

THE ASYMPTOTIC BEHAVIOUR OF THE MEAN-SQUARE OF FRACTIONAL PART SUMS

MANFRED KÜHLEITNER AND WERNER GEORG NOWAK

*Institut für Mathematik u. Ang. Stat., Universität für Bodenkultur, A-1180 Wien,
Austria (kleitner@mail.boku.ac.at; nowak@mail.boku.ac.at)*

(Received 16 April 1998)

Abstract In this article we consider sums $S(t) = \sum_n \psi(tf(n/t))$, where ψ denotes, essentially, the fractional part minus $\frac{1}{2}$, f is a C^4 -function with f'' non-vanishing, and summation is extended over an interval of order t . For the mean-square of $S(t)$, an asymptotic formula is established. If f is algebraic this can be sharpened by the indication of an error term.

Keywords: fractional parts; mean square; lattice points; exponential sums

AMS 1991 *Mathematics subject classification:* Primary 11P21, 11K99

1. Introduction and statement of results

Let f be a real-valued function defined on an interval $[a, b]$ with continuous derivatives up to order 4 and the property that f'' has no zero on $[a, b]$. Further, let ψ denote a row-of-teeth function satisfying

$$\left. \begin{aligned} \psi(w) &= w - [w] - \frac{1}{2}, & \text{for } w \notin \mathbb{Z}, \\ -\frac{1}{2} \leq \psi(w) &\leq \frac{1}{2}, & \text{for } w \in \mathbb{Z}. \end{aligned} \right\} \quad (1.1)$$

Finally, let t be a large real parameter. Then the objective of the present article will be to study the sum

$$S(t) = \sum_{\substack{at < n \leq bt \\ n \in \mathbb{Z}}} \psi\left(tf\left(\frac{n}{t}\right)\right).$$

It is clear that $S(t)$ is connected with the problem of counting the lattice points in the Euclidean (ξ, η) -plane between the curve

$$\frac{\eta}{t} = f\left(\frac{\xi}{t}\right)$$

and the ξ -axis (see also §2 of [15]). In classic times, upper estimates for sums of this type have been obtained by Vinogradov [20] and van der Corput [18, 19], who ultimately proved that

$$S(t) \ll t^{2/3-\delta_0}$$

with some (very small) $\delta_0 > 0$. Under the general conditions stated, there was no improvement until recently when Huxley deduced his ‘Discrete Hardy Littlewood Method’ from earlier ideas due to Bombieri and Iwaniec [1] and Iwaniec and Mozzochi [7]. Huxley’s sharpest version [5] contains the bound

$$S(t) \ll t^{46/73}(\log t)^{315/146} \tag{1.2}$$

(under the additional technical condition that $f^{(3)}$ has no zero on $[a, b]$). For a full account, the reader is referred to Huxley’s textbook [6]. We remark parenthetically that a more general sum which contains a second parameter $U \geq t$, namely

$$S(t, U) = \sum_{at < n \leq bt, n \in \mathbb{Z}} \psi\left(Uf\left(\frac{n}{t}\right)\right),$$

is important for rounding-error estimations in connection with classic numerical integration methods. This, too, has been worked out by Huxley [4]. The upper bounds he established depend in a complicated way on the relative size of U and t and are given in a table exceeding one printed page.

To establish a lower estimate for $S(t)$ and results on its mean-square, one needs one more condition which is most efficiently stated in the following geometric terms.

Definition 1.1. A smooth curve \mathcal{C} is said to satisfy the *tangent condition* if none of the tangents of \mathcal{C} contains the origin.

In the first part of this investigation [15], the upper estimate (1.2) has been complemented by the lower bound

$$\limsup_{t \rightarrow \infty} \left(\frac{-\operatorname{sgn}(f'')S(t)}{t^{1/2}(\log t)^{1/4}} \right) > 0, \tag{1.3}$$

and the mean-square estimate

$$\int_0^T (S(t))^2 dt \ll T^2, \tag{1.4}$$

as $T \rightarrow \infty$. The objective of the present article is to sharpen this latter result by a precise asymptotic formula.

Theorem 1.2. For arbitrary real-valued $f \in C^4[a, b]$, with $f'' < 0$ on $[a, b]$, suppose that the curve $\mathcal{C} : \eta = f(\xi)$, $\xi \in [a, b]$, satisfies the *tangent condition*. Then it follows that

$$\int_0^T (S(t))^2 dt \sim CT^2,$$

with

$$C = \frac{1}{4\pi^2} \sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D} \\ G(\mathbf{u})=G(\mathbf{v})}}^* (u_2 v_2)^{-3/2} \kappa(\mathbf{u}) \kappa(\mathbf{v}),$$

where $u = (u_1, u_2), v = (v_1, v_2)$ are elements of $\mathbb{Z} \times \mathbb{N}$,

$$D = \{(\xi, \eta) \in \mathbb{R} \times \mathbb{R}^+ : -\eta f'(a) \leq \xi \leq -\eta f'(b)\},$$

$$\kappa(\xi, \eta) = \left| f'' \left(\varphi \left(-\frac{\xi}{\eta} \right) \right) \right|^{-1/2}, \quad G(\xi, \eta) = \eta f \left(\varphi \left(-\frac{\xi}{\eta} \right) \right) + \xi \varphi \left(-\frac{\xi}{\eta} \right),$$

where φ denotes the inverse function of f' , and \sum^* means that if u lies on the boundary of D , $\kappa(u)$ gets a factor $\frac{1}{2}$ (and accordingly for v).

Remark 1.3. The condition that $f'' < 0$ is technically convenient, but implies no real loss of generality: if $f'' > 0$ one can replace f by $-f$ and ψ by $-\psi$, with (1.1) still valid.

Remark 1.4. It is in fact sufficient to prove that (under the conditions stated)

$$\int_T^{2T} (S(t))^2 dt \sim 3CT^2. \tag{1.5}$$

To see this, apply (1.5) with T replaced by $T_k = 2^{-k}T, k = 1, \dots, K$, and estimate the integral over $[0, T_K]$ by (1.4). This gives

$$\limsup_{T \rightarrow \infty} \left| \frac{1}{T^2} \int_0^T (S(t))^2 dt - 3C \sum_{k=1}^K 4^{-k} \right| \ll 4^{-K}.$$

Since K can be picked arbitrarily large, Theorem 1.2 is immediate.

Remark 1.5. To fit our result into the frame of the classic theory, it is instructive to refer to the Gaussian circle problem. (For a thorough account on its history, the reader may consult, for example, the book of Krätzel [9].) Let $r(n)$ the number of ways to write $n \in \mathbb{N}_0$ as a sum of two squares, i.e.

$$r(n) = \#\{(u, v) \in \mathbb{Z}^2 : u^2 + v^2 = n\},$$

then it is well known that

$$\sum_{0 \leq n \leq t^2} r(n) = \pi t^2 + P(t),$$

with

$$P(t) = -8S_0(t) + O(1), \quad S_0(t) := \sum_{0 < n \leq t/\sqrt{2}} \psi_* \left(t \sqrt{1 - \left(\frac{n}{t} \right)^2} \right), \tag{1.6}$$

with $\psi_*(w) = w - [w] - \frac{1}{2}$ throughout. Concerning the mean-square of this remainder term, Kátai [8] proved that

$$\int_0^X (P(M^{1/2}))^2 dM = C_0 X^{3/2} + O(X(\log X)^2), \quad C_0 = \frac{1}{3\pi^2} \sum_{n=1}^{\infty} \frac{(r(n))^2}{n^{3/2}},$$

thereby refining earlier results of Cramér [2] and Landau [13]. By (1.6) and a simple integration by parts,

$$\int_0^T (\mathcal{S}_0(t))^2 dt \sim \frac{3}{256} C_0 T^2.$$

It is an easy exercise to verify that, for this special case, $\frac{3}{256} C_0$ is equal to our C . (Note that $f(w) = (1 - w^2)^{1/2}$ implies that $\kappa(u_1, u_2) = (u_1^2 + u_2^2)^{-3/4} u_2^{3/2}$ and $G(u_1, u_2) = (u_1^2 + u_2^2)^{1/2}$.)

Furthermore, we are able to sharpen the statement of Theorem 1.2, provided that the function f and the limits a, b are algebraic.

Theorem 1.6. *For a real-valued function $f \in C^4[a, b]$, with $f'' < 0$ on $[a, b]$, again let the curve $C : \eta = f(\xi)$, $\xi \in [a, b]$, satisfy the tangent condition. Suppose further that a, b are algebraic numbers and that there exists a polynomial $p \in \mathbb{Z}[X, Y]$ such that*

$$p(x, f(x)) = 0,$$

for all $x \in [a, b]$. Then it follows that

$$\int_0^T (\mathcal{S}(t))^2 dt = CT^2 + O(T^{2-\omega}),$$

for some (small) $\omega > 0$ depending on f and a, b (C as in Theorem 1.2).

The method of the proofs of both results will make use of the ideas developed in the first-named author’s papers [10] and [11], where corresponding asymptotics are established related to the arithmetic functions $r_k^\pm(n)$, which count the number of ways to write a positive integer n as the sum (respectively, the difference) of two k th powers.

2. Some auxiliary results

Lemma 2.1 (Vaalers’s Lemma). *For arbitrary $w \in \mathbb{R}$ and $H \in \mathbb{R}^+$, $H \geq 1$, suppose that ψ satisfies (1.1), and put*

$$\psi_H^*(w) = -\frac{1}{\pi} \sum_{1 \leq h \leq H} \frac{\sin(2\pi hw)}{h} \tau\left(\frac{h}{[H] + 1}\right),$$

where

$$\tau(\xi) = \pi\xi(1 - \xi) \cot(\pi\xi) + \xi, \quad \text{for } 0 < \xi < 1.$$

Then the following inequality holds

$$|\psi(w) - \psi_H^*(w)| \leq \frac{1}{[H] + 1} \sum_{1 \leq h \leq H} \left(1 - \frac{h}{[H] + 1}\right) \cos(2\pi hw) + \frac{1}{2[H] + 2}.$$

Proof. For $w \notin \mathbb{Z}$, this is one of the main results in Vaaler [17]. For a very well-readable exposition, see also the book of Graham and Kolesnik [3, p. 116]. The case $w \in \mathbb{Z}$ is an obvious consequence by a limit argument or by direct evaluation. \square

Lemma 2.2 (A sharp form of the van der Corput transformation). Suppose that g is a real-valued function which possesses four continuous derivatives on the interval $[A, B]$. Let L and W be real parameters not less than 1, such that $B - A \asymp L$,

$$g^{(j)}(w) \ll WL^{1-j}, \quad \text{for } w \in [A, B], \quad j = 1, 2, 3, 4,$$

and, for some $C^* > 0$,

$$g''(w) \geq C^*WL^{-1}, \quad \text{for } w \in [A, B].$$

Denote by ϕ the inverse function of g' , and define

$$r(x) = \begin{cases} 0, & \text{if } g'(x) \in \mathbb{Z}, \\ \min\left(\frac{1}{\|g'(x)\|}, \sqrt{\frac{L}{W}}\right), & \text{else,} \end{cases}$$

with $\|\cdot\|$ denoting the distance from the nearest integer. Then it follows that

$$\sum_{A \leq k \leq B} e(g(k)) = \sum_{g'(A) \leq m \leq g'(B)}'' \frac{e(g(\phi(m)) - m\phi(m) + \frac{1}{8})}{\sqrt{g''(\phi(m))}} + O(r(A) + r(B)) + O(\log(2 + W)),$$

with the notation

$$\sum_{a \leq n \leq b}'' \Phi(n) = \frac{1}{2}(\chi_{\mathbb{Z}}(a)\Phi(a) + \chi_{\mathbb{Z}}(b)\Phi(b)) + \sum_{a < n < b} \Phi(n),$$

where $\chi_{\mathbb{Z}}(\cdot)$ is the indicator function of the integers. The O -constant depends on C^* and on the constants implied in the order symbols in the suppositions.

Remark 2.3. Unfortunately, this result does not appear to be available explicitly in the literature. There are versions which state what is needed but impose a complicated condition essentially meaning that g is algebraic (see Krätzel Theorem 2.11 in [9], which is based on ideas due to Vinogradov). Graham and Kolesnik [3] avoid this restriction but unfortunately produce an error term $O(\sqrt{L/W})$ (in our notation) which is too crude for the present purpose. However, it is easy to construct what we need from the ideas in Graham and Kolesnik [3] and Krätzel [9]. Let us first state the following result.

Lemma 2.4. Suppose that F is a real-valued function which possesses four continuous derivatives on the interval $[A, B]$. Let L and W be real parameters not less than 1, such that $B - A \asymp L$,

$$F^{(j)}(w) \ll WL^{1-j}, \quad \text{for } w \in [A, B], \quad j = 1, 2, 3, 4,$$

and, for some $C^* > 0$,

$$F''(w) \geq C^*WL^{-1}, \quad \text{for } w \in [A, B].$$

Suppose further that there exists a value $c \in [A, B]$ for which $F'(c) = 0$. Define

$$R(x) = \min\left(\frac{1}{|F'(x)|}, \sqrt{\frac{L}{W}}\right).$$

Then it follows that

$$\int_A^B e(F(w)) dw = \begin{cases} (F''(c))^{-1/2}e(\frac{1}{8} + F(c)) + O(R(A) + R(B)) + O\left(\frac{1}{W}\right), & \text{if } A < c < B, \\ \frac{1}{2}(F''(A))^{-1/2}e(\frac{1}{8} + F(A)) + O(R(B)) + O\left(\frac{1}{W}\right), & \text{if } c = A, \\ \frac{1}{2}(F''(B))^{-1/2}e(\frac{1}{8} + F(B)) + O(R(A)) + O\left(\frac{1}{W}\right), & \text{if } c = B. \end{cases}$$

Proof of Lemma 2.4. For $A < c < B$, this is explicitly contained in Lemma 3.4 of Graham and Kolesnik [3]. The case $c = B$ can be reduced to $c = A$ by the substitution $w \rightarrow A + B - w$. Finally, to deal with the case $c = A$, it suffices to have a close look at the proof of Lemma 3.4 in [3]. Here, F is approximated by its quadratic Taylor expansion $q(w)$ at the stationary point c . If $c = A$, the integral $\int_A^B e(q(w)) dw$ can be evaluated by an obvious variant of Lemma 3.3 (in [3]), namely

$$\int_0^X e(Uw^2) dw = \frac{e(\frac{1}{8})}{2\sqrt{2U}} + O\left(\frac{1}{UX}\right).$$

The ingenious estimation of the remainder integral then works exactly as in [3]. □

Proof of Lemma 2.2. We start from formula (2.20) in Krätzel [9], and follow the proof of his Theorem 2.11 (p. 48). Thus, first of all,

$$\sum_{A \leq k \leq B} e(g(k)) = \sum_{g'(A) - \frac{1}{2} < m \leq g'(B) + \frac{1}{2}} \int_A^B e(g(t) - mt) dt + O(\log(2 + W)).$$

Here, we want to replace

$$\sum_{g'(A) - \frac{1}{2} < m \leq g'(B) + \frac{1}{2}}$$

with

$$\sum_{g'(A) \leq m \leq g'(B)},$$

and estimate the error we thereby commit. If $\{g'(B)\} \geq \frac{1}{2}$ (with $\{\cdot\}$ denoting the fractional part), we have to take into account the integral corresponding to $m^* = [g'(B)] + 1$:

$$\int_A^B e(g(t) - m^*t) dt \ll \min\left(\frac{1}{m^* - g'(B)}, \sqrt{\frac{L}{W}}\right) = r(B),$$

by the first (respectively, second) derivative test. If $0 < \{g'(A)\} < \frac{1}{2}$, an analogous argument produces the bound $O(r(A))$ for the very first integral. Consequently,

$$\sum_{A \leq k \leq B} e(g(k)) = \sum_{g'(A) \leq m \leq g'(B)} \int_A^B e(g(t) - mt) dt + O(r(A) + r(B)) + O(\log(2 + W)).$$

Each of these integrals is now evaluated by Lemma 2.4, with $F_m(t) = g(t) - mt$ instead of $F(t)$ (and $R(\cdot)$ replaced by $R_m(\cdot)$). The main terms give no difficulty, thus we proceed to estimate the remainder. If $g'(B) \in \mathbb{Z}$, we have

$$\sum_{g'(A) \leq m < g'(B)} R_m(B) \ll \sum_{g'(A) \leq m \leq g'(B)-1} \frac{1}{g'(B) - m} \ll \log(2 + W),$$

otherwise,

$$\sum_{g'(A) \leq m \leq g'(B)} R_m(B) \ll r(B) + \sum_{g'(A) \leq m \leq g'(B)-\frac{1}{2}} \frac{1}{g'(B) - m} \ll r(B) + \log(2 + W).$$

The cases $g'(A) \in \mathbb{Z}$ or $g'(A) \notin \mathbb{Z}$ are dealt with in the same way, producing the bound $r(A) + \log(2 + W)$. This completes the proof of Lemma 2.2. \square

We conclude this section by summarizing some important properties of the function $G(\xi, \eta)$ defined in Theorem 1.2. First of all, it is evident that $G(\xi, \eta)$ is homogeneous, of degree 1, and possesses continuous partial derivatives up to order 3 throughout its domain of definition \mathcal{D} . Furthermore, the following facts have been verified in [15]:

- (i) $G(\xi, \eta)$ is of the same sign throughout \mathcal{D} ; and
- (ii) there exist constants $c_2 > c_1 > 0$, such that, for all $(\xi, \eta) \in \mathcal{D}$,

$$c_1 \sqrt{\xi^2 + \eta^2} \leq G(\xi, \eta) \leq c_2 \sqrt{\xi^2 + \eta^2}. \tag{2.1}$$

3. Proof of Theorem 1.2

We start with a bit of notation. For real $T > 0$ and any function $\Phi \in L^2[T, 2T]$, we define

$$\mathcal{M}_T(\Phi) = \int_T^{2T} (\Phi(t))^2 dt.$$

Throughout the sequel, T (large) and $\varepsilon > 0$ (small) are independent parameters (with $\log T$ sufficiently large compared to $|\log \varepsilon|$ if necessary); all O and \ll -constants may depend on f and $[a, b]$, but not on T and ε . We put $H = [T] + \sqrt{2}$ (so that $H \asymp T$ and $H \notin \mathbb{Q}$). Finally, let $t \in [T, 2T]$ throughout the sequel.

Appealing to Vaaler’s Lemma, we shall approximate $\mathcal{S}(t)$ by

$$\mathcal{S}_H^*(t) = \sum_{\substack{at < n \leq bt \\ n \in \mathbb{Z}}} \psi_H^* \left(t f \left(\frac{n}{t} \right) \right).$$

It readily follows that

$$\mathcal{M}_T(S - S_H^*) \ll \int_T^{2T} \left(\sum_{1 \leq h \leq H} \beta_1(h, H) \sum_{at < n \leq bt} \cos\left(2\pi h t f\left(\frac{n}{t}\right)\right) \right)^2 dt + \frac{T^3}{H^2}, \tag{3.1}$$

with

$$\beta_1(h, H) = \frac{1}{[H] + 1} \left(1 - \frac{h}{[H] + 1}\right).$$

We shall first evaluate

$$\mathcal{M}_T(S_H^*) = \int_T^{2T} \left(\sum_{1 \leq h \leq H} \beta_2(h, H) \sum_{at < n \leq bt} \sin\left(2\pi h t f\left(\frac{n}{t}\right)\right) \right)^2 dt,$$

where

$$\beta_2(h, H) = \frac{1}{\pi} h^{-1} \tau\left(\frac{h}{[H] + 1}\right).$$

Our analysis will yield a bound for the right-hand side of (3.1) as a by-result.

We start by submitting the inner trigonometric sum to a van der Corput transform. Applying Lemma 2.2 with $g(w) = -htf(w/t)$, $L \asymp T$, $W \asymp h$, we readily obtain

$$\sum_{at < n \leq bt} e\left(-htf\left(\frac{n}{t}\right)\right) = \frac{\sqrt{t}}{\sqrt{h}} \sum''_{-hf'(a) \leq m \leq -hf'(b)} \kappa(m, h) e(-tG(m, h) + \frac{1}{8}) + O(\log(1 + h)) + O(r_h(a)) + O(r_h(b)), \tag{3.2}$$

with

$$r_h(c) = \left\{ \begin{array}{ll} 0, & \text{if } hf'(c) \in \mathbb{Z}, \\ \min\left(\|hf'(c)\|^{-1}, \frac{\sqrt{T}}{\sqrt{h}}\right), & \text{else.} \end{array} \right\} \quad (c = a \text{ or } b).$$

First, we estimate the contribution of the error terms to $\mathcal{M}_T(S_H^*)$. If $f'(a) \in \mathbb{Q}$, $r_h(a)$ is bounded, and, thus,

$$\sum_{1 \leq h \leq H} \beta_2(h, H) r_h(a) \ll \sum_{1 \leq h \leq H} h^{-1} \ll \log H.$$

If $f'(a) \notin \mathbb{Q}$,

$$\begin{aligned} & \sum_{1 \leq h \leq (1/\varepsilon)} \beta_2(h, H) r_h(a) + \sum_{(1/\varepsilon) < h \leq H} \beta_2(h, H) r_h(a) \\ & \ll |\log \varepsilon| \max_{1 \leq h \leq (1/\varepsilon)} \|hf'(a)\|^{-1} + \sqrt{T} \sum_{h > (1/\varepsilon)} h^{-3/2} \ll K_1(\varepsilon) + \sqrt{\varepsilon} \sqrt{T}. \end{aligned}$$

Since the same applies to $r_h(b)$, we conclude that

$$\begin{aligned} \mathcal{M}_T\left(\sum_{1 \leq h \leq H} \beta_2(h, H) (O(\log(1 + h)) + O(r_h(a)) + O(r_h(b)))\right) \\ \ll T(\log T)^4 + T(K_1(\varepsilon))^2 + \varepsilon T^2. \end{aligned} \tag{3.3}$$

Considering the imaginary part of (3.2), it remains to evaluate $\mathcal{M}_T(\sqrt{t}\Sigma(\mathcal{D}^{(H)}))$, introducing the useful notation

$$\mathcal{D}^{(H)} := \{(m, h) \in \mathbb{Z} \times \mathbb{N} : -hf'(a) \leq m \leq -hf'(b), 1 \leq h \leq H\},$$

and, for any domain $\mathcal{D}_0 \subseteq \mathcal{D}$,

$$\Sigma(\mathcal{D}_0) := \sum_{(m,h) \in \mathcal{D}_0 \cap \mathbb{Z}}^* h^{-1/2} \beta_2(h, H) \kappa(m, h) \sin(2\pi tG(m, h) - \frac{1}{4}\pi).$$

We split up $\mathcal{D}^{(H)}$ into subsets $\mathcal{D}_1, \mathcal{D}_2$ defined by

$$\begin{aligned} \mathcal{D}_1 &= \{(m, h) \in \mathbb{Z} \times \mathbb{N} : (m, h) \in \mathcal{D}, 1 \leq h \leq (1/\varepsilon)\}, \\ \mathcal{D}_2 &= \{(m, h) \in \mathbb{Z} \times \mathbb{N} : (m, h) \in \mathcal{D}, (1/\varepsilon) < h \leq H\}. \end{aligned}$$

In order to estimate the contribution of \mathcal{D}_2 , we further put

$$S(h, t) := \frac{\beta_2(h, H)}{\sqrt{h}} \sum_{-hf'(a) \leq m \leq -hf'(b)}'' \kappa(m, h) \sin(2\pi tG(m, h) - \frac{1}{4}\pi),$$

$H_i = 2^{-i}H, \mathbb{H}_i =]H_{i+1}, H_i[$ for $i = 0, 1, \dots, I$, with I minimal such that $H_{I+1} < (1/\varepsilon)$. We thus obtain

$$\begin{aligned} \mathcal{M}_T(\sqrt{t}\Sigma(\mathcal{D}_2)) &= \int_T^{2T} t \left(\sum_{(1/\varepsilon) < h \leq H} S(h, t) \right)^2 dt \\ &\ll \int_T^{2T} t \left(\sum_{i=0}^I H_i^{-1/6} \sum_{i=0}^I H_i^{1/6} \left(\sum_{h \in \mathbb{H}_i} S(h, t) \right)^2 \right) dt \\ &\ll \varepsilon^{1/6} \sum_{i=0}^I H_i^{1/6} \mathcal{M}_T \left(\sqrt{t} \sum_{h \in \mathbb{H}_i} S(h, t) \right). \end{aligned} \tag{3.4}$$

We further put

$$\mathcal{A}_i := \{(m, h) \in \mathbb{Z} \times \mathbb{N} : h \in \mathbb{H}_i, -hf'(a) \leq m \leq -hf'(b)\},$$

and write $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2)$, for elements of $\mathbb{Z} \times \mathbb{N}$ in the sequel. We thus obtain

$$\begin{aligned} \mathcal{M}_T \left(\sqrt{t} \sum_{h \in \mathbb{H}_i} S(h, t) \right) &= \mathcal{M}_T(\sqrt{t}\Sigma(\mathcal{A}_i)) \\ &= \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{A}_i} \frac{\beta_2(u_2, H)\beta_2(v_2, H)}{\sqrt{u_2 v_2}} \kappa(\mathbf{u})\kappa(\mathbf{v}) \\ &\quad \times \int_T^{2T} t \sin(2\pi tG(\mathbf{u}) - \frac{1}{4}\pi) \sin(2\pi tG(\mathbf{v}) - \frac{1}{4}\pi) dt. \end{aligned}$$

These integrals can be estimated either trivially or by the second mean-value theorem, on the basis of the elementary identity

$$\sin A \sin B = \frac{1}{2}(\cos(A - B) - \cos(A + B)). \tag{3.5}$$

Observing also that $\beta_2(h, H) \ll h^{-1}$, $\kappa(m, h) \ll 1$, we arrive at

$$\mathcal{M}_T\left(\sqrt{t} \sum_{h \in \mathbb{H}_i} S(h, t)\right) \ll H_i^{-3} \sum_{u, v \in \mathcal{A}_i} \min\left(T^2, \frac{T}{|G(u) - G(v)|}\right). \tag{3.6}$$

But this last expression (apart from an additional factor T) occurred already in [15, formula (5.5)]. As we have shown there (cf. the unnumbered formula below (5.7) in [15]) it is $\ll T^2 H_i^{-1/3} + T \log T$. Inserting this into (3.4), we infer that

$$\begin{aligned} \mathcal{M}_T(\sqrt{t}\Sigma(\mathcal{D}_2)) &\ll \varepsilon^{1/6} \sum_{i=0}^I (H_i^{-1/6} T^2 + H_i^{1/6} T \log T) \\ &\ll \varepsilon^{1/6} H_I^{-1/6} T^2 + H^{1/6} T \log T \\ &\ll \varepsilon^{1/3} T^2 + T^{7/6} \log T. \end{aligned} \tag{3.7}$$

At this point, we notice that our analysis can be readily modified to deal with the right-hand side of (3.1). We only have to replace ε by 2 (say), the sine by the cosine, and β_2 by β_1 . In view of the stronger bound $\beta_1(h, H) \ll H^{-1}$, the factor H_i^{-3} in (3.6) changes to $H^{-2} H_i^{-1}$, and one obtains altogether (taking into account that the error terms coming from the van der Corput transformation give at most the same bound as we had in (3.3)):

$$\begin{aligned} \mathcal{M}_T(S - S_H^*) &\ll \frac{T^3}{H^2} + \sum_{i=0}^I (H^{-2} H_i^{11/6} T^2 + H^{-2} H_i^{13/6} T \log T) \\ &\quad + T(\log T)^4 + T(K_1(\varepsilon))^2 + \varepsilon T^2 \\ &\ll \frac{T^3}{H^2} + H^{-1/6} T^2 + H^{1/6} T \log T + T(\log T)^4 + T(K_1(\varepsilon))^2 + \varepsilon T^2 \\ &\ll T^{11/6} + T(K_1(\varepsilon))^2 + \varepsilon T^2. \end{aligned} \tag{3.8}$$

It remains to evaluate

$$\begin{aligned} \mathcal{M}_T(\sqrt{t}\Sigma(\mathcal{D}_1)) &= \sum_{u, v \in \mathcal{D}_1}^* \frac{\beta_2(u_2, H)\beta_2(v_2, H)}{\sqrt{u_2 v_2}} \kappa(u)\kappa(v) \\ &\quad \times \int_T^{2T} t \sin(2\pi t G(u) - \frac{1}{4}\pi) \sin(2\pi t G(v) - \frac{1}{4}\pi) dt. \end{aligned}$$

The integrals are again computed using (3.5). If we consider only the contribution of ‘ $\cos(A - B)$ ’ and only the terms corresponding to u, v with $G(u) = G(v)$, replace \mathcal{D}_1 by \mathcal{D} , and approximate $\beta_2(h, H)$ by $(\pi h)^{-1}$, we readily obtain the main term $3CT^2$ stated in (1.5), with C as defined in Theorem 1.2. We only have to estimate the several errors we thereby commit as follows.

(i) *The contribution of the ‘cos(A + B)’ part of (3.5).* By the second mean-value theorem, this is

$$\ll \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{D}_1} (u_2 v_2)^{-3/2} \frac{T}{|G(\mathbf{u}) + G(\mathbf{v})|} \ll K_2(\varepsilon)T. \tag{3.9}$$

(ii) *The contribution of the $\mathbf{u}, \mathbf{v} \in \mathcal{D}_1$ with $G(\mathbf{u}) \neq G(\mathbf{v})$.* By the same argument, this is

$$\ll \sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D}_1 \\ G(\mathbf{u}) \neq G(\mathbf{v})}} (u_2 v_2)^{-3/2} \frac{T}{|G(\mathbf{u}) - G(\mathbf{v})|} \ll K_3(\varepsilon)T. \tag{3.10}$$

(iii) *The error term caused by extending \mathcal{D}_1 to all of \mathcal{D} .* To estimate this we notice that, by a result of Swinnerton-Dyer [16],

$$\sum_{\substack{\mathbf{v} \in \mathcal{D}: \\ G(\mathbf{v})=W}} 1 \ll W^{3/5+\delta},$$

for any fixed $\delta > 0$ and large W . Further, $\mathbf{u} \in \mathcal{D}$ implies that $u_1 \ll u_2$ and, thus, $|G(\mathbf{u})| \asymp |\mathbf{u}|_2 \asymp u_2$ (by an appeal to (2.1)), with $|\cdot|_2$ denoting the Euclidean norm. Consequently, the error term in question is

$$\begin{aligned} &\ll T^2 \sum_{\substack{\mathbf{u} \in (\mathcal{D}-\mathcal{D}_1), \mathbf{v} \in \mathcal{D}: \\ G(\mathbf{u})=G(\mathbf{v})}} |\mathbf{u}|_2^{-3} \ll T^2 \sum_{\mathbf{u} \in (\mathcal{D}-\mathcal{D}_1)} |\mathbf{u}|_2^{-12/5+\delta} \\ &\ll T^2 \sum_{n \gg (1/\varepsilon)} r(n)n^{-6/5+\delta/2} \ll T^2 \left(\frac{1}{\varepsilon}\right)^{-1/5+\delta/2} \ll \varepsilon^{1/6} T^2. \end{aligned} \tag{3.11}$$

(This argument also verifies the convergence of the series representing C .)

(iv) *The error coming from approximating β_2 .* Recalling the definition in Vaaler’s Lemma, it is clear that $\tau(w) = 1 + O(w^2)$ on $[0, 1]$, thus

$$\beta_2(h, H) = \frac{1}{\pi h} \tau\left(\frac{h}{[H] + 1}\right) = \frac{1}{\pi h} + O\left(\frac{h}{H^2}\right).$$

The overall error is, therefore,

$$\ll \sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D}_1 \\ G(\mathbf{u})=G(\mathbf{v})}} |\mathbf{u}|_2^{-1} H^{-2} T^2 \ll K_4(\varepsilon). \tag{3.12}$$

We are now ready to collect all partial results. Putting together (3.3) and (3.7)–(3.12) (with Cauchy’s inequality in the back of our mind), we readily see that

$$\limsup_{T \rightarrow \infty} \left| \frac{1}{T^2} \mathcal{M}_T(S) - 3C \right| \ll \varepsilon^{1/12}.$$

Since ε can be chosen arbitrarily small, this completes the proof of (1.5) and thereby that of Theorem 1.2. □

4. Proof of Theorem 1.6

In what follows, ω_0 (small) and C_1, C_2, \dots , (possibly large) denote suitable positive constants depending on f and a, b . (The C_j may depend on each other but not on ω_0 .) It clearly suffices to make the above deduction of Theorem 1.2 effective in the sense that we put

$$\varepsilon = T^{-\omega_0},$$

and replace $K_1(\varepsilon), K_2(\varepsilon), \dots$, by certain powers of $\varepsilon^{-1} = T^{\omega_0}$. We start with two more auxiliary results.

Lemma 4.1. *Assume that the function f satisfies the suppositions of Theorem 1.6 and let $G : \mathcal{D} \rightarrow \mathbb{R}$ be defined as in Theorem 1.2. Then the following conclusions hold true.*

(i) *There exists a polynomial $p^* \in \mathbb{Z}[X, Y, Z]$ such that*

$$p^*(\xi, \eta, G(\xi, \eta)) = 0, \quad \text{for all } (\xi, \eta) \in \mathcal{D}.$$

(ii) *For any $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in \mathcal{D} \cap (\mathbb{Z} \times \mathbb{N})$ with $G(\mathbf{u}) \neq G(\mathbf{v})$,*

$$|G(\mathbf{u}) - G(\mathbf{v})| \gg M^{-C_1},$$

where $M := \max\{|u_1|, u_2, |v_1|, v_2\}$.

Proof of Lemma 4.1. The verification of (i) is a straightforward exercise in classic algebra. To establish (ii), let us write

$$p^*(X, Y, Z) = \sum_{i,j,k} a_{i,j,k} X^i Y^j Z^k = \sum_k A_k(X, Y) Z^k,$$

with $A_k \in \mathbb{Z}[X, Y], k = 0, 1, \dots$. We start from the equations

$$\sum_k A_k(\mathbf{u}) G(\mathbf{u})^k = 0, \quad \sum_k A_k(\mathbf{v}) G(\mathbf{v})^k = 0,$$

and apply a device usually employed to show that $\alpha - \beta$ is algebraic if α, β are also (see, for example, Landau [14, p. 6]). It follows that there exists a polynomial \tilde{p} such that

$$\tilde{p}(G(\mathbf{u}) - G(\mathbf{v})) = 0,$$

with

$$\tilde{p}(Z) = \sum_{k=0}^{K(\mathbf{u}, \mathbf{v})} B_k(\mathbf{u}, \mathbf{v}) Z^k \quad (B_{K(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v}) \neq 0),$$

where $K(\mathbf{u}, \mathbf{v}) \leq C_2$, and the B_k are polynomials in four variables with rational integer coefficients that satisfy $B_k(\mathbf{u}, \mathbf{v}) \ll M^{C_3}$.

We now consider the (minimal) field extension \mathbb{F} over \mathbb{Q} that contains $G(\mathbf{u}) - G(\mathbf{v})$ along with all its conjugates. Let \mathcal{G} denote the corresponding Galois group, then, of

course, $|\mathcal{G}| = [\mathbb{F} : \mathbb{Q}] \leq C_4$. Since $B_{K(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v})(G(\mathbf{u}) - G(\mathbf{v}))$ is a non-zero algebraic integer, the absolute value of its norm is

$$\left| \prod_{\chi \in \mathcal{G}} \chi(B_{K(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v})(G(\mathbf{u}) - G(\mathbf{v}))) \right| = |B_{K(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v})|^{|\mathcal{G}|} \left| \prod_{\chi \in \mathcal{G}} \chi(G(\mathbf{u}) - G(\mathbf{v})) \right| \geq 1.$$

For every $\chi \in \mathcal{G}$, $\chi(G(\mathbf{u}) - G(\mathbf{v}))$ is a zero of \bar{p} , hence

$$|\chi(G(\mathbf{u}) - G(\mathbf{v}))| \leq 1 + \sum_{k=0}^{K(\mathbf{u}, \mathbf{v})-1} |B_k(\mathbf{u}, \mathbf{v})| \ll M^{C_5}.$$

Therefore,

$$|G(\mathbf{u}) - G(\mathbf{v})| \geq |B_{K(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v})|^{-|\mathcal{G}|} \left| \prod_{\substack{\chi \in \mathcal{G} \\ \chi \neq \text{id}}} \chi(G(\mathbf{u}) - G(\mathbf{v})) \right|^{-1} \gg M^{-C_1},$$

which establishes part (ii) of Lemma 4.1. □

Lemma 4.2. For any algebraic number α of degree $N \geq 2$,

$$\sum_{1 \leq h \leq W} \frac{1}{h \|h\alpha\|} \ll W^{N-1} \log W,$$

with a computable \ll -constant depending on α .

Proof of Lemma 4.2. By Liouville’s classic theorem, there exists $c(\alpha) > 0$ such that

$$\|h\alpha\| \geq (c(\alpha)h^{N-1})^{-1},$$

for all $h \in \mathbb{N}$. Thus, Lemma 4.2 is a special case of Lemma 3.3 in Kuipers and Niederreiter [12, p. 123]. □

We are now prepared to continue the proof of Theorem 1.6, i.e. (as we said earlier) to replace all the $K_j(\varepsilon)$ in the deduction of Theorem 1.2 by explicit powers of $\varepsilon^{-1} = T^{\omega_0}$. We start after (3.2) with the estimation of

$$\sum_{1 \leq h \leq H} \beta_2(h, H)r_h(a),$$

for $f'(a) \notin \mathbb{Q}$. Since, by assumption, $f'(a)$ is an algebraic number, this is

$$\begin{aligned} & \sum_{1 \leq h \leq T^{\omega_0}} \beta_2(h, H)r_h(a) + \sum_{T^{\omega_0} < h \leq H} \beta_2(h, H)r_h(a) \\ & \ll \sum_{1 \leq h \leq T^{\omega_0}} \frac{1}{h \|hf'(a)\|} + \sqrt{T} \sum_{h > T^{\omega_0}} h^{-3/2} \ll T^{C_6\omega_0} + T^{1/2-\omega_0/2}, \end{aligned} \tag{4.1}$$

by an appeal to Lemma 4.2. Of course the same applies to $f'(b)$. Therefore, the expression $K_1(\varepsilon)$ in (3.3) and the preceding unnumbered formula can be replaced by $T^{C_6\omega_0}$, and (3.8) becomes

$$\mathcal{M}_T(S - S_H^*) \ll T^{11/6} + T^{1+2C_6\omega_0} + T^{2-\omega_0}. \tag{4.2}$$

Next we consider (3.10). By Lemma 4.1,

$$\sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D}_1 \\ G(\mathbf{u}) \neq G(\mathbf{v})}} (u_2 v_2)^{-3/2} \frac{T}{|G(\mathbf{u}) - G(\mathbf{v})|} \ll T^{1+C_1\omega_0} \left(\sum_{\mathbf{u} \in \mathcal{D}_1} 1 \right)^2 \ll T^{1+C_7\omega_0}. \tag{4.3}$$

The same bound holds for the sums in (3.10) (since G is of constant sign) and (3.12). Combining (4.2), (4.3), and (3.11), and appealing to Cauchy’s inequality, we thus obtain

$$\int_T^{2T} (S(t))^2 dt = 3CT^2 + O(T^{3/2+C_6\omega_0}) + O(T^{2-\omega_0/2}) + O(T^{3/2+C_7\omega_0}) + O(T^{2-\omega_0/12}).$$

Choosing ω_0 sufficiently small, we conclude that

$$\int_T^{2T} (S(t))^2 dt = 3CT^2 + O(T^{2-\omega}) \quad (\omega > 0).$$

A dyadic decomposition of the interval $[0, T]$ now completes the proof of Theorem 1.6. \square

References

1. E. BOMBIERI AND H. IWANIEC, On the order of $\zeta(\frac{1}{2} + it)$, *Ann. Scuola Norm. Sup. Pisa, Ser. IV* **13** (1986), 449–472.
2. H. CRAMÉR, Über zwei Sätze von Herrn G. H. Hardy, *Math. Z.* **15** (1922), 200–210.
3. S. W. GRAHAM AND G. KOLESNIK, *Van der Corput’s method of exponential sums* (Cambridge University Press, Cambridge, 1991).
4. M. N. HUXLEY, Exponential sums and rounding error, *J. Lond. Math. Soc.* **43** (1991), 367–384.
5. M. N. HUXLEY, Exponential sums and lattice points II, *Proc. Lond. Math. Soc.* **66** (1993), 279–301.
6. M. N. HUXLEY, *Area, lattice points, and exponential sums*, LMS Monographs, New Series, vol. 13 (Oxford, 1996).
7. H. IWANIEC AND C. J. MOZZOCHI, On the divisor and circle problems, *J. Number Theory* **29** (1988), 60–93.
8. I. KÁTAI, The number of lattice points in a circle, *Ann. Univ. Sci. Budap. Rolando Eötvös, Sect. Math.* **8** (1965), 39–60.
9. E. KRÄTZEL, *Lattice points* (Kluwer, Dordrecht, 1988).
10. M. KÜHLEITNER, On sums of two k th powers: an asymptotic formula for the mean-square of the error term, *Acta Arithm.* **92** (2000), 263–276.
11. M. KÜHLEITNER, On differences of two k th powers: an asymptotic formula for the mean-square of the error term, *J. Number Theory* **76** (1999), 22–44.
12. L. KUIPERS AND H. NIEDERREITER, *Uniform distribution of sequences* (Wiley, New York, 1974).

13. E. LANDAU, Über die Gitterpunkte in einem Kreise (Vierte Mitteilung), *Nachr. Königl. Ges. Wiss. Göttingen (Math.-Phys. Kl.)* **1923** (1923), 58–65.
14. E. LANDAU, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale* (Chelsea, New York, 1949).
15. W. G. NOWAK, Fractional part sums and lattice points, *Proc. Edinb. Math. Soc.* **41** (1998), 497–515.
16. H. P. F. SWINNERTON-DYER, The number of lattice points on a convex curve, *J. Number Theory* **6** (1974), 128–135.
17. J. D. VAALER, Some extremal problems in Fourier analysis, *Bull. Am. Math. Soc.* **12** (1985), 183–216.
18. J. G. VAN DER CORPUT, Zahlentheoretische Abschätzungen mit Anwendung auf Gitterpunktprobleme, *Math. Z.* **17** (1923), 250–259.
19. J. G. VAN DER CORPUT, Neue zahlentheoretische Abschätzungen, *Math. Ann.* **89** (1923), 215–254.
20. I. M. VINOGRADOV, *Selected works* (ed. L. D. Faddeev, R. V. Gamkrelidze, A. A. Karatsuba, K. K. Mardzhanishvili and E. F. Mishchenko) (Springer, Berlin, 1985).