MOEBIUS-WALSH CORRELATION BOUNDS AND AN ESTIMATE OF MAUDUIT AND RIVAT

J. BOURGAIN

ABSTRACT. We establish small correlation bounds for the Moebius function and the Walsh system, answering affirmatively a question posed by G. Kalai [Ka]. The argument is based on generalizing the approach of Mauduit and Rivat [M-R] in order to treat Walsh functions of ‘large weight’, while the ‘small weight’ case follows from recent work due to B. Green [Gr]. The conclusion is an estimate uniform over the full Walsh system. A similar result also holds for the Liouville function.

§0. Introduction

Fix a large integer $\lambda$ and restrict the Moebius function $\mu$ to the interval $[1, 2^\lambda] \cap \mathbb{Z} = \Omega$. Identifying $\Omega$ with the Boolean cube $\{0, 1\}^\lambda$ by binary expansion $x = \sum_{0 \leq j < \lambda} x_j 2^j$, the Walsh system $\{w_A : A \subset \{0, \ldots, \lambda - 1\}\}$ is defined by $w_\varphi = 1$ and

$$w_A(x) = \prod_{j \in A} (1 - 2x_j) = e^{i\pi \sum_{j \in A} x_j}. \quad (0.1)$$

The Walsh functions on $\Omega$ form an orthonormal basis (the character group of $(\mathbb{Z}/2\mathbb{Z})^\lambda$) and given a function $f$ on $\Omega$, we write

$$f = \sum_{A \subset \{0, \ldots, \lambda - 1\}} \hat{f}(A) w_A \quad (0.2)$$

where $\hat{f}(A) = 2^{-\lambda} \sum_{n \in \Omega} f(n) w_A(n)$ are the Fourier-Walsh coefficients of $f$. Understanding the size and distribution of those coefficients is well-known to be important to various issues, in particular in complexity theory and computer science. Roughly speaking, a $F - W$ spectrum which is ‘spread out’ indicates a high level of complexity for the function $f$. We do not elaborate on this theory here and refer the reader to

This research was partially supported by NSF grants DMS-0808042 and DMS-0835373.
the extensive literature on the subject; see also the preprint of B. Green [Gr], which motivated this Note.

Returning to the Moebius function and the so-called ‘Moebius randomness law’ it seems therefore reasonable to expect that $\mu|\Omega$ will have a $F-W$ spectrum that is not localized. More precisely, we establish the following uniform bound on its $F-W$ coefficients, answering affirmatively a question posed by G. Kalai.

**Theorem 1.** For $\lambda$ large enough,

$$\max_{A \subset \{0, \ldots, \lambda-1\}} \left| \sum_{n<2^\lambda} \mu(n)w_A(n) \right| < 2^{\lambda - \lambda^{1/10}}$$

(a similar estimate is also valid for the Liouville function).

The proof of (0.3) involves different arguments, depending on the size $|A|$. Roughly speaking, one distinguishes between the case $|A| = o(\sqrt{\lambda})$ and $|A| \gtrsim \sqrt{\lambda}$. In the first case, B. Green already obtained an estimate of the type (0.3), see [Gr]. Part of the technique used in [Gr] is borrowed from Harman and Katai’s work [H-K] on prescribing binary digits of the primes. Let us point out that in this range the problem of estimating the correlation of $\mu$ with a Walsh function is reduced to estimates on the usual Fourier spectrum of $\mu$ (by an expansion of $w_A$ in the trigonometric system). The latter is then achieved either by means of Dirichlet $L$-function theory (when the argument $\alpha$ is close to a rational $\frac{a}{q}$ with sufficiently small denominator $q$) or by Vinogradov’s estimate when $q$ is large. At the other end of the spectrum, when $A = \{0, \ldots, \lambda\}$, Mauduit and Rivat proved that

$$\left| \sum_{n<2^\lambda} \Lambda(n)\tilde{w}_A(n) \right| < 2^{(1-\varepsilon)\lambda}$$

for some $\varepsilon > 0$.

Here $\Lambda(n)$ stands for the Van Mangold function ([M-R]). Their motivation was the solution to a problem of Gelfond on the uniform distribution of the sum of the binary digits of the primes. Of course, their argument gives a similar bound for the Moebius
function as well. Thus
\[ \left| \sum_{n<2^\lambda} \mu(n)\widehat{w_{\{0,\ldots,\lambda-1\}}} (n) \right| < 2^{(1-\varepsilon)\lambda}. \]  
(0.5)

A remarkable feature of the [M-R] method is that the usual type-I, type-II sum approach in the study of sums
\[ \sum_{n<X} \Lambda(n)f(n) \quad \text{or} \quad \sum_{n<X} \mu(n)f(n) \]
is applied directly to \( f = w_{\{1,\ldots,\lambda\}} \) without an initial conversion to additive characters (as done in [H-K] and [Gr]). The main idea in what follows is to generalize the Mauduit-Rivat argument in order to treat all Walshes \( w_A \) provided \( A \) is not too small (the latter case being captured by [Gr]).

Needless to say, the \( 2^{-\lambda^{1/10}} \)-saving in (0.3) can surely be improved (this is an issue concerning the treatment of low-weight Walsh functions) and no effort has been made in this respect. We also observe that, assuming \( GRH \), (0.3) may be improved to

**Theorem 2.** Under \( GRH \), assuming \( \lambda \) large, we have
\[ \max_{A \subset \{0,\ldots,\lambda-1\}} \left| \sum_{n<2^\lambda} \mu(n)w_A(n) \right| < 2^{\lambda \left(1 - \frac{c}{(\log \lambda)^2}\right)}. \]  
(0.6)

We will assume the reader familiar with the basic technique, going back to Vinogradov, of type-I and type-II sums, to which sums \( \sum_{n<X} \mu(n)f(n) \) may be reduced; see [I-K] or [M-R]. In fact, we will rely here on the same version as used in [M-R] (see [M-R], Lemma 1). Otherwise, besides referring to the work of B. Green for \( |A| \) small, our presentation is basically self-contained. In particular, all the required lemmas pertaining to bounds on Fourier coefficients of Walsh functions are proven (they include estimates similar to those needed in [M-R] and also some additional ones) and are presented in \( \S 1 \) of the paper.
1. Estimates on Fourier coefficients of Walsh functions

For $A \subset \{0, \ldots, \lambda - 1\}$ and $x = \sum j x_j 2^j \in [1, 2^\lambda] \cap \mathbb{Z}$

$$w_A(x) = \prod_{j \in A} (1 - 2x_j) = e^{i\pi \sum_{j \in A} x_j} = \prod_{j \in A} h\left(\frac{x}{2^{j+1}}\right)$$  \quad (1.0)

where $h : \mathbb{R} \to \{1, -1\}$ is the 1-periodic function

$$
\begin{cases}
  h = 1 & \text{if } 0 \leq x < \frac{1}{2} \\
  h = -1 & \text{if } \frac{1}{2} \leq x < 1
\end{cases}
$$

For $x \in \mathbb{Z}$,

$$h\left(\frac{x}{2^j + 1}\right) = \sum_{|r| < 2^j+1} a_{r,j} e\left(\frac{rx}{2^{j+1}}\right) \text{ with } \sum |a_r| \lesssim j.$$

It follows that

**Lemma 1.** $w_A(x) = \sum_{k < 2^\lambda} \hat{w}_A(k) e\left(\frac{kx}{2}\right)$ with

$$\sum |\hat{w}_A(k)| < (\mathcal{C}\lambda)^{|A|}.$$  \quad (1.1)

From the second equality in (1.0), also

$$\hat{w}_A(k) = 2^{-\lambda} \sum_{\{x_j\}} e^{i\pi \sum_{j \in A} x_j} e^{2\pi i k \sum x_j 2^j} = \prod_{j \not\in A} \left(\frac{1 + e(k2^j - \lambda)}{2}\right) \prod_{j \in A} \left(\frac{1 - e(k2^j - \lambda)}{2}\right)$$

and

$$|\hat{w}_A(k)| = \prod_{j \not\in A} |\cos \pi k 2^j - \lambda| \prod_{j \in A} |\sin \pi k 2^j - \lambda|$$  \quad (1.2)

**Lemma 2.** $||\hat{w}_A||_\infty \lesssim 2^{-c|A|}$ for some constant $c > 0$.  \quad (1.3)

**Proof.** Use (1.2).

Taking some $i_0 \in A$ and assuming

$$\left| \sin \pi \frac{k}{2^\lambda - i_0} \right| \approx 1, \text{ hence } \left| \frac{k}{2^\lambda - i_0} - \frac{1}{2} \right| \approx 0$$
it follows that either
\[ \left\| \frac{k}{2^{\lambda-i_0-1}} - \frac{1}{4} \right\| \approx 0 \]
or
\[ \left\| \frac{k}{2^{\lambda-i_0-1}} - \frac{3}{4} \right\| \approx 0 \]
and in either case
\[ \left| \cos \pi \frac{k}{2^{\lambda-i_0-1}} \right|, \left| \sin \pi \frac{k}{2^{\lambda-i_0-1}} \right| \approx \frac{1}{\sqrt{2}}. \]
The conclusion follows from (1.2). \qed

In addition to (1.1), we have the bound

**Lemma 3.**
\[ \sum_{k < 2^{\lambda}} |\hat{w}_A(k)| \lesssim 2^{(\frac{3}{2}-c)\lambda}. \] \hspace{1cm} (1.4)

for some constant \( c > 0 \).

**Proof.** We have to estimate
\[ \sum_{k \in \mathbb{Z}/2^{\lambda} \mathbb{Z}} \prod_{i \leq \lambda} \left| \cos \pi \left( \frac{u_i}{2} + \frac{k}{2^{\lambda-i}} \right) \right| \] \hspace{1cm} (1.5)
where \( u_i = 1 \) if \( i \in A \) and \( u_i = 0 \) if \( i \notin A \).

Perform a shift \( k \to k + c2^{\lambda-2} + d2^{\lambda-1} \) with \( c, d = 0, 1 \).

This gives
\[ \sum_{k \in \mathbb{Z}/2^{\lambda-2} \mathbb{Z}} \prod_{2 \leq i \leq \lambda} \left| \cos \pi \left( \frac{u_i}{2} + \frac{k}{2^{\lambda-i}} \right) \right|. \] \hspace{1cm} (*)
with
\[
(*) = \frac{1}{4} \sum_{c,d=0,1} \left| \cos \pi \left( \frac{u_0}{2} + \frac{k}{2^\lambda} + \frac{c}{4} + \frac{d}{2} \right) \right| \left| \cos \pi \left( \frac{u_1}{2} + \frac{k}{2^\lambda - 1} + \frac{c}{2} \right) \right|
\]
\[= \frac{1}{4} \sum_{c=0,1} \left( \left| \cos \pi \left( \frac{u_0}{2} + \frac{k}{2^\lambda} + \frac{c}{4} \right) \right| + \left| \sin \pi \left( \frac{u_0}{2} + \frac{k}{2^\lambda} + \frac{c}{4} \right) \right| \right) \left| \cos \pi \left( \frac{u_1}{2} + \frac{k}{2^\lambda - 1} + \frac{c}{2} \right) \right|
\]
\[= \frac{1}{4} \left\{ (|\cos \phi| + |\sin \phi|) \cdot \left| \cos \left( \frac{\pi u_1}{2} + 2\phi \right) \right| + \frac{1}{\sqrt{2}} (|\cos \phi - \sin \phi| + |\sin \phi + \cos \phi|) \cdot \left| \sin \left( \frac{\pi u_1}{2} + 2\phi \right) \right| \right\}
\]
where \( \phi = \pi \left( \frac{u_0}{2} + \frac{k}{2^\lambda} \right) \). Clearly
\[
(1.6) \leq \frac{1}{4} \left\{ (1 + \sin 2\phi) \frac{1}{\sqrt{2}} \left| \cos \left( 2\phi \right) \right| + (1 + \cos 2\phi) \frac{1}{\sqrt{2}} \left| \sin \left( 2\phi \right) \right| \right\}
\]
\[\leq \frac{1}{4} \sqrt{2 + \sqrt{2}}.
\]
Iterating, we obtain the bound
\[
\leq \left( \sqrt{2 + \sqrt{2}} \right)^{\lambda/2}
\]
and hence (1.4). \( \square \)

**Lemma 4.** Let \( r < \lambda, a = 0, 1, \ldots, 2^r - 1 \). Then
\[
\sum_{k \equiv a (mod \, 2^r)} |\hat{w}_A(k)| \lesssim 2^{(\frac{1}{2} - c)(\lambda - r)}.
\] (1.7)

**Proof.** Writing \( k = a + 2^r k_1 \) with \( k_1 < 2^{\lambda - r} \),
\[
|\hat{w}_A(k)| = \prod_{i < \lambda - r} \left| \cos \pi \left( \frac{u_i}{2} + \frac{a}{2^\lambda - i} + \frac{k_1}{2^\lambda - i - r} \right) \right| \prod_{i \geq \lambda - r} \left| \cos \pi \left( \frac{u_i}{2} + \frac{a}{2^\lambda - i} \right) \right|
\]
\[\leq \prod_{i < \lambda - r} \left( \left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2^\lambda - r - i} \right) \right| + 2^{-\lambda + i + r} \right).
\] (1.8)
For fixed $k_1$, denote

$$B(k_1) = \left\{ i < \lambda - r; \left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right| < \left( \frac{1}{\sqrt{2}} \right)^{\lambda - r - i} \right\}$$

Hence, if $i \not\in B_{k_1}$

$$\left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right| < \left( 1 + \left( \frac{1}{\sqrt{2}} \right)^{\lambda - r - i} \right) \left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right|$$

and if $i \in B_{k_1}$

$$\left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right| < \left( 1 + 2 \left( \frac{1}{\sqrt{2}} \right)^{\lambda - r - i} \right) \left| \sin \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right|.$$

Thus certainly

$$|\widehat{w}_A(k)| \lesssim \sum_{B \subset \{0, 1, \ldots, \lambda - r - 1\}} \left( \frac{1}{\sqrt{2}} \right)^{\sum_{i \in B} (\lambda - r - i)} \prod_{i \in B, i < \lambda - r} \left| \cos \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right| \prod_{i \in B} \left| \sin \pi \left( \frac{u_i}{2} + \frac{k_1}{2\lambda - r - i} \right) \right|. \tag{1.9}$$

Given $B \subset [0, \lambda - r - 1]$, define $B_1 \subset [0, \lambda - r - 1]$ as

$$B_1 = (B \cap [u_i = 0]) \cup (B^c \cap [u_i = 1]).$$

Hence

$$(1.9) = \sum_B \left( \frac{1}{\sqrt{2}} \right)^{\sum_{i \in B} (\lambda - r - i)} |\widehat{w}_{B_1}(k_1)|. \tag{1.10}$$

Summation of (1.10) over $k_1 < 2^\lambda - r$ and using the bound (1.4) with $\lambda$ replaced by $\lambda - r$ clearly gives (1.7)

Next, we also need the following ‘approximation property’ for shifts

**Lemma 5.** Let $A \subset [\lambda - \sigma, \lambda] \cap \mathbb{Z}$.

Then

$$\sum_{k < 2^\lambda} |\widehat{w}_A(k)| < C^{(\log \lambda)^2} (2\sigma)^{\frac{1}{2} - c}. \tag{1.11}$$
Moreover, there is a bounded function $W_A$ on $[0, \lambda] \cap \mathbb{Z}$ satisfying $|\hat{W}_A| \leq |\hat{w}_A|$ and

\begin{equation}
2^{-\lambda} \sum_{x < 2^\lambda} |W_A(x) - w_A(x)|^2 \left( \frac{1}{2} \right)^{1/2} < 2^{-c} \tag{1.12}
\end{equation}

\begin{equation}
\hat{W}_A(k) = 0 \text{ if } |k| > 2^{\sigma+t} \tag{1.13}
\end{equation}

Here $t \in \mathbb{Z}$ is a parameter satisfying $C(\log \lambda)^2 < t < \frac{1}{2} (\lambda - \sigma)$.

**Proof.** Writing $k = k_0 + 2^\sigma k_1$ with $k_0 < 2^\sigma, |k_1| < 2^{\lambda-\sigma-1}$ and setting again $u_i = 1$ if $i \in A$, $u_i = 0$ if $i \notin A$, we obtain

\begin{equation}
|\hat{w}_A(k)| = \prod_{i < \lambda - \sigma} \cos \pi \left( \frac{k_0 + 2^\sigma k_1}{2^{\lambda-i}} \right) \cdot \prod_{\lambda - \sigma \leq i < \lambda} \cos \pi \left( \frac{u_i + k_0}{2^{\lambda-i}} \right) \tag{1.14}
\end{equation}

\begin{equation}
= (1.14).|\hat{w}_{A-\lambda+\sigma}(k_0)|. \tag{1.15}
\end{equation}

where

\begin{equation}
A - \lambda + \sigma \subset [0, \sigma] \cap \mathbb{Z}.
\end{equation}

We treat (1.14) as in the proof of Lemma 4, obtaining a bound

\begin{equation}
|(1.14)| < \sum_{B \subset \{0,1,\ldots,\lambda-\sigma-1\}} \left( \frac{1}{\sqrt{2}} \right)^{\sum_{i \in B} (\lambda - \sigma - i)} |\hat{w}_B(k_1)|. \tag{1.16}
\end{equation}

From (1.1), certainly

\begin{equation}
\sum_{k_1 < 2^{\lambda-\sigma}} |\hat{w}_B(k_1)| < (C\lambda)^{|B|} \tag{1.17}
\end{equation}

and substitution of (1.17) in (1.16) implies by (1.15)

\begin{equation}
\|\hat{w}_A\|_1 \leq \|\hat{w}_{A-\lambda+\sigma}\|_1 \cdot \sum_B \left( \frac{1}{\sqrt{2}} \right)^{\sum_{i \in B} (\lambda - \sigma - i)} (C\lambda)^{|B|} \tag{Lemma 3}
\end{equation}

\begin{equation}
< (2^\sigma)^{\frac{1}{2}-c} C(\log \lambda)^2 \tag{1.11}
\end{equation}

which is (1.11).
Next, let \( C(\log \lambda)^2 < \rho < \frac{1}{2}(\lambda - \sigma) \) and estimate

\[
\sum_{k_1} \min_{B \leq \lambda - \sigma - \rho} \left( \frac{1}{\sqrt{2}} \right)^{\frac{\lambda - \sigma - i}{2}} |\hat{w}_B(k_1)| \lesssim 2^{-\rho/4}. \tag{1.18}
\]

If

\[
B \subset [\lambda - \sigma - \rho, \lambda - \sigma]
\]

we establish a bound on \( \hat{w}_B(k_1) \). Write

\[
|\hat{w}_B(k_1)| = \prod_{i < \lambda - \sigma - \rho} \left| \cos \pi \frac{k_1}{2^{\lambda - \sigma - i}} \right| \prod_{\lambda - \sigma - \rho \leq i < \lambda - \sigma} \left| \cos \pi \left( \frac{v_i}{2} + \frac{k_1}{2^{\lambda - \sigma - i}} \right) \right|
\]

with \( v_i = 0, 1 \) if \( i \notin B, i \in B \). Hence, for \( 4^\rho < k_1 < 2^{\lambda - \sigma - 1} \)

\[
|\hat{w}_B(k_1)| \leq \prod_{\rho < j \leq \lambda - \sigma} \left| \cos \frac{v_j}{2^j} \frac{k_1}{2^j} \right| < k_1^{-c} \tag{1.20}
\]

for some \( c < 0 \), as we verify by dyadic expansion of \( k_1 \).

It follows that for \( 4^\rho \leq K_1 < 2^{\lambda - \sigma} \)

\[
\sum_{K_1 < |k_1| < 2^{\lambda - \sigma}} \left\{ \sum_{B(1.19)} \left( \frac{1}{\sqrt{2}} \right)^{\frac{\lambda - \sigma - i}{2}} |\hat{w}_B(k_1)| \right\}^2 <
\]

\[
< C \sum_{B(1.19)} \sum_{K_1 < |k_1| < 2^{\lambda - \sigma}} \left( \frac{1}{\sqrt{2}} \right)^{\frac{\lambda - \sigma - i}{2}} |\hat{w}_B(k_1)|^2 \]

\[
< K_1^{-c} \sum_{B} \left( \frac{1}{\sqrt{2}} \right)^{\frac{\lambda - \sigma - i}{2}} \| \hat{w}_B \|_1 \]

\[
< K_1^{-c} C(\log \lambda)^2. \tag{1.21}
\]

Define \( W_A \) as Fourier restriction of \( w_A \). More specifically, let

\[
W_A(x) = \sum \eta(k)\hat{w}_A(k) e \left( \frac{kx}{2^\lambda} \right) \tag{1.22}
\]

where \( \eta : \mathbb{R} \to [0, 1] \) is trapezoidal with \( \eta(z) = 1 \) for \( |z| < K_1 2^\sigma, \eta(z) = 0 \) for \( |z| \geq 2K_1 2^\sigma \). Hence \( \|W_A\|_\infty \leq 3 \) and \( \hat{w}_A(k) = \hat{w}_A(k) \) for \( |k| \leq K_1 2^\sigma, \hat{w}_A(k) = 0 \) for \( |k| \geq 2K_1 2^\sigma \).
From the preceding
\[
\|\hat{W}_A - \hat{w}_A\|_2^2 \leq \sum_{k_0 < 2^\sigma} |\hat{W}_{A-\lambda - \sigma}(k_0)|^2 \sum_{K_1 \leq |k| < 2^\lambda - \sigma} (1.16)^2
\]
\[
\leq 2^{-\rho/2 + K_1^{-c} C^2 (\log \lambda)^2}.
\]
Taking \(K_1 = 2^{t-1}\), \(\rho = t^2\), Lemma 5 follows. \(\square\)

The role of \(W_A\) is to provide a substitute for \(w_A\) with localized Fourier transform.

**Lemma 6.** If \(J \subset [1, 2^\lambda[\) is an interval, there is a bound
\[
\sum_{k \in J} |\hat{w}_A(k)| \lesssim |J|^{1/2 - c}.
\]

**Proof.** Write
\[
|\hat{w}_A(k)| = \prod_{i < \lambda} \left| \cos \left( \frac{u_i}{2} + \frac{k}{2^\lambda - i} \right) \right|
\]
with \(u_i = 0 (u_i = 1)\) if \(i \notin A (i \in A)\).

Assume \(2^m \sim |J| < 2^m\). Obviously
\[
|\hat{w}_A(k)| \leq \prod_{\lambda - m \leq i < \lambda} \left| \cos \left( \frac{u_i}{2} + \frac{k}{2^\lambda - i} \right) \right| = \prod_{0 \leq i_1 < m} \left| \cos \left( \frac{u_{i_1 + \lambda - m}}{2} + \frac{k}{2m - i_1} \right) \right|
\]
\[
= |\hat{w}_{A_1}(k)|
\]
where
\[
A_1 = \{0 \leq i_1 < m; i_1 \in A + m - \lambda \}.
\]
Hence, since \(\hat{w}_{A_1}\) is \(2^m\)-periodic
\[
\sum_{k \in J} |\hat{w}_A(k)| \leq \sum_{k \in J} |\hat{w}_{A_1}(k)| \leq \sum_{k < 2^m} |\hat{w}_{A_1}(k)| \leq ||\hat{w}_{A_1}||_1 < 2^m(1/2 - c)
\]
by Lemma 3. \(\square\)

2. **Type-II sums**

Let \(X = 2^\lambda\), \(S \subset \{0, \ldots, \lambda - 1\}\), \(w_S(x) = \prod_{i \in S} (1 - 2x_i)\) with \(x = \sum x_i 2^i\).

Specify ranges \(M \sim 2^\mu, N \sim 2^\nu\) such that \(M \leq N\) and \(M.N \sim X\).
Our goal is to bound bilinear sums of the form \( \sum_{m \sim M} \alpha_m \beta_n w_S(m, n) \), where \(|\alpha_m|, |\beta_n| \leq 1\) are arbitrary coefficients.

We fix a relatively small dyadic integer \( L = 2^\rho \) (to be specified). We assume \( \rho < \frac{\mu}{100} \), noting that otherwise our final estimate (2.29) is trivial.

Following [M-R], we proceed with the initial reduction of the problem, crucial to our analysis.

Estimate
\[
\left| \sum_{m \sim M} \sum_{n \sim N} \alpha_m \beta_n w_S(m, n) \right| \leq \sum_{m \sim M} \left| \sum_{n \sim N} \beta_n w(m, n) \right| .
\] (2.1)

Fix \( K \), such that \( L2^K < N \) and write using Cauchy’s inequality
\[
\left| \sum_{n \sim N} \beta_n w(m, n) \right| \leq \frac{1}{L} \sum_{n \sim N} \left| \sum_{\ell = 1}^L \beta_{n + \ell 2^K} w(m + \ell 2^K) \right|
\]
\[
\left( \sum_{n \sim N} \beta_n w(m, n) \right)^2 \lesssim \frac{N}{L} \left[ \sum_{|\ell| < L} \left| \sum_{n \sim N} \beta_n \overline{\beta}_{n + \ell 2^K} w(m, n) w(m + \ell 2^K) \right| \right].
\]

Hence, by another application of Cauchy’s inequality, we obtain
\[
(2.1)^2 \lesssim \frac{M.N}{L} \sum_{|\ell| < L} \left| \sum_{m \sim M} \sum_{n \sim N} w_S(m, n) w_S(m + \ell 2^K) \right| .
\] (2.2)

Comparing the binary expansions of \( mn \) and \( mn + \ell m 2^K \), the \( K \) first digits remain and we can assume that also digits \( j > K + \mu + \rho + \varepsilon \rho \) are unchanged provided in (2.2) we introduce an additional error term of the order \( 2^{-\varepsilon \rho} M^2 N^2 \) (cf. Lemma 5 in [M-R]). Here \( \varepsilon > 0 \) remains to be specified and we assume \( \varepsilon \rho \in \mathbb{Z}_+ \).

Therefore we may write, up to above error
\[
w_S(mn)w_S(m(n + \ell 2^K)) \quad \Rightarrow \quad w_{S'}(mn)w_{S'}(m(n + \ell 2^K))
\]
with
\[
S' = S \cap [K, K + \mu + \rho'] \quad \text{and} \quad \rho' = (1 + \varepsilon) \rho
\]
and in (2.2) we may replace \( w = w_S \) by \( w_{S'} \).

We will either choose \( K = 0 \) or \( \mu - \rho \leq K < \lambda - \mu - \rho \). Hence, by varying \( K \), the intervals \([K, K + \mu + \rho]\) will cover \([0, \lambda]\).

For \( K \neq 0 \), we approximate \( w_{S'} \) by \( W_{S'} \) given by Lemma 5, applied with \( \lambda \) replaced by \( K + \mu + \rho' \) and \( \sigma \) by \( \mu + \rho' \).

Take \( t = \varepsilon \rho \) where \( \rho \) is certainly assumed to satisfy

\[
\frac{\mu}{100} > \rho \gg (\log \lambda)^2.
\]

Thus from (1.12)

\[
\sum_{x<X} |w_{S'}(x) - W_{S'}(x)|^2 < 2^{-ct}X.
\]

From the preceding (since \( W_{S'} \) is bounded)

\[
(2.2) \lesssim \frac{X}{L} \sum_{\substack{n \sim N \atop 0 < \ell < L}} \left| \sum_{m \sim M} W_{S'}(m.n)W_{S'}(m(n + \ell 2^K)) \right| + X \sum_{\substack{m \sim M \atop n \sim N}} |w_{S'}(mn) - W_{S'}(m.n)| + X^2 L^{-\epsilon} \tag{2.3}
\]

where

\[
(2.4) < X \left( \sum_{x<X} |w_{S'}(x) - W_{S'}(x)|^2 \right)^{\frac{1}{4}} \left( \sum_{x<X} d(x)^2 \right)^{\frac{1}{4}} < L^{-c\epsilon} X^2 (\log X)^C < L^{-c\epsilon} X^2.
\]

For \( K = 0 \)

\[
w_{S'}(x) = \sum_{k<2^{\mu+\rho'}} \tilde{w}_{S'}(k)e\left(\frac{kx}{2^{\mu+\rho'}}\right) \tag{2.5}
\]

where, from Lemma 2 and Lemma 3 applied with \( \lambda \) replaced by \( \mu + \rho' \)

\[
\|\tilde{w}_{S'}\|_\infty < 2^{-c|S'|} \tag{2.6}
\]

and

\[
\|\tilde{w}_{S'}\|_1 < 2^{\left(\frac{1}{2}-c\right)(\mu+\rho')} < 2^{\left(\frac{1}{2}-c\right)(\mu+\rho)} \tag{2.7}
\]
for $\varepsilon$ small enough.

For $K \neq 0$,

$$W_S'(x) = \sum_{|k|<2^{\mu+\rho'+t}} \hat{W}_S'(k) e \left( \frac{kx}{2^{\mu+\rho'+K}} \right)$$

(2.8)

where

$$\|\hat{W}_S'\|_\infty \leq \|\hat{w}_S'\|_\infty < 2^{-c|S'|}$$

(2.9)

and by (1.11) and our choice of $\rho$

$$\|\hat{W}_S'\|_1 < 2^{(\frac{1}{2}-c)(\mu+\rho)}.$$ 

(2.10)

Denoting by $w$ either $w_S'$ when $K = 0$ or $W_S'$ for $\mu + \rho \leq K < \lambda - \mu - \rho$, substitution of (2.5), (2.8) and applying a smoothened $m$-summation gives for (2.3), with $M_1 = M^{1-\varepsilon_1}$

$$\frac{M^2N^2}{L} \sum_{|\ell| \leq L} \sum_{n \sim N} |\hat{w}(k)| |\hat{w}(k')| 1 \left[ \| \frac{k \ell}{2^{\mu+\rho'+K}} \| < \frac{M_1}{\lambda} \right]$$

(2.11)

up to a negligible error term.

The condition

$$\left\| \frac{(k-k')n}{2^{\mu+\rho'+K}} - \frac{k'\ell}{2^{\mu+\rho'}} \right\| < \frac{1}{M_1}$$

(2.12)

has to be analyzed.

For $k = k'$ the contribution is

$$\frac{M^2N^2}{L} \sum_{|\ell| \leq L} \sum_{|k|<2^{\mu+\rho'+t}} |\hat{w}(k)|^2 1 \left[ \| \frac{k \ell}{2^{\mu+\rho'}} \| < \frac{M_1}{\lambda} \right].$$

(2.13)

The $\ell = 0$ contribution in (2.2) is at most $\frac{M^2N^2}{L}$.

For $\ell \neq 0$, we get a bound

$$M^{2+\varepsilon_1}N^2L^2 \|\hat{w}\|_\infty^2 < M^2N^2L^2 \|\hat{w}\|_\infty^2 < X^2L^22^{-c|S'|}$$

(2.14)

from (2.6), (2.9) and choosing $\varepsilon_1 > 0$ small enough to ensure $\varepsilon_1 \lambda < \varepsilon \rho$.

In the sequel, we assume $k \neq k'$, $\ell \neq 0$. 
Also, if in (2.11) for given \( k, k', \ell \) there are at most \( O(1) \) values of \( n \) satisfying (2.12), the resulting contribution is at most

\[
M^2 N \| \hat{w} \|_1^2 \leq M^2 N (ML)^{1-2\epsilon} < X^2 LN^{-c} \tag{2.15}
\]

since \( M \leq N \).

Returning to (2.11), consider first the case \( K = 0 \).

We estimate the contribution for

\[
(k - k', 2^{\mu + \rho'}) = 2'.
\]

Thus \( k - k' = k_1 2^r, (k_1, 2) = 1 \) and (2.12) becomes

\[
\left\| \frac{k_1 n}{2^{\mu + \rho' - r}} - \frac{k' \ell}{2^{\mu + \rho'}} \right\| < \frac{1}{M_1} \tag{2.16}
\]

implying also

\[
\left\| \frac{k' \ell}{2^r} \right\| < \frac{L^{1+2\epsilon}}{2^r}. \tag{2.17}
\]

It follows from (2.17) that there are at most \( L^{1+2\epsilon} \) possibilities for \( k' \) (mod \( 2^r \)) and hence for \( (k, k') \) (mod \( 2^r \)).

For fixed \( k, k', \ell \), (2.16) determines \( n \) (mod \( 2^{\mu + \rho' - r} \)) up to \( 1 + L^{1+2\epsilon} 2^{-r} \) possibilities and hence \( n \) up to \( \sum_{a \equiv k (\text{mod } 2^r)} (1 + L^{1+2\epsilon} 2^{-r}) \) possibilities.

Thus the corresponding contribution to (2.11) is at most

\[
\frac{M^2 N}{L} \sum_{|\ell| \leq L} L^{1+2\epsilon} N 2^r \sum_{k \equiv a (\text{mod } 2^r)} \sum_{k' \equiv a (\text{mod } 2^r)} |\hat{w}(k)||\hat{w}(k')| \leq M N^2 (L + 2^r)L^{2\epsilon} \max_a \left[ \sum_{k < 2^{\mu + \rho'}} \left| \hat{w}(k) \right|^2 \right]. \tag{2.18}
\]
MOEBIUS-WALSH CORRELATION BOUNDS AND AN ESTIMATE OF MAUDUIT AND RIVAT

From Lemma 4 applied with \( \lambda \) replaced by \( \mu + \rho' \)

\[
\begin{align*}
(2.18) & \lesssim MN^2(L + 2r)(2^{\mu + \rho' - r})^{1-c}L^{2\epsilon} \\
& = M^2N^2(L^22^{-r} + L)(ML2^{-r})^{-c}L^{3\epsilon}.
\end{align*}
\]

Hence, assuming

\[
ML2^{-r} > L^C
\]

we obtain the bound

\[
\frac{X^2}{L}.
\]

Next, assume

\[
ML2^{-r} < L^C.
\]

From the preceding, there are at most \( L^{1+4\epsilon}(ML2^{-r})^2 < L^C \) possibilities for \((k,k')\).

This gives the contribution

\[ M^2N^2L^C\|\hat{w}\|_\infty^2 < L^C X^2 2^{-c|S'|} \]

and in conclusion (\( K = 0 \)) the bound

\[ X^2(L^{-1} + L^C 2^{-c|S'|}). \]

Next, assume

\[
K \geq \mu - \rho.
\]

Return to (2.11). Fix \( \ell, k, k' \) with \(|k - k'| \sim \Delta k < ML^2 \). Letting \( n \) range over an interval of size \( \frac{ML^2K}{\Delta k} \), the number of possibilities for \( n \) in that interval is at most

\[ 1 + \frac{L^{1+2\epsilon}2^K}{\Delta k}. \]

Assume

\[
N \gtrsim \frac{ML^2K}{\Delta k}.
\]

The number of \( n \)'s satisfying (2.12) is at most (since \( L2^K \geq M > \frac{\Delta k}{L^2} \) by (2.23))

\[
\frac{N\Delta k}{ML^2K} \left( 1 + \frac{L^{1+2\epsilon}2^K}{\Delta k} \right) < \frac{N}{M}L^2.
\]
This gives the contribution in (2.11)

\[
L^2MN^2\|\hat{w}\|^2_1 < L^2MN^2(ML^2)^{1-c} < X^2L^3M^{-c}.
\]

(2.24)

Next, assume

\[ N \ll \frac{ML^2K}{\Delta k}. \]

From (2.12), for \( \ell, k, k' \) given, there are at most

\[ 1 + \frac{2^kL^3}{\Delta k} \sim \frac{2^kL^3}{\Delta k} \]

values of \( n \).

Also

\[
\left\| \frac{k'\ell}{2^{\mu+\rho'}} \right\| < \frac{1}{M_1} + \frac{\Delta k.N}{M.2^{\rho'.2^k}}.
\]

Since \( |k'| < 2^{\mu+\rho}L^2 \), there is some integer \( \ell_1, |\ell_1| < L^2 \) s.t.

\[
\left| \frac{k'\ell}{2^{\mu+\rho'}} - \ell_1 \right| < \frac{1}{M_1} + \frac{\Delta k.N}{M.2^{\rho'.2^k}}
\]

hence

\[
\left| k' - \ell_1 \frac{2^{\mu+\rho'}}{\ell} \right| < L^{1+2^\epsilon} + \frac{\Delta k.N}{2^k}.
\]

This restricts \( k' \) to at most \( L^2 \) intervals of size \( L^{1+2^\epsilon} + \frac{\Delta k.N}{2^k} \).

Using Lemma 6, we obtain the following bound for the contribution to (2.11)

\[
M^2NL^2 \left( L^{1+2^\epsilon} + \frac{\Delta k.N}{2^k} \right)^{1-c} \frac{2^kL^3}{\Delta k} \lesssim
\]

\[
\frac{M^2NL^22^k}{\Delta k} + M^2N^2L^5 \left( \frac{\Delta k.N}{2^k} \right)^{-c} < M^2N^2L^7 \left( \frac{2^k}{N.\Delta k} \right)^c.
\]

(2.25)

If we assume

\[ \frac{N.\Delta k}{2^k} > L^c \]

(2.25) gives the bound

\[
\frac{X^2}{L}.
\]

(2.26)
Assume next
\[ \frac{N \Delta k}{2^k} < L^C. \]

From the preceding, \( k' \) is restricted to \( L^C \) values and the corresponding contribution to (2.11) is bounded by
\[ M^2 N^2 L^C \| \hat{w} \|_\infty^2 < X^2 L^C 2^{-c |S'|}. \]  

(2.27)

Collecting previous bounds gives
\[ (2.11) < X^2 \left( \frac{1}{L} + L^3 M^{-c} + L^C 2^{-c |S'|} \right) \]  

(2.28)

and recalling (2.3), (2.4)
\[ (2.1) < X \left( L^{-c \epsilon} + L^2 M^{-c} + L^C 2^{-c |S'|} \right). \]  

(2.29)

In the estimate (2.29), \( S' \) depends on the choice of \( K \).

Recall that either \( K = 0 \) or \( \mu - \rho \leq K < \lambda - \mu - \rho \) and hence, varying \( K \), the intervals \([K, K + \mu + \rho]\) will cover \([0, \lambda - 1]\). Thus we may choose \( K \) as to ensure that
\[ |S'| \geq \max |S \cap J| \geq \frac{\mu}{\lambda} |S| \]  

(2.30)

with max taken over intervals \( J \subset [0, \lambda - 1] \) of size \( \mu \), in particular (2.29) implies
\[ (2.1) < X \left( L^{-c \epsilon} + L^2 M^{-c} + L^C 2^{-c \frac{\mu}{\lambda} |S|} \right) \]  

(2.31)

where \( L \) is a parameter.

For \( |S| \leq \frac{\lambda^{1/2}}{H} \) with \( H \gg 1 \) a parameter, we apply B. Green’s estimate (see [Gr])
\[ \left| \sum_{x<2^k} w_S(x) \mu(x) \right| < \lambda e^{-cH}. \]  

(2.32)

Thus we assume \( |S| > \frac{\lambda^{1/2}}{H} \). Taking \( L = 2^H \), it follows from (2.29), (2.31) that
\[ (2.1) \lesssim X 2^{-c \epsilon H} \]  

(2.33)

assuming either that
\[ M > 2^{CH^2 \lambda^{1/2}} \]  

(2.34)
or

\[ M > C^H \text{ and } |S'| > CH \left( S' \text{ satisfying (2.30))}. \]  

(2.35)

3. Type-I sums and conclusion

We use Lemma 1 from [M-R] but treat also some of the type-I sums as type-II sums. Indeed, according to (2.33), (2.34), only the range \( M < C H^2 \lambda^{1/2} \) remains to be treated.

Thus we need to bound

\[
\left| \sum_{m \sim M} \left| \sum_{n \sim N} w_S(mn) \right| \right| \leq N \sum_{m \sim M} \sum_{k \sim X} \left| \hat{w}_S(k) \right| \left| \sum_{n \sim N} e \left( \frac{kmn}{2^H} \right) \right| 
\]

(3.1)

where \( M N \sim X = 2^\lambda, M < C H^2 \lambda^{1/2} \). We assume \( |S| > \frac{\lambda^{1/2}}{H} \).

Expanding in Fourier and using a suitable mollifier in the \( n \)-summation, we obtain

\[
(3.1) \leq N \sum_{m \sim M} \sum_{k \sim X} \left| \hat{w}_S(k) \right| \left| \sum_{n \sim N} e \left( \frac{kmn}{2^H} \right) \right| 
\]

\[
< N \sum_{m \sim M} \sum_{k \sim X} \left| \hat{w}_S(k) \right| 1_{\left[ \left| \frac{kmn}{2^H} \right| < \frac{\lambda^2}{N} \right]} + o(1) \quad (3.2)
\]

\[
< NM^2 \lambda^2 ||\hat{w}_S||_\infty \quad (3.2')
\]

\[
< XM2^{-c\lambda^{1/2}H^{-1}} \lambda^2. \quad (3.3)
\]

Taking \( H < \lambda^{1/10} \), (3.3) is certainly conclusive if \( M < C^H \). Hence recalling (2.35), we can assume that

\[ \mu > H \text{ and } \max |S \cap J| < CH \]  

(3.4)

for any interval \( J \subset \{0, \ldots, \lambda - 1\} \) of size \( \mu \), where \( M \sim 2^\mu \).

Assumption (3.4) will provide further information on \( \hat{w}_S \) that will be useful in exploiting (3.2).

Write

\[ S = S_1 \cup S_2 \]

where \( S_1 = S \cap [0, \lambda - 2\mu] \) and \( S_2 = S \cap [\lambda - 2\mu, \lambda] \). Hence by (3.4),

\[ |S_2| < CH. \]
Thus

\[ w_{S_2}(x) = \prod_{j \in S_2} h\left(\frac{x}{2^{j+1}}\right) \]

\[ = \sum_{k_2 \in \mathcal{A}_2} \hat{w}_{S_2}(k_2) e\left(\frac{k_2 x}{2^\lambda}\right) + O_L(2^{-H}) \]  \hspace{1cm} (3.5)

where the set \( \mathcal{A}_2 \) may be taken of size

\[ |\mathcal{A}_2| < 2^{H|S_2|} < C^{H^2} \]  \hspace{1cm} (3.6)

(obtained by truncation of the Fourier expansion of \( h \)).

On the other hand

\[ w_{S_1}(x) = \sum_{k_1 < 2^{\lambda-2\mu}} \hat{w}_{S_1}(k_1) e\left(\frac{k_1 x}{2^{\lambda-2\mu}}\right) \]

and hence

\[ w_S(x) = \sum_{k_1 < 2^{\lambda-2\mu}} \hat{w}_{S_1}(k_1)\hat{w}_{S_2}(k_2) e\left(\frac{2^{2\mu}k_1 + k_2}{2^\lambda} x\right) + O_L(2^{-H}). \]  \hspace{1cm} (3.7)

The bound (3.2) becomes now

\[ N \sum_{m \sim M} \max_{\substack{k_1 < 2^{\lambda-2\mu} \\ k_2 \in \mathcal{A}_2}} \left| \hat{w}_{S_1}(k_1) \right| \left| \hat{w}_{S_2}(k_2) \right| \frac{1}{\left[ \left[ 2^{2\mu}k_1 + k_2 \right] \left[ 2^{4\mu}k_1 + k_2 \right] \right]^{\lambda^2 N} \left( 2^{\lambda^2} \right)^2} \]

< \frac{N|\mathcal{A}_2| \left\| \hat{w}_{S_1} \right\|_\infty \max_{\substack{k_2 \in \mathcal{A}_2}} \sum_{m \sim M} \left\{ k_1 < 2^{\lambda-2\mu}; \right. \left. \frac{2^{2\mu}k_1 + k_2}{2^\lambda m} < \frac{\lambda^2}{N} \right\}}{\left( 2^{\lambda^2} \right)^2}. \]  \hspace{1cm} (3.8)

Clearly

\[ \sum_{m \sim M} \left\{ k_1 < 2^{\lambda-2\mu}; \frac{k_1 m}{2^{\lambda-2\mu}} < \frac{2\lambda^2}{N} \right\} = \mu.M \]

and therefore, since \(|S_1| \gtrsim \frac{\lambda^{1/2}}{NM} \) and (3.6)

\[ (3.8) < \mu C^{H^2} 2^{-c\lambda^{1/2} H^{-1}} NM \]
\[ < 2^{-c\lambda^{1/2}H^{-1}}X. \] (3.9)

From (2.33) and (3.9), we can claim a uniform bound
\[ \left| \sum_{x < X} \mu(x)w_S(x) \right| \lesssim X.2^{-c\lambda^{1/10}} \] (3.10)

hence obtaining Theorem 1.

Under GRH, (3.10) can be improved of course.

First, from a result due to Baker and Harman [B-H], there is a uniform bound
\[ \left\| \sum_{n \in X} \mu(n)e(n\theta) \right\|_\infty \ll X^{\frac{3}{4} + \epsilon}. \] (3.11)

Hence
\[ \left| \sum_{n < X} \mu(n)w_S(n) \right| < \|\hat{w}_S\|_1X^{\frac{3}{4} + \epsilon} < (\log X)|S|X^{\frac{3}{4} + \epsilon} \] (3.12)

and we may assume
\[ |S| > c \frac{\log X}{\log \log X}. \] (3.13)

If (3.13), apply the type-I-II analysis above.

From (2.31), assuming
\[ M \sim 2^\mu > X^{c_1 \frac{1}{\log \log X}} \] (3.14)

and choosing \( L \) appropriately, we obtain
\[ (2.1) < X.2^{-c\frac{\log X}{(\log \log X)^2}}. \] (3.15)

If \( M \) fails (3.14) the type-I bound (3.2’) gives
\[ (3.1) < X.M\|\hat{w}_S\|_\infty \]
\[ < \frac{1}{c_1 \log \log X} X.X^{c_1 \frac{1}{\log \log X}} 2^{-c\frac{\log X}{\log \log X}} \]
\[ < X^{1-c_2 \frac{1}{\log \log X}} \] (3.16)

for appropriate choice of \( c_1 \) in (3.14).
In either case
\[ \left| \sum_{n < X} \mu(n)w_S(n) \right| < X^{1 - \frac{c}{(\log \log X)^2}} \quad (3.17) \]

which is Theorem 2.

REFERENCES

[B-H] R. Baker, G. Harman, Exponential sums formed with the Möbius function, *JLMS.* (2), 43 (1991), no 2, 193–198.

[Gr] B. Green, On (not) computing the Möbius function using bounded depth circuits, preprint, March 2011.

[H-K] G. Harman, I. Katai, Primes with preassigned digits II, Acta Arith. 133 (2008), no 2, 171–184.

[I-K] H. Iwaniec, E. Kowalski, Analytic Number Theory, AMS, 2004.

[Ka] G. Kalai, Private communication.

[M-R] C. Mauduit, J. Rivat, Sur un problème de Gelfond: la somme des chiffres des nombres premiers, Annals Math (2), 171 (2010), no 3, 1591–1646.

INSTITUTE FOR ADVANCED STUDY, 1 EINSTEIN DRIVE, PRINCETON, NJ 08540