

THE DISTRIBUTION OF CONSECUTIVE PRIME BIASES AND SUMS OF SAWTOOTH RANDOM VARIABLES

ROBERT J. LEMKE OLIVER AND KANNAN SOUNDARARAJAN

ABSTRACT. In recent work, we considered the frequencies of patterns of consecutive primes (mod q) and numerically found biases toward certain patterns and against others. We made a conjecture explaining these biases, the dominant factor in which permits an easy description but fails to distinguish many patterns that have seemingly very different frequencies. There was a secondary factor in our conjecture accounting for this additional variation, but it was given only by a complicated expression whose distribution was not easily understood. Here, we study this term, which proves to be connected to both the Fourier transform of classical Dedekind sums and the error term in the asymptotic formula for the sum of $\phi(n)$.

1. INTRODUCTION

Let p_n denote the sequence of primes in ascending order. Given $q \geq 3$ and $\mathbf{a} = (a_1, \dots, a_r)$ satisfying $(a_i, q) = 1$ for all $1 \leq i \leq r$, in recent work [7] we studied biases in the occurrence of the pattern \mathbf{a} in strings of r consecutive primes reduced (mod q). Thus, we defined

$$\pi(x; q, \mathbf{a}) := \#\{p_n \leq x : p_{n+i-1} \equiv a_i \pmod{q} \text{ for } 1 \leq i \leq r\},$$

and conjectured that

$$(1.1) \quad \pi(x; q, \mathbf{a}) = \frac{\text{li}(x)}{\phi(q)^r} \left(1 + c_1(q; \mathbf{a}) \frac{\log \log x}{\log x} + c_2(q; \mathbf{a}) \frac{1}{\log x} + O((\log x)^{-7/4}) \right),$$

where $c_1(q; \mathbf{a})$ and $c_2(q; \mathbf{a})$ are certain explicit constants. The term $c_1(q; \mathbf{a})$ is easily described,

$$c_1(q; \mathbf{a}) = \frac{\phi(q)}{2} \left(\frac{r-1}{\phi(q)} - \#\{i \leq r-1 : a_i \equiv a_{i+1} \pmod{q}\} \right),$$

and it acts as a bias against immediate repetitions in the pattern \mathbf{a} . The term $c_2(q; \mathbf{a})$ is more complicated, and the goal of this paper is to understand its distribution. If $r \geq 3$ then

$$c_2(q; \mathbf{a}) = \sum_{i=1}^{r-1} c_2(q; (a_i, a_{i+1})) + \frac{\phi(q)}{2} \sum_{j=1}^{r-2} \frac{1}{j} \left(\frac{r-1-j}{\phi(q)} - \#\{i : a_i \equiv a_{i+j+1} \pmod{q}\} \right),$$

so that it is sufficient to understand the case $r = 2$; that is, $c_2(q; (a, b))$ with $(a, q) = (b, q) = 1$.

For the sake of simplicity, we shall confine ourselves to the case when q is prime. For any character $\chi \pmod{q}$ we define

$$(1.2) \quad A_{q, \chi} = \prod_{p \nmid q} \left(1 - \frac{(1 - \chi(p))^2}{(p-1)^2} \right).$$

Robert Lemke Oliver was partially supported by NSF grant DMS-1601398. Kannan Soundararajan was partially supported by NSF grant DMS-1500237 and by a Simons Investigator grant from the Simons Foundation.

Then the quantity $c_2(q; (a, b))$ is given by

$$c_2(q; (a, a)) = \frac{q-2}{2} \log(q/2\pi),$$

and when $a \not\equiv b \pmod{q}$ by

$$(1.3) \quad c_2(q; (a, b)) = \frac{1}{2} \log \frac{2\pi}{q} + \frac{q}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \left(\bar{\chi}(b-a) + \frac{1}{\phi(q)} (\bar{\chi}(b) - \bar{\chi}(a)) \right) L(0, \chi) L(1, \chi) A_{q, \chi}.$$

The diagonal term $c_2(q; (a, a))$ is thus completely explicit, and of size $q \log q$. Our work here shows that the off-diagonal terms $c_2(q; (a, b))$ can also be large; usually they are of size about q , occasionally getting to size $q \log \log q$ (attaining both positive and negative values), which we believe is their maximal size.

Before stating our result, we make one more simplification. Define

$$(1.4) \quad C(k) = C(k; q) = \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \overline{\chi(k)} L(0, \chi) L(1, \chi) A_{q, \chi}.$$

Since $A_{q, \chi} \ll 1$, $L(1, \chi) \ll \log q$, and (upon using the functional equation) $L(0, \chi) \ll \sqrt{q} \log q$, from (1.3) it follows that for $a \not\equiv b \pmod{q}$

$$(1.5) \quad \frac{c_2(q; (a, b))}{q} = C(b-a) + O\left(\frac{(\log q)^2}{\sqrt{q}}\right).$$

Thus for large q it is enough to understand the distribution of $C(k)$ as k varies over all non-zero residue classes \pmod{q} . Since $L(0, \chi) = 0$ for even characters χ , in (1.4) only odd characters χ make a contribution, and therefore $C(k) = -C(-k)$ is an odd function of k .

Theorem 1.1. (1) *As $q \rightarrow \infty$ the distribution of $C(k)$ tends to a continuous probability distribution, symmetric around 0. Precisely, there is a continuous function Φ_C with $\Phi_C(-x) + \Phi_C(x) = 1$ such that uniformly for all $x \in [-X, X]$ one has*

$$\frac{1}{q} \#\{k \pmod{q} : C(k) \leq \frac{e^\gamma}{2} x\} = \Phi_C(x) + o(1).$$

(2) *Uniformly for all $e \leq x \leq (\frac{1}{2} - \epsilon) \log \log q$ one has*

$$\exp(-A_1 e^x/x) \geq \frac{1}{q} \#\{k \pmod{q} : C(k) \geq \frac{e^\gamma}{2} x\} \geq \exp(-A_2 e^x \log x)$$

for some positive constants A_1 and A_2 .

(3) *For all large q , there exists $k \pmod{q}$ with*

$$-C(-k) = C(k) \geq \left(\frac{e^\gamma}{4} - \epsilon\right) \log \log q.$$

(4) *For all $k \pmod{q}$ we have*

$$C(k) \ll (\log q)^{\frac{2}{3}} (\log \log q)^2.$$

(5) *The values $C(k)$ have an ‘‘almost periodic’’ structure. Precisely, suppose $1 \leq m \leq q/4$ is a multiple of every natural number below $B \geq 2$. Then*

$$\frac{1}{q} \sum_{k \pmod{q}} |C(k) - C(k+m)|^2 \ll \frac{1}{B^{1-\epsilon}} + \frac{m}{q} \log B.$$

We make a few comments concerning Theorem 1.1 before proceeding to related results. In part 1, we believe that the distribution for $C(k)$ has a density, which is to say that Φ_C is in fact differentiable. Our proof falls just a little short of establishing this. In part 2, there is a gap between the upper and lower bounds for the tail frequencies. With a little more care, we can improve the lower bound there to $\exp(-A_3 e^x)$ for a suitable positive constant A_3 , but there still remains a gap between the two bounds. The distribution of $C(k)$, and especially the double exponential decay seen in part 2, are reminiscent of the distribution of values of $L(1, \chi_d)$ (see [4]). Motivated by this analogy, or by extrapolating the lower bounds in part 2, we believe that in part 3 there should exist values of $C(k)$ as large as $(\frac{e^\gamma}{2} - \epsilon) \log \log q$. We also conjecture that $(\frac{e^\gamma}{2} + \epsilon) \log \log q$ should be the largest possible value of $C(k)$, which would be a substantial strengthening of part 4. Finally, in addition to the almost periodic structure given in part 5 (where k varies), there should be an almost periodic structure as q varies. That is, if q_1 and q_2 are two large random primes with $q_1 - q_2$ being a multiple of the numbers below B , then $C(k; q_1)$ and $C(k; q_2)$ will be close to each other (on average over k). We hope that an interested reader will embrace some of these remaining problems.

While the quantity $C(k)$ is the main focus of this paper, closely related objects arise in two other seemingly unrelated contexts. The first of these concerns Dedekind sums. Let $\psi(x)$ denote the sawtooth function defined by

$$\psi(x) = \begin{cases} \{x\} - 1/2 & \text{if } x \notin \mathbf{Z}, \\ 0 & \text{if } x \in \mathbf{Z}, \end{cases}$$

which is an odd function, periodic with period 1. If q is prime and a is a reduced residue (mod q), then the Dedekind sum $s_q(a)$ is defined by

$$(1.6) \quad s_q(a) := \sum_{x \pmod{q}} \psi\left(\frac{x}{q}\right) \psi\left(\frac{ax}{q}\right).$$

The Dedekind sum arises naturally in number theory when studying the modular transformation properties of the Dedekind η -function, but it also appears in other contexts and satisfies many interesting properties [1, 11]. We study here the discrete Fourier transform of the Dedekind sum $s_q(a)$. Thus for a prime q and residue class $t \pmod{q}$ we define

$$(1.7) \quad \widehat{s}_q(t) := \frac{1}{q} \sum_{a \pmod{q}} s_q(a) e(at/q),$$

where $e(z) = e^{2\pi iz}$ throughout. In Lemma 2.1 we shall see that

$$\widehat{s}_q(t) = \frac{-1}{\pi i \phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t) L(0, \chi) L(1, \chi),$$

so that $\widehat{s}_q(t)$ is indeed a simpler version of $C(k)$. An alternative useful expression is

$$(1.8) \quad \widehat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{n=1 \\ (n,q)=1}}^{\infty} \frac{\psi(t\bar{n}/q)}{n},$$

where \bar{n} denotes the multiplicative inverse of the reduced residue class $n \pmod{q}$ and the sum converges since the partial sums $\sum_{n \leq x, (n,q)=1} \psi(t\bar{n}/q)$ are bounded.

Theorem 1.2. (1) As $q \rightarrow \infty$ the distribution of $\pi i \widehat{s}_q(t)$ tends to a continuous probability distribution, symmetric around 0. Precisely, there is a continuous function Φ_s with $\Phi_s(-x) + \Phi_s(x) = 1$ such that uniformly for all $x \in [-X, X]$ one has

$$\frac{1}{q} \#\{t \pmod{q} : \pi i \widehat{s}_q(t) \leq \frac{e^\gamma}{2} x\} = \Phi_s(x) + o(1).$$

(2) Uniformly for all $e \leq x \leq (\frac{1}{2} - \epsilon) \log \log q$ one has

$$\exp(-A_1 e^x/x) \geq \frac{1}{q} \#\{t \pmod{q} : \pi i \widehat{s}_q(t) \geq \frac{e^\gamma}{2} x\} \geq \exp(-A_2 e^x \log x)$$

for some positive constants A_1 and A_2 .

(3) For all large q , there exists $t \pmod{q}$ with

$$-\pi i \widehat{s}_q(-t) = \pi i \widehat{s}_q(t) \geq \left(\frac{e^\gamma}{4} - \epsilon\right) \log \log q.$$

(4) For all $t \pmod{q}$ we have

$$\widehat{s}_q(t) \ll (\log q)^{\frac{2}{3}} (\log \log q)^2.$$

(5) The values $\widehat{s}_q(t)$ have an ‘‘almost periodic’’ structure. Precisely, suppose $1 \leq m \leq q/4$ is a multiple of every natural number below $B \geq 2$. Then

$$\frac{1}{q} \sum_{t \pmod{q}} |\widehat{s}_q(t) - \widehat{s}_q(k+m)|^2 \ll \frac{1}{B^{1-\epsilon}} + \frac{m}{q} \log B.$$

Theorem 1.2 exactly parallels the results of Theorem 1.1, with the same deficiencies discussed there. The proofs of Theorems 1.1 and 1.2 are nearly identical, and so we give details only for Theorem 1.1.

Our third topic concerns the remainder term in the asymptotic for the mean value of Euler’s ϕ -function. Define the quantity $R(x)$ by the relation

$$\sum_{n \leq x} \phi(n) = \frac{3}{\pi^2} x^2 + R(x).$$

Simple arguments show that $R(x) \ll x \log x$, and Walfisz [12] established that $R(x) \ll x(\log x)^{2/3}(\log \log x)^{4/3}$, which is presently the best known estimate. Montgomery [8] conjectured that $R(x) \ll x \log \log x$ and $R(x) = \Omega_\pm(x \log \log x)$, and he showed that $R(x) = \Omega_\pm(x \sqrt{\log \log x})$. Key to Montgomery’s work is the expression

$$R(x) = \frac{\phi(x)}{2} - x \sum_{n \leq x} \frac{\mu(n) \psi(x/n)}{n} + O\left(x \exp(-c \sqrt{\log x})\right)$$

for some positive constant c , where $\phi(x) = 0$ if $x \notin \mathbf{Z}$. The sum in this expression is akin to the equation (1.8) for $\widehat{s}_q(t)$ with \bar{n}/q replaced by $1/n$ and with the weight $1/n$ replaced with $\mu(n)/n$. Accordingly, many of the techniques used to prove Theorems 1.1 and 1.2 apply to $R(x)$ as well, though unfortunately with less precision owing to the presence of $\mu(n)$. For convenience, we define $\widetilde{R}(x) = R(x)/x - \phi(x)/2x$.

Theorem 1.3. *As $y \rightarrow \infty$ the distribution of $\tilde{R}(u)$ for real $u \leq y$ tends to a probability distribution, symmetric around 0. Precisely, there is a function Φ_R with $\Phi_R(-x) + \Phi_R(x) = 1$ such that uniformly for all $x \in [-X, X]$ one has*

$$\frac{1}{y} \text{meas}(\{u \leq y : \tilde{R}(u) \leq \frac{3e^\gamma}{\pi^2} x\}) = \Phi_R(x) + o(1),$$

where $\text{meas}(I)$ denotes the Lebesgue measure of $I \subseteq \mathbf{R}$. Moreover, uniformly for all $e \leq x \leq (\frac{1}{2} - \epsilon) \log \log y$ one has

$$\frac{1}{y} \text{meas}(\{u \leq y : \tilde{R}(u) \geq \frac{3e^\gamma}{\pi^2} x\}) \leq \exp(-A_1 e^x/x)$$

for some positive constant A_1 .

We prove Theorem 1.3 by showing that all positive integral moments of $\tilde{R}(n)$ exist and are not too large. The moment calculation refines earlier work of Pillai and Chowla [10] and Chowla [3], who computed the mean and variance respectively:

$$\sum_{n \leq x} \tilde{R}(n) = o(x) \quad \text{and} \quad \frac{1}{y} \int_0^y \tilde{R}(u)^2 du \sim \frac{1}{2\pi^2}.$$

In Theorem 1.3, using Montgomery's construction in his Ω -result, we can obtain a lower bound for the frequency of large values of $\tilde{R}(u)$ of the form $\exp(-e^{x^{2+\epsilon}})$, which is very far from the upper bound. We expect that there is a lower bound similar to that in Theorems 1.1 and 1.2 in this situation also, and this would be in keeping with Montgomery's conjecture on the true size of $\tilde{R}(u)$.

Organization. Our main focus is the proofs of Theorems 1.1 and 1.2. We establish preliminary results useful for both in Sections 2 and 3. We then prove Theorem 1.1 in Sections 4-6; since the proof of Theorem 1.2 follows along identical lines, we omit it. In Section 7, we discuss the modifications that lead to Theorem 1.3.

2. FIRST STEPS

Here we establish some formulae for $\hat{s}_q(t)$ and $C(k)$ which will be the basis for our subsequent work.

Lemma 2.1. *Let q be prime. For any $(t, q) = 1$, we have*

$$\hat{s}_q(t) = \frac{-1}{\pi i \phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t) L(0, \chi) L(1, \chi) = \frac{1}{\pi i} \sum_{\substack{n=1 \\ (n,q)=1}}^{\infty} \frac{\psi(t\bar{n}/q)}{n}.$$

Moreover, for any $x \geq 1$ we have

$$\hat{s}_q(t) = \frac{1}{\pi i} \sum_{\substack{n \leq x \\ (n,q)=1}} \frac{\psi(t\bar{n}/q)}{n} + O\left(\frac{q}{x}\right).$$

Proof. For any non-principal character $\chi \pmod{q}$, we have (see, e.g., [13, Theorem 4.2])

$$(2.1) \quad L(0, \chi) = - \sum_{a \pmod{q}} \chi(a) \psi(a/q).$$

Notice that $L(0, \chi) = 0$ if χ is an even character, and that right side of the formula in (2.1) evaluates to 0 if χ is principal. The functional equation for odd characters gives

$$L(1, \chi) = -\frac{\tau(\chi)\pi i}{q}L(0, \bar{\chi}),$$

where $\tau(\chi) = \sum_{m \pmod{q}} \chi(m)e(m/q)$ denotes the Gauss sum. Thus we obtain

$$\begin{aligned} \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t)L(0, \chi)L(1, \chi) &= -\frac{\pi i}{q} \sum_{\chi \pmod{q}} \tau(\chi) \left| \sum_{a \pmod{q}} \chi(a)\psi\left(\frac{a}{q}\right) \right|^2 \\ &= -\frac{\pi i}{q} \sum_{a, b, m \pmod{q}} e(m/q)\psi\left(\frac{a}{q}\right)\psi\left(\frac{b}{q}\right) \sum_{\chi \pmod{q}} \chi(am)\bar{\chi}(bt) \\ &= -\frac{\phi(q)\pi i}{q} \sum_{a, b \neq 0 \pmod{q}} e\left(\frac{tb\bar{a}}{q}\right)\psi\left(\frac{a}{q}\right)\psi\left(\frac{b}{q}\right) \\ &= -\phi(q)\pi i \hat{s}_q(t). \end{aligned}$$

The first identity in the lemma follows.

To obtain the second identity, note that using (2.1) and the orthogonality relation for characters

$$\begin{aligned} - \sum_{\chi \neq \chi_0 \pmod{q}} \bar{\chi}(t)L(0, \chi) \sum_{n \leq N} \frac{\chi(n)}{n} &= \sum_{\chi \pmod{q}} \bar{\chi}(t) \sum_{a \pmod{q}} \chi(a)\psi(a/q) \sum_{n \leq N} \frac{\chi(n)}{n} \\ (2.2) \qquad \qquad \qquad &= \phi(q) \sum_{\substack{n \leq N \\ (n, q)=1}} \frac{1}{n} \psi(t\bar{n}/q). \end{aligned}$$

Letting $N \rightarrow \infty$, the second identity follows.

To obtain the truncated version, note that

$$\left| \sum_{\substack{n \leq x \\ (n, q)=1}} \psi(t\bar{n}/q) \right| \leq q$$

trivially, and therefore

$$\sum_{\substack{n > x \\ (n, q)=1}} \frac{\psi(t\bar{n}/q)}{n} = \int_x^\infty \frac{1}{y^2} \sum_{x < n \leq y} \psi(t\bar{n}/q) dy \ll \frac{q}{x}.$$

□

Recall the definition of $A_{q, \chi}$ from (1.2). Expanding this product out, we find

$$(2.3) \quad A_{q, \chi} = (2\chi(2) - \chi(2)^2) \prod_{p \nmid 2q} \left(1 - \frac{1}{(p-1)^2}\right) \left(1 + \frac{2\chi(p) - \chi(p)^2}{p^2 - 2p}\right) = C \sum_{n=1}^{\infty} a(n)\chi(2n).$$

Here

$$(2.4) \quad C = 2 \prod_{\substack{p \geq 3 \\ p \nmid q}} \left(1 - \frac{1}{(p-1)^2}\right),$$

and $a(n)$ is a multiplicative function defined by $a(2) = -1/2$ and $a(2^v) = 0$ for all $v \geq 2$, and for odd primes p we have

$$(2.5) \quad a(p) = \frac{2}{p(p-2)}, \quad a(p^2) = -\frac{1}{p(p-2)}, \quad \text{and} \quad a(p^v) = 0 \text{ for all } v \geq 3.$$

From the definition of $a(n)$ it is easy to check that $\sum_{n=1}^{\infty} |a(n)|n^{\sigma}$ converges for all $\sigma < 1/2$ so that

$$(2.6) \quad \sum_{n \geq N} |a(n)| \ll N^{-\frac{1}{2}+\epsilon} \quad \text{and} \quad C \sum_{\substack{n \leq N \\ (n,q)=1}} a(n) = 1 + O(N^{-\frac{1}{2}+\epsilon}).$$

Lemma 2.2. *Define the multiplicative function $b(n)$ by setting $b(n) = \sum_{uv=n} a(u)/v$, so that $b(n) = 0$ unless n is odd and square-free, and $b(p) = 1/(p-2)$ for all odd primes p . Then for any natural number N we have*

$$C(k) = -C \sum_{\substack{n \leq N \\ (n,q)=1}} b(n)\psi(k\overline{2n}/q) + O(q^{\frac{3}{2}+\epsilon}N^{-\frac{1}{4}+\epsilon}).$$

Proof. Arguing as in (2.2) we find

$$(2.7) \quad \frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \overline{\chi(k)} L(0, \chi) \sum_{\substack{n \leq N \\ (n,q)=1}} b(n)\chi(2n) = - \sum_{\substack{n \leq N \\ (n,q)=1}} b(n)\psi(k\overline{2n}/q).$$

Now if $n = uv \leq N$ then either $u \leq \sqrt{N}$ or $v \leq \sqrt{N}$ and $\sqrt{N} < u \leq N/v$. Therefore

$$(2.8) \quad \sum_{n \leq N} b(n)\chi(2n) = \sum_{u \leq \sqrt{N}} a(u)\chi(2u) \sum_{v \leq N/u} \frac{\chi(v)}{v} + \sum_{v \leq \sqrt{N}} \frac{\chi(v)}{v} \sum_{\sqrt{N} < u \leq N/v} a(u)\chi(2u).$$

Bounding the partial sums of characters trivially, we find

$$(2.9) \quad L(1, \chi) = \sum_{n \leq x} \frac{\chi(n)}{n} + \int_x^{\infty} \sum_{x < n \leq y} \chi(n) \frac{dy}{y^2} = \sum_{n \leq x} \frac{\chi(n)}{n} + O\left(\frac{q}{x}\right),$$

and so the first term in (2.8) is (using (2.6))

$$\begin{aligned} & \sum_{u \leq \sqrt{N}} a(u)\chi(2u) \left(L(1, \chi) + O\left(\frac{qu}{N}\right) \right) \\ &= C^{-1} A_{q,\chi} L(1, \chi) + O((\log q)N^{-\frac{1}{4}+\epsilon}) + O(qN^{-\frac{1}{2}}) \\ &= C^{-1} A_{q,\chi} L(1, \chi) + O(qN^{-\frac{1}{4}+\epsilon}). \end{aligned}$$

As for the second term in (2.8), using (2.6) we may bound this by

$$\ll \sum_{v \leq \sqrt{N}} \frac{1}{v} N^{-\frac{1}{4}+\epsilon} \ll N^{-\frac{1}{4}+\epsilon}.$$

We conclude that

$$C \sum_{\substack{n \leq N \\ (n,q)=1}} b(n)\psi(k\overline{2n}/q) = -\frac{1}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \overline{\chi(k)} L(0, \chi) \left(A_{q,\chi} L(1, \chi) + O(qN^{-\frac{1}{4}+\epsilon}) \right),$$

and since $L(0, \chi) \ll \sqrt{q} \log q$, the lemma follows. \square

Lemmas 2.1 and 2.2 give crude approximations to $\widehat{s}_q(t)$ and $C(k)$ by long sums (for example taking $x = q^2$ in Lemma 2.1, or taking $N = q^8$ in Lemma 2.2). However, on average over t or k , it is possible to approximate these quantities by very short sums.

Lemma 2.3. *Let $1 \leq B < q$ be a real number. Then*

$$\frac{1}{\phi(q)} \sum_{k \pmod{q}} \left| C(k) + C \sum_{n \leq B} b(n) \psi(k\overline{2n}/q) \right|^2 \ll B^{-1+\epsilon},$$

and

$$\frac{1}{\phi(q)} \sum_{t \pmod{q}} \left| \widehat{s}_q(t) - \frac{1}{\pi i} \sum_{n \leq B} \frac{\psi(t\overline{n}/q)}{n} \right|^2 \ll B^{-1+\epsilon}.$$

Proof. We shall content ourselves with proving the estimate for $C(k)$, the situation for $\widehat{s}_q(t)$ being entirely similar. Using (2.7) and Lemma 2.2 we see that

$$\begin{aligned} & \frac{1}{\phi(q)} \sum_{k \pmod{q}} \left| C(k) + C \sum_{n \leq B} b(n) \psi(k\overline{2n}/q) \right|^2 \\ &= \frac{1}{\phi(q)} \sum_{k \pmod{q}} \left| \frac{C}{\phi(q)} \sum_{\chi \neq \chi_0 \pmod{q}} \overline{\chi(k)} L(0, \chi) \sum_{B < n \leq q^{10}} b(n) \chi(2n) + O(q^{-1+\epsilon}) \right|^2. \end{aligned}$$

Using the orthogonality of characters to evaluate the sum over k , this is

$$\ll \frac{1}{\phi(q)^2} \sum_{\chi \neq \chi_0 \pmod{q}} |L(0, \chi)|^2 \left| \sum_{B < n \leq q^{10}} b(n) \chi(n) \right|^2 + q^{-2+\epsilon},$$

and using (2.9) and the functional equation this is

$$\ll \frac{1}{q} \sum_{\chi \neq \chi_0 \pmod{q}} \left| \sum_{m \leq q^2} \frac{\chi(m)}{m} \sum_{B < n \leq q^{10}} b(n) \chi(n) \right|^2 + q^{-2+\epsilon}.$$

Write temporarily

$$\sum_{m \leq q^2} \frac{\chi(m)}{m} \sum_{B < n \leq q^{10}} b(n) \chi(n) = \sum_{B < n \leq q^{12}} \frac{\alpha(n)}{n} \chi(n),$$

for some coefficients $\alpha(n) \ll n^\epsilon$. Then (including also the contribution of χ_0 below)

$$\frac{1}{q} \sum_{\chi \neq \chi_0 \pmod{q}} \left| \sum_{B < n \leq q^{12}} \frac{\alpha(n)}{n} \chi(n) \right|^2 \ll \sum_{\substack{B < n_1, n_2 \leq q^{12} \\ n_1 \equiv n_2 \pmod{q}}} \frac{|\alpha(n_1) \alpha(n_2)|}{n_1 n_2}.$$

The terms with n_1, n_2 both below q (so that $n_1 = n_2$) contribute

$$\ll \sum_{B < n < q} \frac{n^\epsilon}{n^2} \ll B^{-1+\epsilon}.$$

The terms with $\max(n_1, n_2) \geq q$ contribute (assume without loss of generality that n_2 is the larger one)

$$\ll q^\epsilon \sum_{B < n_1 \leq q^{12}} \frac{1}{n_1} \sum_{\substack{q < n_2 \leq q^{12} \\ n_2 \equiv n_1 \pmod{q}}} \frac{1}{n_2} \ll q^\epsilon \log q \frac{\log q}{q} \ll q^{-1+\epsilon}.$$

Assembling these estimates, the lemma follows. \square

3. A KEY QUANTITY

We shall study $\widehat{s}_q(t)$ and $C(k)$ by computing their moments, and the following key quantity will arise in this context. Let ℓ be a natural number, and suppose n_1, \dots, n_ℓ are ℓ natural numbers. Then set

$$(3.1) \quad \mathcal{B}(n_1, \dots, n_\ell) = \frac{1}{n_1 \cdots n_\ell} \int_0^{n_1 \cdots n_\ell} \prod_{j=1}^{\ell} \psi(x/n_j) dx.$$

Proposition 3.1. *The quantity $\mathcal{B}(n_1, \dots, n_\ell)$ satisfies the following properties.*

(1) *If ℓ is odd then $\mathcal{B}(n_1, \dots, n_\ell) = 0$. For even ℓ we have*

$$(3.2) \quad \mathcal{B}(n_1, \dots, n_\ell) = \left(\frac{i}{2\pi}\right)^\ell \sum_{\substack{k_1, \dots, k_\ell \neq 0 \\ \sum k_j/n_j = 0}} \frac{1}{k_1 \cdots k_\ell},$$

where the sum is over all non-zero integers k_j , and this sum is absolutely convergent. In the case $\ell = 2$ one has

$$\mathcal{B}(n_1, n_2) = \frac{(n_1, n_2)^2}{12n_1n_2}.$$

(2) *If p is a prime dividing n_j and such that p does not divide any other n_i , then*

$$\mathcal{B}(n_1, \dots, n_j, \dots, n_\ell) = \frac{1}{p} \mathcal{B}(n_1, \dots, n_j/p, \dots, n_\ell).$$

(3) *If we write $n_1 \cdots n_\ell = rs$ where r and s are coprime and r is square-free while s is square-full then*

$$|\mathcal{B}(n_1, \dots, n_\ell)| \leq 2^{-\ell} r^{-1}.$$

We begin by recalling the Fourier expansion of the sawtooth function. Note that $\widehat{\psi}(0) = 0$ and for $k \neq 0$ we have

$$(3.3) \quad \widehat{\psi}(k) = \int_0^1 \psi(x) e(-kx) dx = \frac{1}{-2\pi ik} = \frac{i}{2\pi k},$$

and so

$$(3.4) \quad \psi(x) = i \sum_{k \neq 0} \frac{e(kx)}{2\pi k}.$$

This series converges conditionally pointwise for each $x \notin \mathbf{Z}$, and also in the L^2 -sense. For any non-negative integer N , recall also the Fejer kernel

$$(3.5) \quad K_N(x) = \sum_{j=-N}^N \left(1 - \frac{|j|}{N+1}\right) e(jx) = \frac{1}{N+1} \left(\frac{\sin(\pi(N+1)x)}{\sin \pi x}\right)^2.$$

We shall find it convenient to replace $\psi(x)$ by the approximation $\psi_N(x)$ defined by

$$(3.6) \quad \psi_N(x) = i \sum_{0 < |k| \leq N} \frac{e(kx)}{2\pi k} \left(1 - \frac{|k|}{N+1}\right).$$

Note that ψ_N is the convolution of ψ with the Fejer kernel K_N

$$\psi_N(x) = \int_0^1 \psi(y) K_N(x-y) dy,$$

and so

$$(3.7) \quad |\psi_N(x) - \psi(x)| \ll \min\left(1, \frac{1}{N\|x\|}\right),$$

which implies that

$$(3.8) \quad \int_0^1 |\psi_N(x) - \psi(x)| dx \ll \frac{1 + \log N}{N}.$$

Note also that $|\psi_N(x)| \leq 1/2$ always.

Proof of Proposition 3.1: Part 1. Since ψ is an odd function, it is clear that $\mathcal{B}(n_1, \dots, n_\ell) = 0$ for odd ℓ . Now suppose ℓ is even. By Parseval it follows that

$$(3.9) \quad \frac{1}{n_1 \cdots n_\ell} \int_0^{n_1 \cdots n_\ell} \psi_N(x/n_1) \cdots \psi_N(x/n_\ell) dx = \left(\frac{i}{2\pi}\right)^\ell \sum_{\substack{0 < |k_j| \leq N \\ \sum k_j/n_j = 0}} \frac{1}{k_1 \cdots k_\ell} \prod_{j=1}^{\ell} \left(1 - \frac{|k_j|}{N+1}\right).$$

For any complex numbers $\alpha_1, \dots, \alpha_\ell$ and $\beta_1, \dots, \beta_\ell$ note the simple identity

$$(3.10) \quad \alpha_1 \cdots \alpha_\ell - \beta_1 \cdots \beta_\ell = (\alpha_1 - \beta_1)\alpha_2 \cdots \alpha_\ell + \beta_1(\alpha_2 - \beta_2)\alpha_3 \cdots \alpha_\ell + \beta_1 \cdots \beta_{\ell-1}(\alpha_\ell - \beta_\ell).$$

Applying this, we obtain

$$|\psi(x/n_1) \cdots \psi(x/n_\ell) - \psi_N(x/n_1) \cdots \psi_N(x/n_\ell)| \leq \frac{1}{2^{\ell-1}} \sum_{j=1}^{\ell} |\psi(x/n_j) - \psi_N(x/n_j)|,$$

and so by (3.9) and (3.8) we conclude that

$$(3.11) \quad \begin{aligned} \mathcal{B}(n_1, \dots, n_\ell) &= \frac{1}{n_1 \cdots n_\ell} \int_0^{n_1 \cdots n_\ell} \psi(x/n_1) \cdots \psi(x/n_\ell) dx \\ &= \left(\frac{i}{2\pi}\right)^\ell \sum_{\substack{0 < |k_j| \leq N \\ \sum k_j/n_j = 0}} \frac{1}{k_1 \cdots k_\ell} \prod_{j=1}^{\ell} \left(1 - \frac{|k_j|}{N+1}\right) + O\left(\frac{1 + \log N}{N}\right). \end{aligned}$$

We now show that

$$\sum_{\substack{0 < |k_j| \leq N \\ \sum k_j/n_j = 0}} \frac{1}{|k_1 \cdots k_\ell|}$$

is bounded, so that (3.11) will imply (letting $N \rightarrow \infty$) the stated formula (3.2) for $\mathcal{B}(n_1, \dots, n_\ell)$ and that the sum there converges absolutely. By Parseval

$$\sum_{\substack{0 < |k_j| \leq N \\ \sum k_j/n_j = 0}} \frac{1}{|k_1 \cdots k_\ell|} = \frac{1}{n_1 \cdots n_\ell} \int_0^{n_1 \cdots n_\ell} \prod_{j=1}^{\ell} \left(\sum_{0 < |k_j| \leq N} \frac{e(k_j x/n_j)}{|k_j|} \right) dx.$$

One may check that (with $\|x\|$ denoting the distance of x from the nearest integer)

$$\sum_{0 < |k| \leq N} \frac{e(k\theta)}{|k|} \ll \log \min\left(N, \frac{1}{\|\theta\|}\right) \ll \log \frac{N}{1 + N\|\theta\|}.$$

Using this and the arithmetic-geometric mean inequality above, we find

$$\begin{aligned} \sum_{\substack{0 < |k_j| \leq N \\ \sum k_j/n_j = 0}} \frac{1}{|k_1 \cdots k_\ell|} &\ll \sum_{j=1}^{\ell} \frac{1}{n_1 \cdots n_\ell} \int_0^{n_1 \cdots n_\ell} \left(\log \frac{N}{1 + N\|x/n_j\|} \right)^\ell dx \\ &\ll \int_0^1 \left(\log \frac{N}{1 + N\|x\|} \right)^\ell dx \ll 1. \end{aligned}$$

This proves our claim, and establishes (3.2).

If $\ell = 2$ then the condition $k_1/n_1 + k_2/n_2 = 0$ means that $k_1 = rn_1/(n_1, n_2)$ and $k_2 = -rn_2/(n_1, n_2)$ for some non-zero integer r . Therefore

$$\mathcal{B}(n_1, n_2) = -\frac{1}{4\pi^2} \sum_{r \neq 0} \frac{-1}{r^2} \frac{(n_1, n_2)^2}{n_1 n_2} = \frac{(n_1, n_2)^2}{12n_1 n_2}.$$

□

Proof of Proposition 3.1: Parts 2 and 3. If p divides n_j and no other n_i , then, in (3.2), k_j must necessarily be a multiple of p . Cancelling p from k_j and n_j , Part 2 follows. Part 3 follows from Part 2, and noting that $|\mathcal{B}(n_1, \dots, n_\ell)| \leq 2^{-\ell}$ always. □

For computing the moments of $\widehat{s}_q(t)$ and $C(k)$ the following proposition, which connects correlations of the sawtooth function with \mathcal{B} , will be very useful.

Proposition 3.2. *Let n_1, \dots, n_ℓ be positive integers. Define $K = n_1 \cdots n_\ell / \min(n_1, \dots, n_\ell)$. If $K < q/\ell$ then*

$$\frac{1}{q} \sum_{k \pmod{q}} \psi(k\overline{n_1}/q) \cdots \psi(k\overline{n_\ell}/q) = \mathcal{B}(n_1, \dots, n_\ell) + O\left(\frac{\ell K}{q} \log\left(\frac{eq}{K}\right)\right).$$

Proof. Take $N = \lfloor q/(\ell K) \rfloor$. The identity (3.10) gives

$$\begin{aligned} \sum_{k \pmod{q}} |\psi(k\overline{n_1}/q) \cdots \psi(k\overline{n_\ell}/q) - \psi_N(k\overline{n_1}/q) \cdots \psi_N(k\overline{n_\ell}/q)| \\ \leq \frac{1}{2^{\ell-1}} \sum_{j=1}^{\ell} \sum_{k \pmod{q}} |\psi(k\overline{n_j}/q) - \psi_N(k\overline{n_j}/q)|. \end{aligned}$$

Using now (3.7), the above is

$$(3.12) \quad \ll \frac{1}{2^\ell} \sum_{j=1}^{\ell} \sum_{k \pmod{q}} \min\left(1, \frac{1}{N\|k\overline{n_j}/q\|}\right) \ll \frac{q}{N} \log(eN).$$

By Parseval

$$(3.13) \quad \frac{1}{q} \sum_{k \pmod{q}} \psi_N(k\overline{n_1}/q) \cdots \psi_N(k\overline{n_\ell}/q) = \left(\frac{i}{2\pi}\right)^\ell \sum_{\substack{0 < |k_j| \leq N \\ \sum_j k_j \overline{n_j} \equiv 0 \pmod{q}}} \frac{1}{k_1 \cdots k_\ell} \prod_{j=1}^{\ell} \left(1 - \frac{|k_j|}{N+1}\right),$$

which bears a striking resemblance to (3.9). With our choice for N , we claim that in fact the right side of (3.13) is exactly equal to the expression in (3.9). Multiplying through

by $n_1 \cdots n_\ell$, the congruence $\sum k_j \bar{n}_j \equiv 0 \pmod{q}$ becomes $\sum_j k_j (n_1 \cdots n_\ell / n_j) \equiv 0 \pmod{q}$. Since $|k_j| < q/(\ell K)$ and $(n_1 \cdots n_\ell / n_j) \leq K$ for all j , it follows that $|\sum_j k_j (n_1 \cdots n_\ell / n_j)| < q$ so that the congruence becomes the equality $\sum_j k_j (n_1 \cdots n_\ell / n_j) = 0$, which is the same as the criterion $\sum_j k_j / n_j = 0$ of (3.9). Combining this observation with (3.11) and (3.12), our proposition follows. \square

4. THE MOMENTS OF $\widehat{s}_q(t)$ AND $C(k)$

We now state our main result on computing the moments of $\widehat{s}_q(t)$ and $C(k)$.

Theorem 4.1. *Let q be a prime, and ℓ a natural number. Then, uniformly in the range $\ell \leq \sqrt{\log q} / \log \log q$,*

$$(4.1) \quad \frac{1}{q} \sum_{k \pmod{q}} C(k)^\ell = M_C(\ell) + O(q^{-1/(20\ell \log \ell)}),$$

where

$$M_C(\ell) = C^\ell \sum_{n_1, \dots, n_\ell \geq 1} b(n_1) \cdots b(n_\ell) \mathcal{B}(n_1, \dots, n_\ell).$$

The quantity $M_C(\ell)$ equals zero for all odd ℓ , and for even ℓ satisfies

$$(4.2) \quad \frac{e^\gamma}{2} (\log \ell - \log \log \ell + O(1)) \leq M_C(\ell)^{\frac{1}{\ell}} \leq \frac{e^\gamma}{2} \log \ell + O(1).$$

Theorem 4.2. *Let q be a prime, and ℓ a natural number. Then, uniformly in ℓ ,*

$$\frac{1}{q} \sum_{t \pmod{q}} (\pi i \widehat{s}_q(t))^\ell = M_s(\ell) + O(q^{-1/(20\ell \log \ell)}),$$

where

$$M_s(\ell) = \sum_{n_1, \dots, n_\ell \geq 1} \frac{\mathcal{B}(n_1, \dots, n_\ell)}{n_1 \cdots n_\ell}.$$

The quantity $M_s(\ell)$ equals zero for all odd ℓ , and for even ℓ satisfies

$$\frac{e^\gamma}{2} (\log \ell - \log \log \ell + O(1)) \leq M_s(\ell)^{\frac{1}{\ell}} \leq \frac{e^\gamma}{2} \log \ell + O(1).$$

We confine ourselves to proving Theorem 4.1, and the proof of Theorem 4.2 follows along similar lines. In the rest of this section, we establish the asymptotic (4.1) and the upper bound in (4.2); the lower bound in (4.2) needs more work, and will be treated in the next section.

Proof of (4.1). Since $C(-k) = -C(k)$ the odd moments of $C(k)$ vanish. When ℓ is odd, $\mathcal{B}(n_1, \dots, n_\ell) = 0$ and so the quantity $M_C(\ell)$ is also zero here. In what follows, we may therefore assume that ℓ is an even natural number.

Let $1 \leq B \leq q$ be a parameter to be chosen shortly. Note that

$$\begin{aligned} & \left| C(k)^\ell - \left(-C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right)^\ell \right| \\ & \leq \left| C(k) + C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right| \cdot \sum_{j=0}^{\ell-1} |C(k)|^j \left| C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right|^{\ell-1-j} \\ & \leq (C_0 \log q)^{\ell-1} \left| C(k) + C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right|, \end{aligned}$$

for some absolute constant C_0 . By Cauchy-Schwarz and Lemma 2.3,

$$\frac{1}{q} \sum_{k \pmod{q}} \left| C(k) + C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right| \ll B^{-\frac{1}{2} + \epsilon}.$$

We choose $B = q^{1/\ell}$, and (in the range $\ell \leq \sqrt{\log q} / \log \log q$) deduce that

$$\frac{1}{q} \sum_{k \pmod{q}} C(k)^\ell = \frac{1}{q} \sum_{k \pmod{q}} \left(C \sum_{n \leq B} b(n) \psi \left(\frac{k\overline{2n}}{q} \right) \right)^\ell + O(q^{-\frac{1}{4\ell}}).$$

Expand out the main term above, replace $k \pmod{q}$ by $2k \pmod{q}$, and appeal to Proposition 3.2 with K there being $\leq q^{(\ell-1)/\ell}$. It follows that

$$(4.3) \quad \frac{1}{q} \sum_{k \pmod{q}} C(k)^\ell = C^\ell \sum_{n_1, \dots, n_\ell \leq q^{1/\ell}} b(n_1) \cdots b(n_\ell) \mathcal{B}(n_1, \dots, n_\ell) + O(q^{-\frac{1}{4\ell}}).$$

It remains now to bound the difference between the main term in (4.3) and the expression for $M_C(\ell)$, which is

$$\leq C^\ell \sum_{n > q^{1/\ell}} \sum_{n_1 \cdots n_\ell = n} b(n_1) \cdots b(n_\ell) |\mathcal{B}(n_1, \dots, n_\ell)| \leq (C/2)^\ell \sum_{n > q^{1/\ell}} \sum_{n_1 \cdots n_\ell = n} b(n_1) \cdots b(n_\ell) \frac{1}{\text{sf}(n)},$$

where $\text{sf}(n)$ is the largest squarefree divisor d of n that is coprime to n/d . We estimate the sum above by Rankin's trick; with $\alpha = 1/(10 \log \ell)$ the above is

$$\begin{aligned} & \leq (C/2)^\ell q^{-\alpha/\ell} \sum_{n=1}^{\infty} \sum_{n_1 \cdots n_\ell = n} b(n_1) \cdots b(n_\ell) \frac{n^\alpha}{\text{sf}(n)} \\ & \leq e^{O(\ell)} q^{-\alpha/\ell} \prod_{p \geq 3} \left(1 + \frac{\ell p^\alpha}{p(p-2)} + \sum_{j=2}^{\ell} \binom{\ell}{j} \frac{p^{j\alpha}}{(p-2)^j} \right), \end{aligned}$$

upon recalling the definition of $b(n)$. The contribution of primes $p \leq \ell$ to the product above is

$$\leq \prod_{3 \leq p \leq \ell} \left(1 + \frac{p^\alpha}{p-2} \right)^\ell \leq (\log \ell)^\ell e^{O(\ell)},$$

while the contribution of primes $p > \ell$ to the product above is

$$\ll \prod_{p > \ell} \exp \left(O \left(\frac{\ell^2 p^{2\alpha}}{p^2} \right) \right) = e^{O(\ell)}.$$

We conclude that the difference between the main term in (4.3) and the expression for $M_C(\ell)$ is

$$\ll (\log \ell)^\ell e^{O(\ell)} q^{-\alpha/\ell} \ll q^{-1/(20\ell \log \ell)},$$

completing the proof of (4.1). \square

Proof of the upper bound in (4.2). Note that

$$\begin{aligned} M_C(\ell) &\leq C^\ell \sum_{n_1, \dots, n_\ell} b(n_1) \cdots b(n_\ell) |\mathcal{B}(n_1, \dots, n_\ell)| \leq (C/2)^\ell \sum_{n_1, \dots, n_\ell} \frac{b(n_1) \cdots b(n_\ell)}{\text{sf}(n)} \\ &\leq (C/2)^\ell \prod_{p \geq 3} \left(1 + \frac{\ell}{p(p-2)} + \sum_{j=2}^{\ell} \binom{\ell}{j} \frac{1}{(p-2)^j} \right). \end{aligned}$$

The contribution of primes $p \leq \ell$ is

$$\leq \prod_{3 \leq p \leq \ell} \left(1 + \frac{1}{p-2} \right)^\ell = \left(\prod_{3 \leq p \leq \ell} \left(1 - \frac{1}{(p-1)^2} \right)^{-1} \left(1 - \frac{1}{p} \right)^{-1} \right)^\ell = C^{-\ell} (e^\gamma \log \ell + O(1))^\ell,$$

upon using Mertens's theorem. The contribution of primes $p > \ell$ is

$$\exp \left(\sum_{p > \ell} O\left(\frac{\ell^2}{p^2}\right) \right) = \exp \left(O\left(\frac{\ell}{\log \ell}\right) \right),$$

and so the upper bound in (4.2) follows. \square

5. COMPLETING THE PROOF OF THEOREM 4.1: PROOF OF THE LOWER BOUND IN (4.2)

To obtain the lower bound in (4.2) we take an indirect approach, working with a continuous model that has the same moments as $C(k)$. Let B be a positive integer, and let $L(B)$ denote the least common multiple of the natural numbers $n \leq B$. For a real number x , define

$$C(x; B) = C \sum_{n \leq B} b(n) \psi(x/n).$$

It follows readily that

$$\frac{1}{L(B)} \int_0^{L(B)} C(x; B)^\ell dx = C^\ell \sum_{n_1, \dots, n_\ell \leq B} b(n_1) \cdots b(n_\ell) \mathcal{B}(n_1, \dots, n_\ell),$$

so that

$$(5.1) \quad M_C(\ell) = \lim_{B \rightarrow \infty} \frac{1}{L(B)} \int_0^{L(B)} C(x; B)^\ell dx.$$

We shall obtain a lower bound for the right side of (5.1); naturally, we may assume that ℓ is even and large.

Suppose that $B > \ell$, and put $\ell_0 = \ell / \log \ell$. Let \mathcal{I} denote the subset of $[0, L(B)]$ consisting of points $x = kL(\ell_0) - y$ with $1 \leq k \leq L(B)/L(\ell_0)$, and $0 < y \leq 1/10$. Let $\psi^+(t) = \psi(t)$ whenever t is not an integer, and $\psi^+(t) = 1/2$ when t is an integer. Then for $x = kL(\ell_0) - y \in \mathcal{I}$ note that

$$C(x; B) = C \sum_{n \leq B} b(n) \psi((kL(\ell_0) - y)/n) = C \sum_{n \leq B} b(n) \left(\psi^+(kL(\ell_0)/n) - y/n \right).$$

Since, for $n \leq B$,

$$\sum_{k=1}^{L(B)/L(\ell_0)} \psi^+\left(\frac{kL(\ell_0)}{n}\right) = \frac{1}{2} \frac{L(B)}{L(\ell_0)} \frac{(n, L(\ell_0))}{n},$$

it follows that (note $|\mathcal{I}| = L(B)/(10L(\ell_0))$)

$$\frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} C(x; B) dx = \frac{C}{2} \sum_{n \leq B} b(n) \frac{(n, L(\ell_0)) - 1/20}{n},$$

and therefore by Hölder's inequality that

$$\frac{1}{L(B)} \int_0^B C(x, B)^\ell dx \geq \frac{1}{10L(\ell_0)} \frac{1}{|\mathcal{I}|} \int_{\mathcal{I}} C(x; B)^\ell dx \geq \frac{1}{10L(\ell_0)} \left(\frac{C}{2} \sum_{n \leq B} b(n) \frac{(n, L(\ell_0)) - 1/20}{n} \right)^\ell.$$

Now letting $B \rightarrow \infty$, we find by (5.1) that

$$M_C(\ell) \geq \frac{1}{10L(\ell_0)} \left(\frac{C}{2} \sum_{n=1}^{\infty} \frac{b(n)(n, L(\ell_0))}{n} + O(1) \right)^\ell \geq e^{-O(\ell_0)} \left(\frac{C}{2} \prod_{3 \leq p \leq \ell_0} \left(1 + \frac{1}{p-2} \right) + O(1) \right)^\ell,$$

upon using the prime number theorem to estimate $L(\ell_0)$, and recalling the definition of b . Now

$$\frac{C}{2} \prod_{3 \leq p \leq \ell_0} \left(1 + \frac{1}{p-2} \right) = \prod_{3 \leq p \leq \ell_0} \left(1 - \frac{1}{p} \right)^{-1} \left(1 + O\left(\frac{1}{\ell_0}\right) \right) = \frac{e^\gamma}{2} \log \ell_0 + O(1),$$

and therefore the lower bound in (4.2) follows.

6. PROOF OF THEOREM 1.1

Proof of Part 1. Theorem 4.1 shows that all the moments of $C(k)$ exist, and do not grow too rapidly. The moment generating function $\sum_{\ell=0}^{\infty} x^\ell M_C(\ell)/\ell!$ converges for all x , and therefore the sequence of moments $M_C(\ell)$ uniquely determines a distribution, which is the limiting distribution for $C(k)$. Since $C(k) = -C(-k)$, the limiting distribution is clearly symmetric around 0.

To gain an understanding of this limiting distribution, and to establish its continuity, it is helpful to think of the continuous model $C(x; B)$ discussed in Section 5. Consider the characteristic function (that is, Fourier transform) of $C(x; B)$; namely

$$\mathbb{E}(e^{itC(x, B)}) = \frac{1}{L(B)} \int_0^{L(B)} e^{itC(x, B)} dx.$$

Omit the measure zero set of integers x , and write $x = k - y$ with $1 \leq k \leq L(B)$ and $0 < y < 1$. Then, with ψ^+ as in Section 5 and $C^+(x; B) = C \sum_{b \leq B} b(n) \psi^+(x/b)$, we have $C(x; B) = C^+(k; B) - y \sum_{n \leq B} b(n)/n$, and so

$$(6.1) \quad \frac{1}{L(B)} \int_0^{L(B)} e^{itC(x, B)} dx = \frac{1}{L(B)} \sum_{k=1}^{L(B)} e^{itC^+(k, B)} \int_0^1 e^{-ity \sum_{n \leq B} b(n)/n} dy \ll \frac{1}{1 + |t|}.$$

Given an interval $I = (\alpha - \epsilon, \alpha + \epsilon)$ with $\epsilon < 1/2$, we can readily find a majorant $\Psi(x)$ of the indicator function of I , with $|\widehat{\Psi}(x)| \ll \epsilon/(1 + (\epsilon x)^2)$. For example take $\Psi(x) =$

$\max(2 - |x - \alpha|/\epsilon, 0)$, which is a relative of the Fejer kernel. Then by Fourier inversion

$$\begin{aligned} \frac{1}{L(B)} \int_{\substack{x \in [0, L(B)] \\ C(x, B) \in I}} dx &\leq \frac{1}{L(B)} \int_0^{L(B)} \Psi(C(x, B)) dx \\ &= \int_{-\infty}^{\infty} \widehat{\Psi}(t) \mathbb{E}(e^{itC(x, B)}) dt \ll \int_{-\infty}^{\infty} \frac{1}{1 + |t|} \frac{\epsilon}{1 + (\epsilon t)^2} dt \ll \epsilon \log(1/\epsilon). \end{aligned}$$

Therefore $C(x, B)$ has a continuous distribution, and the continuity is uniform in B , so that letting $B \rightarrow \infty$, we conclude that the limiting distribution for $C(k)$ is also continuous. \square

Proof of Parts 2 and 3. Since Part 3 follows upon taking $x = (\frac{1}{2} - \epsilon) \log \log q$ in Part 2, it is enough to prove Part 2. For any even $\ell \leq \sqrt{\log q} / \log \log q$, we see using Theorem 4.1 that

$$\frac{1}{q} \#\{k \pmod{q} : C(k) \geq \frac{e^\gamma}{2} x\} \leq \left(\frac{e^\gamma}{2} x\right)^{-\ell} (M_C(\ell) + o(1)) \ll \left(\frac{\log \ell + O(1)}{x}\right)^\ell.$$

Choosing ℓ to be an even integer around Ae^x for a suitably small positive constant A , the upper bound in Part 2 follows.

To establish the lower bound in Part 2, note that for even $\ell \leq \sqrt{\log q} / (2 \log \log q)$, we have by Theorem 4.1

$$(6.2) \quad \left(\frac{e^\gamma}{2} (\log \ell - \log \log \ell + O(1))\right)^\ell \ll \frac{1}{q} \sum_{k \pmod{q}} C(k)^\ell.$$

The contribution from terms k with $|C(k)| \leq \frac{e^\gamma}{2} (\log \ell - \log \log \ell - A)$ for a suitably large constant A is clearly negligible compared to the right side of (6.2). The contribution from terms k with $|C(k)| \geq \frac{e^\gamma}{2} (\log \ell + \log \log \ell + A)$ for a suitably large constant A is

$$\begin{aligned} &\leq \left(\frac{e^\gamma}{2} (\log \ell + \log \log \ell + A)\right)^{-\ell} \frac{1}{q} \sum_{k \pmod{q}} C(k)^{2\ell} \\ &\ll \left(\frac{e^\gamma}{2} (\log \ell + \log \log \ell + A)\right)^{-\ell} \left(\frac{e^\gamma}{2} \log \ell + O(1)\right)^{2\ell}, \end{aligned}$$

upon using Theorem 4.1 to estimate the 2ℓ -th moment. If A is suitably large, then this too is negligible in comparison to the right side of (6.2). Therefore it is the terms with $|C(k)|$ lying between $\frac{e^\gamma}{2} (\log \ell - \log \log \ell - A)$ and $\frac{e^\gamma}{2} (\log \ell + \log \log \ell + A)$ that account for the bulk of the contribution to (6.2), and so

$$\begin{aligned} &\left(\frac{e^\gamma}{2} (\log \ell + \log \log \ell + A)\right)^\ell \frac{1}{q} \#\{k : |C(k)| \geq \frac{e^\gamma}{2} (\log \ell - \log \log \ell - A)\} \\ &\gg \left(\frac{e^\gamma}{2} (\log \ell - \log \log \ell + O(1))\right)^\ell. \end{aligned}$$

Choosing ℓ of size xe^x , the lower bound in Part 2 follows. \square

Proof of Part 4. First suppose that $C(k)$ is negative. From [9] (Chapter 1, page 6) we recall that for each natural number K there is a trigonometric polynomial

$$B_K(x) = \frac{1}{2(K+1)} + \sum_{1 \leq |j| \leq K} c_j e(jx)$$

with $c_j \ll 1/j$, such that $B_K(x) \geq \psi(x)$ for all x . Using Lemma 2.2 with $N = q^8$ we obtain

$$0 \leq -C(k) = C \sum_{\substack{n \leq q^8 \\ (n,q)=1}} b(n) \psi(k\overline{2n}/q) + O(1) \leq C \sum_{n \leq q^8} b(n) B_K(k\overline{2n}/q) + O(1).$$

Thus, for some positive constant A ,

$$(6.3) \quad -C(k) \leq A \left(1 + \frac{1}{K+1} \sum_{n \leq q^8} b(n) + \sum_{1 \leq |j| \leq K} \frac{1}{j} \left| \sum_{\substack{n \leq q^8 \\ (n,q)=1}} b(n) e\left(\frac{kj\overline{2n}}{q}\right) \right| \right).$$

At this stage, we need the following result which follows from work of Bourgain and Garaev [2] (refining earlier work of Karatsuba [5]; see also Korolev [6]).

Lemma 6.1. *Let q be a prime, and a be any integer coprime to q . Then for all $N \geq 1$*

$$\left| \sum_{\substack{n \leq N \\ (n,q)=1}} \frac{1}{n} e\left(\frac{a\overline{n}}{q}\right) \right| \ll (\log q)^{\frac{2}{3}} (\log \log q)^2.$$

Proof. Theorem 16 of Bourgain and Garaev [2] gives

$$\left| \sum_{n \leq x} e\left(\frac{a\overline{n}}{q}\right) \right| \ll \frac{x}{(\log x)^{\frac{3}{2}}} \log q (\log \log q)^3.$$

Partial summation using this bound for $x \geq \exp((\log q)^{\frac{2}{3}} (\log \log q)^2)$, and the trivial bound (that the sum is at most x) for smaller x yields the lemma. \square

Returning to (6.3), take there $K = \lfloor \log q \rfloor$. Then the right side of (6.3) is (recalling the definition $b(n) = \sum_{uv=n} a(u)/v$)

$$\ll 1 + \sum_{j \leq K} \frac{1}{j} \sum_{\substack{u \leq q^8 \\ (u,q)=1}} |a(u)| \left| \sum_{\substack{v \leq q^8/u \\ (v,q)=1}} \frac{1}{v} e\left(\frac{kj\overline{2uv}}{q}\right) \right| \ll (\log q)^{\frac{2}{3}} (\log \log q)^3,$$

using Lemma 6.1 and since $\sum_n |a(n)| \ll 1$. This proves that $-C(k) \leq A(\log q)^{\frac{2}{3}} (\log \log q)^3$, which is the desired bound in the case $C(k)$ negative. Arguing similarly with a minorant for $\psi(x)$ instead of a majorant, leads to the same bound for $C(k)$ in the case when it is positive. \square

Proof of Part 5. Applying Lemma 2.3 we find that

$$\frac{1}{q} \sum_{k \pmod{q}} |C(k) - C(k+m)|^2 \ll B^{-1+\epsilon} + \frac{1}{q} \sum_{k \pmod{q}} \left| \sum_{n \leq B} b(n) \left(\psi\left(\frac{(k+m)\overline{2n}}{q}\right) - \psi\left(\frac{k\overline{2n}}{q}\right) \right) \right|^2.$$

Using Cauchy-Schwarz the second term above is

$$(6.4) \quad \ll \frac{1}{q} \left(\sum_{n \leq B} b(n) \right) \sum_{n \leq B} b(n) \sum_{k \pmod{q}} \left(\psi\left(\frac{k+m\overline{2n}}{q}\right) - \psi\left(\frac{k}{q}\right) \right)^2,$$

where in the inner sum we replaced k by $2kn$. Since $|\psi((k+a)/q) - \psi(k/q)| \leq |a|/q$ unless there is an integer between k/q and $(k+a)/q$, we may check that

$$\frac{1}{q} \sum_{k \pmod{q}} \left(\psi\left(\frac{k+a}{q}\right) - \psi\left(\frac{k}{q}\right) \right)^2 \ll \frac{a}{q}.$$

Since m is a multiple of all numbers B (and recalling that $b(n) = 0$ unless n is odd), we may write $m\overline{2n} = qr + a$ with $a = m/(2n)$. Therefore the quantity in (6.4) is

$$\ll (\log B) \sum_{n \leq B} \frac{m}{nq} \ll \frac{m}{q} \log B,$$

completing our proof. \square

7. PROOF OF THEOREM 1.3

As in the proofs of Theorems 1.1 and 1.2, the main result is to compute the moments of $\tilde{R}(u)$. The proof of Theorem 1.3 then follows in exactly the same way as the corresponding parts of Theorem 1.1.

Theorem 7.1. *There is a positive number $c < 1$ such that uniformly for all natural numbers ℓ in the range $\ell \leq \frac{c}{9} \sqrt{\log y} / \log \log y$, we have*

$$\frac{1}{y} \int_0^y \tilde{R}(u)^\ell du = M_R(\ell) + O\left(\exp\left(-\frac{c}{8} \sqrt{\log y}\right)\right),$$

where

$$M_R(\ell) = \sum_{n_1, \dots, n_\ell} \frac{\mu(n_1) \dots \mu(n_\ell)}{n_1 \dots n_\ell} \mathcal{B}(n_1, \dots, n_\ell).$$

For odd ℓ , $M_R(\ell) = 0$, while $M_R(2) = 1/2\pi^2$ and for even $\ell \geq 4$ we have

$$M_R(\ell) \leq \left(\frac{3e^\gamma}{\pi^2} \log \ell + O(1)\right)^\ell.$$

We begin with a lemma, which will allow us to truncate $\tilde{R}(u)$ by a short sum of sawtooth functions.

Lemma 7.2. *For all $1 \leq N \leq y$ we have*

$$\sum_{N < n_1, n_2 \leq 2N} \left| \frac{1}{y} \int_0^y \psi(x/n_1) \psi(x/n_2) dx \right| \ll (\log y)^2 \left(N + N^2 \frac{\sqrt{N}}{\sqrt{y}} \right).$$

Proof. Let $K \geq 2$ be a parameter to be chosen shortly, and let $\psi_K(x)$ be as in (3.6). First note that

$$\frac{1}{y} \int_0^y |\psi(x/n_1) \psi(x/n_2) - \psi_K(x/n_1) \psi_K(x/n_2)| dx \leq \frac{1}{y} \int_0^y \sum_{j=1}^2 |\psi(x/n_j) - \psi_K(x/n_j)| dx \ll \frac{1}{K},$$

upon using (3.8), and since n_1 and n_2 are at most $N \leq y$. Next, from the Fourier expansion of ψ_K (see (3.6)) it follows that

$$\begin{aligned} \frac{1}{y} \left| \int_0^y \psi_K(x/n_1) \psi_K(x/n_2) dx \right| &\ll \sum_{0 < |k_1|, |k_2| \leq K} \frac{1}{|k_1 k_2|} \left| \frac{1}{y} \int_0^y e\left(x \left(\frac{k_1}{n_1} + \frac{k_2}{n_2}\right)\right) dx \right| \\ &\ll \sum_{0 < |k_1|, |k_2| \leq K} \frac{1}{|k_1 k_2|} \min\left(1, \frac{1}{y|k_1/n_1 + k_2/n_2|}\right). \end{aligned}$$

From these two estimates it follows that the sum to be bounded is

$$\ll \frac{N^2}{K} + \sum_{0 < |k_1|, |k_2| \leq K} \frac{1}{|k_1 k_2|} \sum_{N < n_1, n_2 \leq 2N} \min\left(1, \frac{1}{y|k_1/n_1 + k_2/n_2|}\right).$$

To estimate the sum above, we split the terms into two groups: those with $|k_1/n_1 + k_2/n_2| \geq K/y$ and those terms with $|k_1/n_1 + k_2/n_2| < K/y$. The first group contributes

$$\ll \frac{N^2}{K} \sum_{0 < |k_1|, |k_2| \leq K} \frac{1}{|k_1 k_2|} \frac{1}{|k_1 k_2|} \ll \frac{N^2}{K} (\log K)^2.$$

Terms in the second group only exist for k_1 and k_2 of opposite sign, and here $|k_1 n_2 + k_2 n_1| \ll KN^2/y$, so that if k_1 , n_1 , and k_2 are fixed, then n_2 has $\ll 1 + KN^2/y$ choices. Therefore the second group contributes

$$\ll \left(1 + \frac{KN^2}{y}\right) N \sum_{0 < |k_1|, |k_2| \leq K} \frac{1}{|k_1 k_2|} \ll (\log K)^2 N \left(1 + \frac{KN^2}{y}\right).$$

Choosing $K = 2\lceil \sqrt{y/N} \rceil$, the lemma follows. \square

Proof of Theorem 7.1. From Theorem 1 and Lemma 1 of [8] (but beware of the changes in notation, especially that his saw tooth function differs from ours in sign) it follows that with $N = y \exp(-c\sqrt{\log y})$ for a suitable positive constant $c < 1$, one has

$$\tilde{R}(u) = - \sum_{n \leq N} \frac{\mu(n)}{n} \psi(u/n) + O(\exp(-c\sqrt{\log y})),$$

for all $N \leq u \leq y$. Since $\tilde{R}(u)$ and the sum over n above are $\ll \log y$, it follows that for $\ell \leq \frac{c}{9} \sqrt{\log y} / \log \log y$

$$(7.1) \quad \frac{1}{y} \int_0^y \tilde{R}(u)^\ell du = \frac{(-1)^\ell}{y} \int_0^y \left(\sum_{n \leq N} \frac{\mu(n)}{n} \psi(u/n) \right)^\ell du + O(\exp(-\frac{c}{2} \sqrt{\log y})).$$

Now applying (3.10) we see that

$$(7.2) \quad \begin{aligned} \frac{1}{y} \int_0^y \left(\sum_{n \leq N} \frac{\mu(n)}{n} \psi(u/n) \right)^\ell du &= \frac{1}{y} \int_0^y \left(\sum_{n \leq y^{1/(2\ell)}} \frac{\mu(n)}{n} \psi(u/n) \right)^\ell du \\ &+ O\left(\frac{\ell(\log y)^{\ell-1}}{y} \int_0^y \left| \sum_{y^{1/(2\ell)} \leq n \leq N} \frac{\mu(n)}{n} \psi(u/n) \right| du \right). \end{aligned}$$

Expanding out, the main term in (7.2) is

$$\begin{aligned} &\sum_{n_1, \dots, n_\ell \leq y^{1/(2\ell)}} \frac{\mu(n_1) \cdots \mu(n_\ell)}{n_1 \cdots n_\ell} \frac{1}{y} \int_0^y \prod_{j=1}^{\ell} \psi(u/n_j) du \\ &= \sum_{n_1, \dots, n_\ell \leq y^{1/(2\ell)}} \frac{\mu(n_1) \cdots \mu(n_\ell)}{n_1 \cdots n_\ell} (\mathcal{B}(n_1, \dots, n_\ell) + O(n_1 \cdots n_\ell)). \end{aligned}$$

Arguing as in the proof of Theorem 4.1, this may be seen to equal $M_R(\ell) + O(y^{-1/(40\ell \log \ell)})$.

As for the remainder term in (7.2), splitting the terms $y^{1/(2\ell)} \leq n \leq N$ into dyadic blocks, we may bound this by

$$\ll \exp\left(\frac{c}{8} \sqrt{\log y}\right) \max_{\substack{y^{1/(2\ell)} \leq M \leq N \\ I \subset [M, 2M]}} \frac{1}{y} \int_0^y \left| \sum_{n \in I} \frac{\mu(n)}{n} \psi(u/n) \right| du,$$

where the maximum is over subintervals I of $[M, 2M]$. By Cauchy-Schwarz and Lemma 7.2, this is

$$\ll \exp\left(\frac{\epsilon}{8}\sqrt{\log y}\right) \max_{\substack{y^{1/(2\epsilon)} \leq M \leq N \\ I \subset [M, 2M]}} (\log y) \left(\frac{1}{M} + \frac{\sqrt{M}}{\sqrt{y}}\right)^{\frac{1}{2}} \ll \exp\left(-\frac{\epsilon}{8}\sqrt{\log y}\right).$$

This justifies the first claim of the theorem. It is also clear that $M_R(\ell) = 0$ for odd ℓ , and the formula for $M_R(2)$ follows from our knowledge of $\mathcal{B}(n_1, n_2)$. Lastly, the claimed upper bound on $M_R(\ell)$ follows exactly as the upper bound for $M_C(\ell)$ in Theorem 4.1. \square

REFERENCES

- [1] T. M. Apostol. *Modular functions and Dirichlet series in number theory*, volume 41 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.
- [2] J. Bourgain and M. Z. Garaev. Sumsets of reciprocals in prime fields and multilinear Kloosterman sums. *Izv. Ross. Akad. Nauk Ser. Mat.*, 78(4):19–72, 2014.
- [3] S. Chowla. Contributions to the analytic theory of numbers. *Math. Z.*, 35(1):279–299, 1932.
- [4] A. Granville and K. Soundararajan. The distribution of values of $L(1, \chi_d)$. *Geom. Funct. Anal.*, 13(5):992–1028, 2003.
- [5] A. A. Karatsuba. New estimates for short Kloosterman sums. *Mat. Zametki*, 88(3):384–398, 2010.
- [6] M. A. Korolëv. On Karatsuba’s method of estimating Kloosterman sums. *Mat. Sb.*, 207(8):117–134, 2016.
- [7] R. J. Lemke Oliver and K. Soundararajan. Unexpected biases in the distribution of consecutive primes. *Proc. Natl. Acad. Sci. USA*, 113(31):E4446–E4454, 2016.
- [8] H. L. Montgomery. Fluctuations in the mean of Euler’s phi function. *Proc. Indian Acad. Sci. Math. Sci.*, 97(1-3):239–245 (1988), 1987.
- [9] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.
- [10] S. S. Pillai and S. D. Chowla. On the Error Terms in some Asymptotic Formulae in the Theory of Numbers (1). *J. London Math. Soc.*, S1-5(2):95.
- [11] I. Vardi. Dedekind sums have a limiting distribution. *Internat. Math. Res. Notices*, (1):1–12, 1993.
- [12] A. Walfisz. *Weylsche Exponentialsummen in der neueren Zahlentheorie*. Mathematische Forschungsberichte, XV. VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.
- [13] L. C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1997.

DEPARTMENT OF MATHEMATICS, TUFTS UNIVERSITY
E-mail address: robert.lemke_oliver@tufts.edu

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY
E-mail address: ksound@stanford.edu