (1992), no. 1, 131–164.
36. ———, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), no. 2, 379–399.
37. ———, *The application of a new mean value theorem to the fractional parts of polynomials*, Acta Arith. **65** (1993), no. 2, 163–179.
38. ———, *Quasi-diagonal behaviour in certain mean value theorems of additive number theory*, J. Amer. Math. Soc. **7** (1994), no. 1, 221–245.
39. ———, *Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), no. 3, 421–451.

40. ______, *New estimates for smooth Weyl sums*, J. London Math. Soc. (2) **51** (1995), no. 1, 1–13.
41. ______, *New estimates for Weyl sums*, Quart. J. Math. Oxford Ser. (2) **46** (1995), no. 181, 119–127.
42. ______, *Some remarks on Vinogradov's mean value theorem and Tarry's problem*, Monatsh. Math. **122** (1996), no. 3, 265–273.
43. ______, *On exponential sums over smooth numbers*, J. Reine Angew. Math. **488** (1997), 79–140.

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA