A New k-th Derivative Estimate for Exponential Sums via Vinogradov's Mean Value

D.R. Heath-Brown
Mathematical Institute, Oxford

In celebration of the 125th anniversary of the birth of Ivan Matveevich Vinogradov

1 Introduction

The familiar van der Corput k-th derivative estimate for exponential sums (Titchmarsh [10, Theorems 5.9, 5.11, & 5.13], for example), may be stated as follows. Let $k \geq 2$ be an integer, and suppose that $f(x) : [0, N] \to \mathbb{R}$ has continuous derivatives of order up to k on $(0, N)$. Suppose further that $0 < \lambda_k \leq f^{(k)}(x) \leq A\lambda_k$ on $(0, N)$. Then

$$\sum_{n \leq N} e(f(n)) \ll A^{2^{2-k}} N\lambda_k^{1/(2^k-2)} + N^{1-2^{2-k}}\lambda_k^{-1/(2^k-2)}, \tag{1}$$

with an implied constant independent of k. One usually chooses k so that the first term dominates, and one often has $A^{2^{2-k}} \ll 1$, so that the bound is merely $O(N\lambda_k^{1/(2^k-2)})$. Clearly one can only get a non-trivial bound when $\lambda_k < 1$. A typical application is the series of estimates

$$\zeta(\sigma + it) \ll t^{1/(2^k-2)} \log t, \quad \left(\sigma = 1 - \frac{k}{2^k - 2},\ t \geq 2\right)$$

for $k = 2, 3, \ldots$. Again the implied constant is independent of k.

One can improve on the standard k-th derivative bound somewhat. Thus Robert and Sargos [7] show roughly that if $k = 4$ then

$$\sum_{n \leq N} e(f(n)) \ll_\varepsilon N^\varepsilon (N\lambda_4^{1/13} + \lambda_4^{-7/13}),$$

for any $\varepsilon > 0$. In the corresponding version of (1) one would have a term $N\lambda_4^{1/14}$ in place of $N\lambda_4^{1/13}$. Similarly for $k = 8$ and 9, Sargos [9, Theorems 3 & 4] gives bounds

$$\sum_{n \leq N} e(f(n)) \ll_\varepsilon N^\varepsilon (N\lambda_8^{1/204} + \lambda_8^{-95/204}),$$

and

$$\sum_{n \leq N} e(f(n)) \ll_\varepsilon N^\varepsilon (N\lambda_9^{7/2640} + \lambda_9^{-1001/2640}),$$

respectively. Here the exponents 1/204 and 7/2640 should be compared with the values 1/254 and 1/510 produced by (1).

There are also results in which the second term in the van der Corput bound is improved. Thus when $k = 3$ the bound (1) reduces to an estimate $O(N\lambda_3^{1/6})$ for $\lambda_3 \geq N^{-3/2}$. However it was shown independently by Gritsenko [3] and Sargos [8] that the weaker hypothesis $\lambda_3 \geq N^{-2}$ suffices.

There are quite different approaches to exponential sums, using estimates for the Vinogradov mean value integral

$$J_{s,l}(P) = \int_0^1 \ldots \int_0^1 \left| \sum_{n \leq P} e(\alpha_1 n + \ldots + \alpha_l n^l) \right|^{2s} d\boldsymbol{\alpha}, \tag{2}$$

see Vinogradov [11], [12], and Korobov [4], amongst others. The first of these methods is described by Titchmarsh [10, Chapter 6] for example. The Vinogradov-Korobov machinery has been used by Ford [2, Theorem 2] to show that

$$\sum_{N < n \leq 2N} n^{-it} \ll N^{1-1/134k^2} \tag{3}$$

for $N^k \geq t \geq 2$. (Ford's result is somewhat more precise, and more general.) One may think of this as corresponding very roughly to a bound of the form (1) with first term $N\lambda_k^{1/134k^2}$.

A slightly refined version of the original method of Vinogradov [11] coupled with new estimates for the Vinogradov mean value integral, leads to distinctly stronger bounds. For example, Wooley [13, Theorem 1.2] gives

$$J_{s,l}(P) \ll_{\varepsilon,l} P^{2s-l(l+1)/2+\varepsilon} \quad (s \geq l(l-1),$$

and Robert [6, Theorem 10] used this to show that if $k \geq 4$ then

$$\sum_{n \leq N} e(f(n)) \ll_{A,k,\varepsilon} N^{1+\varepsilon}(\lambda_k^{1/2(k-1)(k-2)} + N^{-1/2(k-1)(k-2)})$$

for $N \geq \lambda_k^{-(k-1)/(2k-3)}$. This is a remarkable improvement on the classical k-th derivative estimate. The exponent of λ_k is better than $1/(2^k - 2)$ for all $k \geq 4$, and decreases quadratically rather than exponentially.

The purpose of this paper is to further refine the original method of Vinogradov [11] and to input the very recent optimal bounds for the Vinogradov mean value integral, due to Wooley [13] (for $l = 3$), and to Bourgain, Demeter and Guth [1] (for $l \geq 4$). These theorems show that

$$J_{s,l}(P) \ll_{\varepsilon,l} P^{2s-l(l+1)/2+\varepsilon} \quad (s \geq \tfrac{1}{2}l(l+1),\ l \geq 1), \tag{4}$$

the cases $l = 1$ and $l = 2$ being elementary. The range for s is optimal, and it is this feature that represents the dramatic culmination of many previous works over the past 80 years. Unfortunately neither result gives an explicit dependence on l and s, nor gives an explicit form for the factor P^ε. Results prior to the advent of Wooley's efficient congruencing method had required s to be larger, but had given an explicit dependence on l. Thus for example, Ford [2, Theorem 3] implies in particular that

$$J_{s,l}(P) \ll l^{13l^3/4} P^{2s-l(l+1)/2+l^2/1000} \quad (s \geq \tfrac{5}{2}l^2,\ l \geq 129),$$

in which one has an additional term $l^2/1000$ in the exponent, and more restrictive conditions on l and s. An important application of bounds for Weyl sums is to the zero-free region for $\zeta(s)$, as described by Ford. However for this it is crucial to have a suitable dependence on the parameter l, so that the new result of Bourgain, Demeter and Guth is not applicable.

Our first result gives a new k-th derivative estimate

Theorem 1 *Let $k \geq 3$ be an integer, and suppose that $f(x) : [0, N] \to \mathbb{R}$ has continuous derivatives of order up to k on $(0, N)$. Suppose further that*

$$0 < \lambda_k \leq f^{(k)}(x) \leq A\lambda_k, \quad x \in (0, N).$$

Then

$$\sum_{n \leq N} e(f(n)) \ll_{A,k,\varepsilon} N^{1+\varepsilon}(\lambda_k^{1/k(k-1)} + N^{-1/k(k-1)} + N^{-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}).$$

If one thinks of $N\lambda_k^{1/k(k-1)}$ as being the leading term here, then one needs to compare the exponent $1/k(k-1)$ with the corresponding exponent $1/(2^k-2)$ in (1). These agree for $k = 3$, but for larger values of k the new exponent tends to zero far more slowly than the old one. It may perhaps be something of a surprise that an analysis via Vinogradov's mean value integral reproduces the same term $N\lambda_3^{1/6}$ as in the classical third-derivative estimate.

We should emphasize that the stength of Theorem 1 comes almost entirely from the new bound (4). One could have injected (4) into the method of Robert [6], to produce an estimate with the same terms $\lambda_k^{1/k(k-1)} + N^{-1/k(k-1)}$ as in Theorem 1, but valid only for $N \geq \lambda_k^{-(k-1)/(2k-3)}$. Our result, incorporating a slightly better way of using the Vinogradov mean value, gives the terms $\lambda_k^{1/k(k-1)} + N^{-1/k(k-1)}$ in the substantially longer range $N \geq \lambda_k^{-2/k}$. However for our application to Theorems 2–5 below, Robert's range would have been very nearly sufficient.

The secondary terms in the bound given by Theorem 1 are somewhat awkward. The classical estimate (1) leads easily to an exponent pair,

$$\left(\frac{1}{2^k-2}, \frac{2^k-k-1}{2^k-2}\right)$$

in which the term $N^{1-2^{2-k}}\lambda_k^{-1/(2^k-2)}$ has no effect. However the situation with Theorem 1 is more complicated. None the less we are able to produce a series of new exponent pairs.

Before stating the result we remind the reader of the necessary background. Let s and c be positive constants, and let $\mathcal{F}(s, c)$ be the set of quadruples (N, I, f, y) where $y \geq N^s$ are positive real numbers, I is a subinterval of $(N, 2N]$, and f is an infinitely differentiable function on I, with

$$\left| f^{(n+1)}(x) - \frac{d^n}{dx^n}(yx^{-s}) \right| \leq c \left| \frac{d^n}{dx^n}(yx^{-s}) \right|$$

for $x \in I$, for all $n \geq 0$. We then say that (p, q) is an exponent pair, if p and q lie in the range $0 \leq p \leq \frac{1}{2} \leq q \leq 1$, and for each s there is a corresponding $c = c(p, q, s) > 0$ such that

$$\sum_{n \in I} e(f(n)) \ll_{p,q,s} (yN^{-s})^p N^q,$$

uniformly for all quadruples $(N, I, f, y) \in \mathcal{F}(s, c)$.

We then have the following.

Theorem 2 *For any integer $k \geq 3$ and any real $\varepsilon > 0$ there is an exponent pair given by*

$$p = \frac{2}{(k-1)^2(k+2)}, \tag{5}$$

and

$$q = \frac{k^3 + k^2 - 5k + 2}{k(k-1)(k+2)} + \varepsilon = 1 - \frac{3k - 2}{k(k-1)(k+2)} + \varepsilon. \tag{6}$$

In fact we are able to handle a much weaker condition on f. Let $\mathcal{A} = (a_k)_3^\infty$ and $\mathcal{B} = (b_k)_3^\infty$ be sequences of positive real numbers, and let $\mathcal{G}(\mathcal{A}, \mathcal{B})$ be the set of quadruples (N, I, g, T) where $T \geq N$ are positive real numbers, I is a subinterval of $(N, 2N]$, and g is an infinitely differentiable function on I, with

$$a_k T N^{-k} \leq \left| g^{(k)}(x) \right| \leq b_k T N^{-k}$$

for $x \in I$, and for all $k \geq 3$. We then have the following.

Theorem 3 *For any integer $k \geq 3$ and any real $\varepsilon > 0$, let p and q be given by (5) and (6). Then*

$$\sum_{n \in I} e(g(n)) \ll_{k,\varepsilon,\mathcal{A},\mathcal{B}} (TN^{-1})^p N^q,$$

uniformly for $(N, I, g, T) \in \mathcal{G}(\mathcal{A}, \mathcal{B})$.

If $(N, I, g, y) \in \mathcal{F}(s, \frac{1}{4})$, then $(N, I, g, yN^{1-s}) \in \mathcal{G}(\mathcal{A}, \mathcal{B})$ with

$$a_k = \frac{3 \times 2^{1-2-k}}{4s(s+1)\ldots(s+k-2)}, \quad b_k = \frac{5}{4s(s+1)\ldots(s+k-2)}.$$

The sequences $\mathcal{A}$ and $\mathcal{B}$ depend only on s, and we immediately see that Theorem 2 follows from Theorem 3.

We next present a slightly weaker version of Theorem 3, which is somewhat more immediately intelligible. It will be convenient to write $T = N^\tau$.

Theorem 4 *Let sequences $\mathcal{A}$ and $\mathcal{B}$, and a real number $\varepsilon > 0$ be given, then*

$$\sum_{n \in I} e(g(n)) \ll_{\varepsilon,\mathcal{A},\mathcal{B}} N^{1-49/(80\tau^2)+\varepsilon},$$

uniformly for quadruples $(N, I, g, T) \in \mathcal{G}(\mathcal{A}, \mathcal{B})$ with $N \leq T^{1/2}$.

The constant 49/80 arises from the use of an exponent pair

$$(\tfrac{1}{20}\,,\ \tfrac{33}{40}) = A^2BA^2B(0,1)$$

when $\tau = \frac{7}{2}$. One could improve the constant slightly by employing a better exponent pair. As will be clear from the proof, the constant $\frac{49}{80}$ may be replaced by $1-\delta$ for any small $\delta > 0$, if we restrict to sufficiently large values $\tau \geq \tau(\delta)$.

As an example of Theorem 4, if $t \geq 2$ we find that

$$\sum_{n\in I} n^{-it} \ll_\varepsilon N^{1-49/80\tau^2+\varepsilon}, \tag{7}$$

for $\tau = (\log t)/(\log N) \geq 2$. This should be compared with (3). Using (7) we produce the following result.

Theorem 5 *Let $\kappa = \frac{8}{63}\sqrt{15} = 0.4918\ldots$. Then for any fixed $\varepsilon > 0$ we have*

$$\zeta(\sigma+it) \ll_\varepsilon t^{\kappa(1-\sigma)^{3/2}+\varepsilon} \tag{8}$$

uniformly for $t \geq 1$ and $\frac{1}{2} \leq \sigma \leq 1$. Moreover we have

$$\zeta(\sigma+it) \ll_\varepsilon t^{\frac{1}{2}(1-\sigma)^{3/2}+\varepsilon} \tag{9}$$

uniformly for $t \geq 1$ and $0 \leq \sigma \leq 1$.

One sees from the proof that κ may be reduced to $2/\sqrt{27}+\delta = 0.3849\ldots$ for any small $\delta > 0$, if we restrict σ to a suitably small range $\sigma(\delta) \leq \sigma \leq 1$. The corresponding result in the work of Ford [2, Theorem 1] states that

$$|\zeta(\sigma+it)| \leq 76.2t^{4.45(1-\sigma)^{3/2}}(\log t)^{2/3}$$

for $t \geq 3$ and $\frac{1}{2} \leq \sigma \leq 1$. Thus we have reduced the constant 4.45 to $0.4918\ldots$. Unfortunately our result yields no useful information when σ tends to 1, which is a critical situation in many applications. Moreover we do not have the explicit order constant that Ford finds.

As Ford explains, there are a number of interesting corollaries, for which we merely have to replace the constant $B = 4.45$ by $B = 0.492$ in the arguments given in [2, Pages 566 and 567]. We can feed our bound into the zero-density theorem of Montgomery [5, Theorem 12.3] (with $1-\alpha = 4.93(1-\sigma)$ as used by Ford [2, Page 566]) to give the following.

Corollary 1 *We have*

$$N(\sigma,T) \ll_\varepsilon T^{6.42(1-\sigma)^{3/2}+\varepsilon}$$

for $\frac{9}{10} \leq \sigma \leq 1$.

For moments of the Riemann Zeta-function we have:

Corollary 2 *For any positive integer k one has*

$$\int_0^T |\zeta(\sigma+it)|^{2k}dt \sim T\sum_1^\infty d_k(n)^2 n^{-2\sigma},$$

as $t \to \infty$, for any fixed $\sigma \geq 1 - 0.534k^{-2/3}$.

For the generalized divisor problem we have:

Corollary 3 *For any positive integer k the error term $\Delta_k(x)$ in the generalized divisor problem satisfies*

$$\Delta_k(x) \ll_k x^{1-0.849k^{-2/3}}.$$

In Section 2 we will reduce the proof of Theorem 1 to a two-variable counting problem involving fractional parts of the derivatives $f^{(j)}(n)$. Section 3 shows how this counting problem is tackled, and finally Section 4 completes the proof of our theorems.

Acknowledgement. This work was supported by EPSRC grant number EP/K021132X/1

2 Initial Steps

Our goal in the first stage of the proof is to estimate the sum

$$\Sigma = \sum_{n \le N} e(f(n))$$

in terms of $J_{s,l}(P)$, together with, a counting function involving the fractional parts of numbers of the form $f^{(j)}(n)/j!$.

Lemma 1 *Let $k \ge 2$ be an integer, and suppose that $f(x) : [0, N] \to \mathbb{R}$ has continuous derivatives of order up to k on $(0, N)$. Suppose further that*

$$0 < \lambda_k \le f^{(k)}(x) \le A\lambda_k, \quad x \in (0, N),$$

and that $A\lambda_k \le \frac{1}{4}$. Then

$$\Sigma \ll H + k^2 N^{1-1/s} \mathcal{N}^{1/2s} \left\{ H^{-2s+k(k-1)/2} J_{s,k-1}(H) \right\}^{1/2s},$$

where $H = [(A\lambda_k)^{-1/k}]$ and

$$\mathcal{N} = \# \left\{ m, n \le N : \left|\left| \frac{f^{(j)}(m)}{j!} - \frac{f^{(j)}(n)}{j!} \right|\right| \le 2H^{-j} \text{ for } 1 \le j \le k-1 \right\}.$$

If $J_{s,k-1}(H) \ll_{\varepsilon,k} H^{2s-k(k-1)/2+\varepsilon}$ as in (4) the estimate in the lemma reduces to

$$\Sigma \ll_{\varepsilon,k} H + N^{1-1/s+\varepsilon} \mathcal{N}^{1/2s}. \tag{10}$$

Here we would want to choose s to be as small as possible, and since we are taking $l = k-1$ this means that we will have $s = k(k-1)/2$.

The lemma is clearly trivial if $H \ge N$, and we may therefore suppose for the proof that $H \le N$. For any positive integer $H \le N$ we will have

$$H\Sigma = \sum_{h \le H} \sum_{-h < n \le N-h} e(f(n+h)) = \sum_{h \le H} \sum_{1 \le n \le N-H} e(f(n+h)) + O(H^2),$$

so that

$$\Sigma = H^{-1} \sum_{n \le N-H} \sum_{h \le H} e(f(n+h)) + O(H). \tag{11}$$

We proceed to approximate $f(n+h)$ by the polynomial

$$f_n(h) := f(n) + f'(n)h + \ldots + \frac{f^{(k-1)}(n)}{(k-1)!}h^{k-1}.$$

To do this we set $g_n(x) = f(n+x) - f_n(x)$ and use summation by parts to obtain the bound

$$\sum_{h \le H} e(f(n+h)) \ll |S_n(H)| + \int_0^H |S_n(x) g_n'(x)| dx,$$

where we have written

$$S_n(x) = \sum_{h \le x} e(f_n(h))$$

for convenience.

If $0 \le x \le H$ we may use Taylor's Theorem with Lagrange's form of the remainder to show that

$$f'(n+x) = f_n'(x) + \frac{f^{(k)}(\xi)}{(k-1)!}x^{k-1}$$

for some $\xi \in (n, n+x) \subseteq (0, N)$. It follows that

$$g_n'(x) \ll A\lambda_k H^{k-1}$$

on $[0, H]$. With the choice $H = [(A\lambda_k)^{-1/k}]$ we find that

$$\sum_{h \le H} e(f(n+h)) \ll |S_n(H)| + H^{-1} \int_0^H |S_n(x)| dx.$$

The bound (11) now yields

$$\Sigma \ll H + H^{-1} \sum_{n \le N-H} |S_n(H)| + H^{-2} \int_0^H \left\{ \sum_{n \le N-H} |S_n(x)| \right\} dx.$$

It then follows that there is a positive integer $H_0 \le H$ such that

$$\Sigma \ll H + H^{-1} \sum_{n \le N-H} |S_n(H_0)|. \tag{12}$$

Now suppose that $\boldsymbol{\alpha} \in [0,1]^{k-1}$ and

$$||f^{(j)}(n)/j! - \alpha_j|| \le H^{-j} \text{ for } 1 \le j \le k-1 \tag{13}$$

where

$$||\theta|| = \min_{n \in \mathbb{Z}} |\theta - n|$$

as usual. We proceed to replace $f_n(h)$ by

$$f(h; \boldsymbol{\alpha}) = \alpha_1 h + \ldots + \alpha_{k-1} h^{k-1}$$

as follows. Firstly we remove the constant term $f(n)$ from $f_n(h)$. This has no effect on $|S_n(H_0)|$. Next, we replace each coefficient $f^{(j)}(n)/j!$ by c_j, say, with $f^{(j)}(n)/j!-c_j \in \mathbb{Z}$, so that $|c_j-\alpha_j| \le H^{-j}$, and denote the resulting polynomial by $f_n^*(h)$. If we write

$$S_n^*(H_0) = \sum_{h\le H_0} e(f_n^*(h))$$

then clearly $|S_n(H_0)| = |S_n^*(H_0)|$. Moreover

$$\frac{d}{dx}\left(f(x;\boldsymbol{\alpha}) - f_n^*(x)\right) \ll k^2 \max_{j\le k-1} |c_j-\alpha_j|H^{j-1} \ll k^2H^{-1}.$$

It therefore follows on summing by parts that

$$S_n^*(H_0) \ll |S(H_0;\boldsymbol{\alpha})| + k^2H^{-1}\int_0^{H_0} |S(x;\boldsymbol{\alpha})|dx,$$

where we have set

$$S(x;\boldsymbol{\alpha}) = \sum_{h\le x} e(f(h;\boldsymbol{\alpha})).$$

We may therefore conclude that

$$S_n^*(H_0) \ll 2^{-k}H^{k(k-1)/2}\left\{\int_{\boldsymbol{\alpha}} |S(H_0;\boldsymbol{\alpha})|d\boldsymbol{\alpha} + k^2H^{-1}\int_0^{H_0}\int_{\boldsymbol{\alpha}} |S(x;\boldsymbol{\alpha})|d\boldsymbol{\alpha}dx\right\},$$

where the integral over $\boldsymbol{\alpha}$ is for vectors in $[0,1]^{k-1}$ satisfying (13).

For each $\boldsymbol{\alpha} \in [0,1]^{k-1}$ we now define

$$\nu(\boldsymbol{\alpha}) = \#\{n \le N-H : ||f^{(j)}(n)/j! - \alpha_j|| \le H^{-j} \text{ for } 1 \le j \le k-1\}.$$

We then find that

$$\sum_{n\le N-H} |S_n(H_0)| \ll 2^{-k}H^{k(k-1)/2}\left\{I(H_0) + k^2H^{-1}\int_0^{H_0} I(x)dx\right\}, \qquad (14)$$

with

$$I(x) = \int_0^1 \cdots \int_0^1 |S(x;\boldsymbol{\alpha})|\nu(\boldsymbol{\alpha})d\boldsymbol{\alpha}.$$

We easily see that

$$\int_0^1 \cdots \int_0^1 \nu(\boldsymbol{\alpha})d\boldsymbol{\alpha} = 2^{k-1}H^{-k(k-1)/2}(N-H),$$

and that

$$\int_0^1 \cdots \int_0^1 \nu(\boldsymbol{\alpha})^2 d\boldsymbol{\alpha} \le 2^{k-1}H^{-k(k-1)/2}\mathcal{N},$$

where $\mathcal{N}$ is defined in Lemma 1. Moreover

$$\int_0^1 \cdots \int_0^1 |S(x;\boldsymbol{\alpha})|^{2s} d\boldsymbol{\alpha} = J_{s,k-1}(x)$$

in the notation of (2). Since $J_{s,k-1}(P)$ is non-decreasing in P this last integral may be bounded by $J_{s,k-1}(H)$.

Hence, by Hölder's inequality, for any positive integer s we have

$$I(x) \ll 2^k H^{-k(k-1)/2} N^{1-1/s} \mathcal{N}^{1/2s} \left\{ H^{k(k-1)/2} J_{s,k-1}(H) \right\}^{1/2s}.$$

Thus (14) yields

$$\sum_{n \le N-H} |S_n(H_0)| \ll k^2 N^{1-1/s} \mathcal{N}^{1/2s} \left\{ H^{k(k-1)/2} J_{s,k-1}(H) \right\}^{1/2s}$$

and (12) gives us

$$\Sigma \ll H + k^2 N^{1-1/s} \mathcal{N}^{1/2s} \left\{ H^{-2s+k(k-1)/2} J_{s,k-1}(H) \right\}^{1/2s}$$

as required.

3 The counting function $\mathcal{N}$

Naturally our next task is to bound $\mathcal{N}$. The original approach taken by Vinogradov, as described in Titchmarsh [10, Chapter 6], merely used an L^∞ bound for $\nu(\boldsymbol{\alpha})$. One discards all the information on $f^{(j)}(n)/j!$ for $j \le k-2$ and uses only the case $j = k-1$. One then employs a standard procedure given by the following trivial variant of [10, Lemma 6.11], for example.

Lemma 2 *Let N be a positive integer, and suppose that $g(x) : [0, N] \to \mathbb{R}$ has a continuous derivative on $(0, N)$. Suppose further that*

$$0 < \mu \le g'(x) \le A_0 \mu, \quad x \in (0, N).$$

Then

$$\#\{n \le N : ||g(n)|| \le \theta\} \ll (1 + A_0 \mu N)(1 + \mu^{-1}\theta).$$

We fix m and take

$$g(x) = \frac{f^{(k-1)}(x) - f^{(k-1)}(m)}{(k-1)!}$$

and $\mu = \lambda_k/(k-1)!$, $A_0 = A$. This leads to a bound

$$\mathcal{N} \ll (k-1)! N (1 + AN\lambda_k)(1 + H^{1-k}\lambda_k^{-1}).$$

Under the assumption $A\lambda_k \le \frac{1}{4}$ in Lemma 1 we have

$$H^{1-k}\lambda_k^{-1} \asymp (A\lambda_k)^{1-1/k}\lambda_k^{-1} = A(A\lambda_k)^{-1/k} \ge A \ge 1,$$

whence our bound produces

$$\mathcal{N} \ll A^2 (k-1)! N \lambda_k^{-1/k} (1 + N\lambda_k). \tag{15}$$

If one inserts this into (10) with $s = k(k-1)/2$ one gets an estimate

$$\begin{aligned} \Sigma &\ll_{\varepsilon,k} (A\lambda_k)^{-1/k} + N^{1-1/s+\varepsilon}\{A^2 N \lambda_k^{-1/k}(1 + N\lambda_k)\}^{1/2s} \\ &\ll_{\varepsilon,k} AN^{\varepsilon}\{\lambda_k^{-1/k} + N^{1-1/k(k-1)}\lambda_k^{-1/k^2(k-1)} + N\lambda_k^{1/k^2}\}. \end{aligned}$$

In fact the first term can be dropped, giving

$$\Sigma \ll_{\varepsilon,k} AN^{\varepsilon}\{N\lambda_k^{1/k^2} + N^{1-1/k(k-1)}\lambda_k^{-1/k^2(k-1)}\}. \tag{16}$$

To see this we note that we have

$$\Sigma \ll N^{1-1/k(k-1)}\lambda_k^{-1/k^2(k-1)}$$

trivially unless

$$N^{1-1/k(k-1)}\lambda_k^{-1/k^2(k-1)} \le N.$$

In this latter case however one sees that

$$\lambda_k^{-1/k} \le N^{1-1/k(k-1)}\lambda_k^{-1/k^2(k-1)}.$$

We may therefore regard (16) as being the result that Vinogradov's method achieves, given the results of Wooley [13] and Bourgain, Demeter and Guth [1]. It is already a remarkable improvement on (1), replacing the critical exponent $1/(2^k-2)$ by $1/k^2$. Thus, in appropriate circumstances, we get an improvement as soon as $k \ge 5$. Our goal in this section is to make the following small further sharpening in the estimation of $\mathcal{N}$.

Lemma 3 *When $k \ge 3$ we have*

$$\mathcal{N} \ll \big((k-1)!A\big)^4(N + \lambda_k N^2 + \lambda_k^{-2/k})\log N.$$

Apart from the term $\lambda_k^{-2/k}$, which is insignificant in applications, this represents an improvement of (15) by a factor $\ll_{A,k} \lambda_k^{1/k}$.

On the one hand our proof will use the fact that $\mathcal{N}$ is a counting function of two variables m and n. On the other we shall use information about both $f^{(k-1)}$ and $f^{(k-2)}$. The reader may find it slightly surprising in the light of this that our bound depends on λ_k only, and not on estimates for other derivatives $f^{(j)}$. The introduction of $\mathcal{N}$, and our procedure for estimating it, are the only really new aspects to this paper.

We begin our analysis by assuming that $k \ge 3$ and noting that $\mathcal{N}$ is at most

$$\mathcal{N}_1 = \#\left\{m, n \le N : \left|\left|\frac{f^{(j)}(m)}{j!} - \frac{f^{(j)}(n)}{j!}\right|\right| \le 2H^{-j} \text{ for } j = k-2, k-1\right\}.$$

We proceed to show that it suffices to consider pairs m, n of integers that are relatively close. It will be convenient to write $B = 4H^{2-k}$ and $C = 4H^{1-k}$ and to set

$$g_1(x) = \frac{f^{(k-2)}(x)}{(k-2)!}, \quad g_2(x) = \frac{f^{(k-1)}(x)}{(k-1)!}.$$

We also define the doubly-periodic function

$$\phi(x,y) = \max\left(1 - B^{-1}||x||, 0\right)\max\left(1 - C^{-1}||y||, 0\right),$$

so that

$$\mathcal{N}_1 \ll \sum_{m,n\le N} \phi\big(g_1(m) - g_1(n)\,,\, g_2(m) - g_2(n)\big).$$

The function $\phi(x,y)$ has an absolutely convergent Fourier series

$$\phi(x,y)=\sum_{r,s\in\mathbb{Z}} c_{r,s}e(rx+sy)$$

with non-negative coefficients

$$c_{r,s}=BC\left(\frac{\sin(\pi rB)\sin(\pi sC)}{\pi^2 rsBC}\right)^2.$$

Thus

$$\begin{aligned}\mathcal{N}_1 &\ll \sum_{r,s\in\mathbb{Z}} c_{r,s}\sum_{m,n\le N} e\big(r(g_1(m)-g_1(n))+s(g_2(m)-g_2(n))\big)\\ &= \sum_{r,s\in\mathbb{Z}} c_{r,s}\left|\sum_{n\le N} e\big(rg_1(n)+sg_2(n)\big)\right|^2.\end{aligned}$$

Let K be a positive integer parameter, to be chosen later. We proceed to partition the range $(0,N]$ into K intervals $I_i=(a_i,b_i]$ for $i\le K$, having integer endpoints, and length $b_i-a_i\le 1+N/K$. An application of Cauchy's inequality then yields

$$\begin{aligned}\mathcal{N}_1 &\ll K\sum_{i\le K}\sum_{r,s\in\mathbb{Z}} c_{r,s}\left|\sum_{n\in I_i} e\big(rg_1(n)+sg_2(n)\big)\right|^2\\ &= K\sum_{i\le K}\sum_{r,s\in\mathbb{Z}} c_{r,s}\sum_{m,n\in I_i} e\big(r(g_1(m)-g_1(n))+s(g_2(m)-g_2(n))\big)\\ &= K\sum_{i\le K}\sum_{m,n\in I_i}\phi\big(g_1(m)-g_1(n)\,,\,g_2(m)-g_2(n)\big)\\ &\le K\sum_{\substack{m,n\le N\\ |m-n|\le 1+N/K}}\phi\big(g_1(m)-g_1(n)\,,\,g_2(m)-g_2(n)\big).\end{aligned}$$

We may therefore conclude that $\mathcal{N}_1\ll K\mathcal{N}_2$, where $\mathcal{N}_2$ counts pairs of integers $m,n\le N$ with $|m-n|\le 1+N/K$ for which

$$\left|\left|\frac{f^{(j)}(m)}{j!}-\frac{f^{(j)}(n)}{j!}\right|\right|\le 4H^{-j}\text{ for } j=k-2,k-1.$$

If $|m-n|\le 1+N/K$ we will have

$$\left|\frac{f^{(k-1)}(m)}{(k-1)!}-\frac{f^{(k-1)}(n)}{(k-1)!}\right|\le\frac{|m-n|}{(k-1)!}\sup|f^{(k)}|\le A\lambda_k(1+N/K),$$

by the mean-value theorem. We will choose

$$K=1+[4A\lambda_k N],$$

so that

$$\left|\frac{f^{(k-1)}(m)}{(k-1)!}-\frac{f^{(k-1)}(n)}{(k-1)!}\right|\le\frac{1}{2},$$

in view of our assumption that $A\lambda_k \le \frac{1}{4}$. Thus if

$$\left|\left|\frac{f^{(k-1)}(m)}{(k-1)!} - \frac{f^{(k-1)}(n)}{(k-1)!}\right|\right| \le 4H^{1-k}$$

we must have

$$\left|\frac{f^{(k-1)}(m)}{(k-1)!} - \frac{f^{(k-1)}(n)}{(k-1)!}\right| \le 4H^{1-k}.$$

However the mean-value theorem also tells us that

$$\left|\frac{f^{(k-1)}(m)}{(k-1)!} - \frac{f^{(k-1)}(n)}{(k-1)!}\right| \ge \frac{|m-n|}{(k-1)!}\inf|f^{(k)}| \ge \lambda_k \frac{|m-n|}{(k-1)!}.$$

We therefore conclude that

$$|m-n| \le \frac{4(k-1)!}{\lambda_k H^{k-1}}$$

for any pair m, n counted by $\mathcal{N}_2$.

There are N pairs $m = n$ counted by $\mathcal{N}_2$. We consider the remaining pairs with $m > n$, the alternative case producing the same estimates by symmetry. Then $m = n + d$ with $1 \le d \le D$, where

$$D = \min\left(N\,,\,\left[\frac{4(k-1)!}{\lambda_k H^{k-1}}\right]\right).$$

For each available value of d we estimate the number of corresponding integers n via Lemma 2, taking

$$g(x) = \frac{f^{(k-2)}(x+d) - f^{(k-2)}(x)}{(k-2)!}.$$

Then

$$g'(x) = \frac{f^{(k-1)}(x+d) - f^{(k-1)}(x)}{(k-2)!},$$

so that

$$d\frac{\lambda_k}{(k-2)!} \le d\frac{\inf|f^{(k)}|}{(k-2)!} \le g'(x) \le d\frac{\sup|f^{(k)}|}{(k-2)!} \le d\frac{A\lambda_k}{(k-2)!},$$

by the mean-value theorem. We therefore apply the lemma with $\mu = \lambda_k d/(k-2)!$ and $A_0 = A$. This shows that each $d \ge 1$ contributes

$$\begin{aligned}
&\ll (k-2)!(1 + AN\lambda_k d)(1 + H^{2-k}\lambda_k^{-1}d^{-1}) \\
&\ll (k-2)!(1 + AN\lambda_k D)(D + H^{2-k}\lambda_k^{-1})d^{-1} \\
&\ll \big((k-1)!\big)^3 A(1 + NH^{1-k})H^{2-k}\lambda_k^{-1}d^{-1} \\
&\ll \big((k-1)!A\big)^3(1 + N\lambda_k^{1-1/k})\lambda_k^{-2/k}d^{-1}.
\end{aligned}$$

Summing for $d \le D$ we therefore find that

$$\mathcal{N}_2 \ll N + \big((k-1)!A\big)^3(1 + N\lambda_k^{1-1/k})\lambda_k^{-2/k}\log D.$$

Since $k \geq 3$, $\lambda_k \leq 1$ and $D \leq N$ this simplifies to give

$$\mathcal{N}_2 \ll \big((k-1)!A\big)^3 (N + \lambda_k^{-2/k}) \log N,$$

whence

$$\begin{aligned}\mathcal{N} \leq \mathcal{N}_1 &\ll K\mathcal{N}_2 \ll (1 + A\lambda_k N)\big((k-1)!A\big)^3 (N + \lambda_k^{-2/k}) \log N \\ &\ll \big((k-1)!A\big)^4 (N + \lambda_k N^2 + \lambda_k^{-2/k} + N\lambda_k^{1-2/k}) \log N.\end{aligned}$$

Since $k \geq 3$ and $\lambda_k \leq 1$ we have $N\lambda_k^{1-2/k} \leq N$, and Lemma 3 follows.

4 Proof of the Theorems

If we insert Lemma 3 into Lemma 1, and use the bound (4) with the choices $l = k-1, s = k(k-1)/2$, we see that

$$\Sigma \ll_{A,k,\varepsilon} N^{\varepsilon}(\lambda_k^{-1/k} + N^{1-1/k(k-1)} + N\lambda_k^{1/k(k-1)} + N^{1-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}).$$

The term $\lambda_k^{-1/k}$ may be omitted, since the resulting bound

$$\Sigma \ll_{A,k,\varepsilon} N^{\varepsilon}(N^{1-1/k(k-1)} + N\lambda_k^{1/k(k-1)} + N^{1-2/k(k-1)}\lambda_k^{-2/k^2(k-1)})$$

holds trivially when $N \leq N^{1-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}$, while

$$\lambda_k^{-1/k} \leq N^{1-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}$$

when $N \geq N^{1-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}$. This suffices for Theorem 1.

We turn next to Theorem 3. Suppose that $(N, I, g, T) \in \mathcal{G}(\mathcal{A}, \mathcal{B})$, and let I have end points N_0 and $N_0 + N_1$, so that $N_1 \leq N$. We apply Theorem 1 to the function $f(x) = g(N_0 + x)$, taking $\lambda_k = a_k T N^{-k}$ and $A = b_k/a_k$. (Since $f^{(k)}$ is differentiable it is continuous, and hence it cannot change sign if $|f^{(k)}(x)| \geq a_k T N^{-k} > 0$. Taking complex conjugates of our sum if necessary we may therefore assume that $f^{(k)}(x)$ is positive on I.) It follows that if $k \geq 3$ then

$$\begin{aligned}&\sum_{n \in I} e(g(n)) \\ &\quad \ll_{\varepsilon,k,\mathcal{A},\mathcal{B}} \quad N^{1+\varepsilon}(\lambda_k^{1/k(k-1)} + N^{-1/k(k-1)} + N^{-2/k(k-1)}\lambda_k^{-2/k^2(k-1)}) \\ &\quad \ll_{\varepsilon,k,\mathcal{A},\mathcal{B}} \quad N^{1+\varepsilon}(N^{-1/(k-1)}T^{1/k(k-1)} + N^{-1/k(k-1)} + T^{-2/k^2(k-1)}).\end{aligned}$$

We use the above bound for

$$\frac{(k-1)^2+1}{k} \leq \tau < \frac{k^2+1}{k+1}$$

where we define τ by $T = N^{\tau}$. For this range of τ we find that

$$\begin{aligned}&\max\left(\frac{\tau - k}{k(k-1)}, \frac{-1}{k(k-1)}, \frac{-2\tau}{k^2(k-1)}\right) \\ &\quad = \left\{\begin{array}{ll} -1/k(k-1), & ((k-1)^2+1)/k \leq \tau \leq k-1, \\ (\tau - k)/k(k-1), & k-1 \leq \tau < (k^2+1)/(k+1), \end{array}\right. \qquad (17) \\ &\quad \leq \quad A_k\tau + B_k,\end{aligned}$$

where the coefficients A_k and B_k are chosen so that

$$A_k \frac{(k-1)^2+1}{k} + B_k = \frac{-1}{k(k-1)}$$

and

$$A_k \frac{k^2+1}{k+1} + B_k = \frac{(k^2+1)/(k+1)-k}{k(k-1)} = \frac{-1}{k(k+1)}.$$

One then calculates that

$$A_k = \frac{2}{(k-1)^2(k+2)} \quad \text{and} \quad B_k = -\frac{3k^2-3k+2}{k(k-1)^2(k+2)}.$$

If we now define $\phi(\tau) : [2, \infty) \to \mathbb{R}$ by taking $\phi(\tau) = A_k\tau + B_k$ on

$$\left[\frac{(k-1)^2+1}{k}, \frac{k^2+1}{k+1}\right)$$

for each integer $k \geq 3$, we conclude that

$$\sum_{n \in I} e(g(n)) \ll_{\varepsilon, \tau_0, \mathcal{A}, \mathcal{B}} N^{1+\phi(\tau)+\varepsilon}, \tag{18}$$

uniformly for $2 \leq \tau \leq \tau_0$. The function ϕ is continuous, and since the coefficients A_k are monotonic decreasing ϕ is also convex. It follows that $\phi(\tau) \leq A_k\tau + B_k$ for any $\tau \in [2, \infty)$ and any $k \geq 3$. Thus

$$\sum_{n \in I} e(g(n)) \ll_{\varepsilon, \tau_0, \mathcal{A}, \mathcal{B}} N^{1+B_k+\varepsilon} T^{A_k} = (TN^{-1})^p N^q,$$

with p, q given by (5) and (6). As before, this is uniform in any finite range $2 \leq \tau \leq \tau_0$. However if we set $\tau_0 = 1 + (1-q)/p$ then τ_0 will depend on ε and k alone. Moreover, if $\tau \geq \tau_0$ then we trivially have

$$\sum_{n \in I} e(g(n)) \ll N \leq (TN^{-1})^p N^q.$$

Finally, if $\tau \leq 2$ we use the well known exponent pair $(\frac{1}{6}, \frac{2}{3})$ to show that

$$\sum_{n \in I} e(g(n)) \ll T^{1/6} N^{1/2}.$$

When $k \geq 3$ one easily verifies that $q \geq p + 1/2$ and $p + q \geq 5/6$ for the values (5) and (6), whence $T^{1/6}N^{1/2} \leq T^p N^{q-p}$ for $N \geq T^{1/2}$. It then follows that

$$\sum_{n \in I} e(g(n)) \ll T^{1/6} N^{1/2} \leq T^p N^{q-p}$$

for the remaining range $1 \leq \tau \leq 2$. This completes the proof of Theorem 3.

We move now to the proof of Theorem 4. Let $\tau_0 = \sqrt{49/80\varepsilon^2}$. Then if $\tau \geq \tau_0$ we will trivially have

$$\sum_{n \in I} e(g(n)) \ll N \leq N^{1-49/80\tau^2+\varepsilon}.$$

When $\tau \le \tau_0$ we begin by handling the range $\frac{13}{3} \le \tau \le \tau_0$, for which we claim that $\phi(\tau) \le -49/80\tau^2$. This will clearly suffice, in view of the estimate (18). Since $\phi(\tau)$ is piecewise linear, while the function $-49/80\tau^2$ is convex, it suffices to verify that $\phi(\tau) \le -49/80\tau^2$ at each of the points $\tau = (k^2+1)/(k+1)$, for $k \ge 5$. This condition is equivalent to

$$\frac{(k^2+1)^2}{k(k+1)^3} \ge \frac{49}{80}.$$

However the fraction on the right is increasing for $k \ge 5$, and takes the value $169/270 > 49/80$ at $k = 5$.

When $\frac{7}{2} \le \tau \le \frac{13}{3}$ we will use the bounds

$$\sum_{n\in I} e(g(n)) \ll_\varepsilon N^{1-1/20+\varepsilon}, \quad (\tfrac{7}{2} \le \tau \le 4)$$

and

$$\sum_{n\in I} e(g(n)) \ll_\varepsilon N^{1-(5-\tau)/20+\varepsilon}, \quad (4 \le \tau \le \tfrac{13}{3})$$

which come from the case $k = 5$ of (17). Note that the first of these is valid in the longer range $\frac{17}{5} \le \tau \le 4$, but we shall only use it when $\frac{7}{2} \le \tau \le 4$. We therefore need to verify that $-1/20 \le -49/80\tau^2$ for $\frac{7}{2} \le \tau \le 4$ and that $-(5-\tau)/20 \le -49/80\tau^2$ for $4 \le \tau \le \frac{13}{3}$. This is routine, but we observe that we have equality at $\tau = \frac{7}{2}$.

We next consider the case in which $\frac{59}{22} \le \tau \le \frac{7}{2}$, for which we use the bound

$$\sum_{n\in I} e(g(n)) \ll (T/N)^{1/20} N^{33/40} = N^{1+(2\tau-9)/40}$$

corresponding to the exponent pair $(\frac{1}{20}, \frac{33}{40})$. (This pair is $A^2BA^2B(0,1)$ in the usual notation, see Titchmarsh [10, §5.20], for example.) Again, it is routine to check that

$$\frac{2\tau-9}{40} \le -\frac{49}{80\tau^2}, \quad (\frac{59}{22} \le \tau \le \frac{7}{2}).$$

Finally we examine the range $2 \le \tau \le \frac{59}{22}$, and here we use the bound

$$\sum_{n\in I} e(g(n)) \ll (T/N)^{1/9} N^{13/18} = N^{1+(2\tau-7)/18}$$

corresponding to the exponent pair $(\frac{1}{9}, \frac{13}{18})$. (This pair is $ABA^2B(0,1)$ in the usual notation, see Titchmarsh [10, §5.20], for example.) Another routine check shows that

$$\frac{2\tau-7}{18} \le -\frac{49}{80\tau^2}, \quad (2 \le \tau \le \frac{59}{22}),$$

thereby completing the proof of Theorem 4.

We turn now to Theorem 5. If $\tau \ge 2$ we may use (7) along with a partial summation to obtain

$$\sum_{n\in J} n^{-\sigma-it} \ll_\varepsilon N^{1-49/80\tau^2-\sigma+\varepsilon} \le t^{(1-\sigma)\tau^{-1}-\frac{49}{80}\tau^{-3}+\varepsilon/2}$$

for any $\sigma \in [\frac{1}{2}, 1]$, and for any interval $J \subseteq (N, 2N]$. As a function of $\tau \in (0, \infty)$ the exponent of t is maximal at

$$\tau = \sqrt{\frac{147}{80(1-\sigma)}},$$

whence

$$\sum_{n \in J} n^{-\sigma - it} \ll_\varepsilon t^{\kappa(1-\sigma)^{3/2} + \varepsilon/2}.$$

Using a dyadic subdivision of $(0, N]$ we therefore have

$$\sum_{n \le N} n^{-\sigma - it} \ll_\varepsilon t^{\kappa(1-\sigma)^{3/2} + 3\varepsilon/4}$$

for any $N \le t^{1/2}$. A further summation by parts then shows that

$$\sum_{n \le M} n^{-1+\sigma - it} \ll_\varepsilon M^{2\sigma - 1} t^{\kappa(1-\sigma)^{3/2} + \varepsilon} \ll_\varepsilon t^{\sigma - \frac{1}{2} + \kappa(1-\sigma)^{3/2} + \varepsilon}$$

for any $M \le t^{1/2}$. The required bound (8) then follows from the approximate functional equation for $\zeta(s)$.

The bound (9) follows from (8) when $\frac{1}{2} \le \sigma \le 1$, since $\kappa < \frac{1}{2}$. For the remaining range we use the functional equation, which shows that

$$\zeta(\sigma + it) \ll t^{\frac{1}{2} - \sigma} |\zeta(1 - \sigma + it)| \ll_\varepsilon t^{\frac{1}{2} - \sigma + \frac{1}{2}\sigma^{3/2} + \varepsilon}.$$

However one can readily verify that

$$\frac{1}{2} - \sigma + \frac{\sigma^{3/2}}{2} \le \frac{(1-\sigma)^{3/2}}{2}$$

for $0 \le \sigma \le \frac{1}{2}$, which completes the proof of Theorem 5.

References

[1] J. Bourgain, C. Demeter and L. Guth, Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, `arXiv:1512.01565`.

[2] K. Ford, Vinogradov's integral and bounds for the Riemann zeta function, *Proc. London Math. Soc. (3)*, 85 (2002), no. 3, 565–633.

[3] S.A. Gritsenko, On estimates for trigonometric sums with respect to the third derivative, *Mat. Zametki*, 60 (1996), 383–389.

[4] N.M. Korobov, Estimates of trigonometric sums and their applications, *Uspehi Mat. Nauk*, 13 (1958) no. 4 (82), 185–192.

[5] H.L. Montgomery, *Topics in multiplicative number theory,* Lecture Notes in Mathematics 227, (Springer, Berlin, 1971).

[6] O. Robert, On van der Corputs k-th derivative test for exponential sums, *Indag. Math. (N.S.)*, to appear.

[7] O. Robert and P. Sargos, A fourth derivative test for exponential sums, *Compositio Math.*, 130 (2002), no. 3, 275–292.

[8] P. Sargos, Points entiers au voisinage d'une courbe, sommes trigonométriques courtes et paires d'exposants, *Proc. London Math. Soc. (3)*, 70 (1995), 285–312.

[9] P. Sargos, An analog of van der Corput's A^4-process for exponential sums, *Acta Arith.*, 110 (2003), no. 3, 219–231.

[10] E.C. Titchmarsh, *The theory of the Riemann zeta-function*, Second edition, (Clarendon Press, Oxford University Press, New York, 1986).

[11] I.M. Vinogradov, New Estimates for Weyl Sums, *Doklady Nauk SSSR*, 8 (1935), no. 5, 195–198.

[12] I.M. Vinogradov, A new estimate of the function $\zeta(1+it)$, *Izv. Akad. Nauk SSSR. Ser. Mat.*, 22 (1958), 161–164.

[13] T.D. Wooley, The cubic case of the main conjecture in Vinogradov's mean value theorem, `arXiv:1401.3150`.

Mathematical Institute,
Radcliffe Observatory Quarter,
Woodstock Road,
Oxford
OX2 6GG
UK

rhb@maths.ox.ac.uk