AUTOCORRELATION OF THE MOBIUS FUNCTION

N. A. CARELLA

Abstract. Let \( x \geq 1 \) be a large integer, and let \( \mu : \mathbb{N} \rightarrow \{-1,0,1\} \) be the Mobius function. This article proposes an effective asymptotic result for the autocorrelation function
\[
\sum_{n \leq x} \mu(n)\mu(n+t) = O\left(e^{-c\sqrt{\log x}}\right),
\]
where \( t \neq 0 \) be a small fixed integer, and \( c > 0 \) is a constant.

Contents

1. Introduction 1
2. Average Orders of Mobius Functions 2
3. Integers in Arithmetic Progressions 2
4. Proof of Theorem 1.1 4
References 7

1. Introduction

The correlation function induced by a pair of arithmetic functions \( f, g : \mathbb{N} \rightarrow \mathbb{C} \) is defined by the arithmetic average
\[
\sum_{n \leq x} f(n)g(n).
\]

The Chowla conjecture, Elliot conjecture, and Sarnak conjecture are correlation problems associated with the Mobius function \( g(n) = \mu(n) \) and bounded multiplicative functions \( f(n) \ll 1 \). These problems, which are the topics of current research, have many similarities and common structures. The corresponding arithmetic averages are, respectively,
\[
\sum_{n \leq x} \mu(n)\mu(n+t) = o(x),
\]
where \( t \neq 0 \), see [2], [12];
\[
\sum_{n \leq x} f(n)f(n+t) = o(x),
\]
where \( f : \mathbb{N} \rightarrow \mathbb{C} \) is a bounded arithmetic function, [5], [9], and
\[
\sum_{n \leq x} \mu(n)f(T^n(z)) = o(x),
\]
where \( T : \mathbb{Z} \rightarrow \mathbb{Z} \) is a Borel measurable transformation, see [13]. Each of these conjectures has an extensive literature.

2010 Mathematics Subject Classification. Primary 11N37; Secondary 11P32.
Key words and phrases. Autocorrelation, Correlation, Arithmetic function, Mobius function, Liouville function, Chowla conjecture.
The best estimates for (1.2) in the literature have the asymptotic formulae of the forms
\[ \sum_{n \leq x} \mu(n)\mu(n+t) = O\left(\frac{x}{\sqrt{\log \log x}}\right), \quad (1.5) \]
where \( t \neq 0 \) is a fixed integer, see [7, Corollary 5]. Here, the following result is proposed.

**Theorem 1.1.** Let \( x \geq 1 \) be a large prime number, and let \( \mu : \mathbb{N} \to \{-1, 0, 1\} \) be the Mobius function. If \( t \neq 0 \) is an integer, then,
\[ \sum_{n \leq x} \mu(n)\mu(n+t) = O\left(e^{-c\sqrt{\log x}}\right), \]
where \( c > 0 \) is a constant.

The proof of Theorem 1.1, based on a new idea and standard results in analytic number theory, is given in Section 4.

2. **Average Orders of Mobius Functions**

**Theorem 2.1.** If \( \mu : \mathbb{N} \to \{-1, 0, 1\} \) is the Mobius function, then, for any large number \( x > 1 \), the following statements are true.

(i) \[ \sum_{n \leq x} \mu(n) = O\left(xe^{-c\sqrt{\log x}}\right), \]
unconditionally,

(ii) \[ \frac{1}{n} \sum_{n \leq x} \mu(n) = O\left(e^{-c\sqrt{\log x}}\right), \]
unconditionally,

where \( c > 0 \) is an absolute constant.

**Proof.** See, [11, p. 182], [8, p. 347], et alii.

There are many sharp bounds of the summatory function of the Mobius function, say, \( O\left(xe^{-c(\log x)^\delta}\right) \), and the conditional estimate \( O\left(x^{1/2+\varepsilon}\right) \) presupposes that the nontrivial zeros of the zeta function \( \zeta(\rho) = 0 \) in the critical strip \( \{0 < \Re(s) < 1\} \) are of the form \( \rho = 1/2 + it, t \in \mathbb{R} \). However, the simpler notation will be used whenever it is convenient.

3. **Integers in Arithmetic Progressions**

An effective asymptotic formula for the number of integers in arithmetic progressions is derived in Lemma 3.1. The derivation is based on a version of the basic large sieve inequality stated below.

**Theorem 3.1.** Let \( x \) be a large number and let \( Q \leq x \). If \( \{a_n : n \geq 1\} \) is a sequence of real number, then
\[ \sum_{q \leq Q} q \sum_{1 \leq a \leq q} \left| \sum_{n \leq x} a_n - \frac{1}{q} \sum_{n \leq x} a_n \right|^2 \leq Q \left(10Q + 2\pi x\right) \sum_{n \leq x} |a_n|^2. \]

**Proof.** The essential technical details are covered in [3, Chapter 23]. This inequality is discussed in [6] and the literature in the theory of the large sieve.
**Lemma 3.1.** If $x \geq 1$ is a large number and $1 \leq a < q \leq x$, then

$$\max_{1 \leq a \leq q} \left| \sum_{n \leq x \atop n \equiv a \mod q} 1 - \frac{1}{q} \sum_{n \leq x} 1 \right| = O \left( \frac{x}{q} e^{-c\sqrt{\log x}} \right), \quad (3.1)$$

where $c > 0$ is a constant. In particular,

$$\sum_{n \leq x \atop n \equiv a \mod q} 1 = \left[ \frac{x}{q} \right] + O \left( \frac{x}{q} e^{-c\sqrt{\log x}} \right), \quad (3.2)$$

**Proof.** Trivially, the basic finite sum satisfies the asymptotic

$$\sum_{n \leq x} 1 = [x] = x - \{x\}, \quad (3.3)$$

where $[x] = x - \{x\}$ is the largest integer function, and the number of integers in any equivalent class satisfies the asymptotic formula

$$\sum_{n \leq x \atop n \equiv a \mod q} 1 = \frac{x}{q} + E(x). \quad (3.4)$$

Let $Q = x$ and let the sequence of real numbers be $a_n = 1$ for $n \geq 1$. Now suppose that the error term is of the form

$$E(x) = E_0(x) = O \left( x^\alpha \right), \quad (3.5)$$

where $\alpha \in (0, 1]$ is a constant. Then, the large sieve inequality, Theorem 3.1, yields the lower bound

$$\sum_{q \leq x} q \sum_{1 \leq a \leq q} \left| \sum_{n \leq x \atop n \equiv a \mod q} 1 - \frac{1}{q} \sum_{n \leq x} 1 \right|^2 = \sum_{q \leq x} q \sum_{1 \leq a \leq q} \left| \frac{x}{q} + O \left( x^\alpha \right) - \frac{x - \{x\}}{q} \right|^2$$

$$\gg \sum_{q \leq x} q \sum_{1 \leq a \leq q} \left| x^\alpha + \frac{\{x\}}{q} \right|^2$$

$$\gg \sum_{q \leq x} q \sum_{1 \leq a \leq q} |x^\alpha|^2$$

$$\gg x^{2\alpha} \sum_{q \leq x} q \sum_{1 \leq a \leq q} 1$$

$$\gg x^{2\alpha} \sum_{q \leq x} q^2$$

$$\gg x^{3+2\alpha}. \quad (3.6)$$
On the other direction, it yields the upper bound

\[ \sum_{q \leq x} q \sum_{1 \leq a \leq q} \left| \sum_{n \leq x, \ n \equiv a \ mod \ q} 1 - \frac{1}{q} \sum_{n \leq x} 1 \right|^2 \leq Q (10Q + 2\pi x) \sum_{n \leq x} |a_n|^2 \]

\[ \leq x (10x + 2\pi x) \sum_{n \leq x} |1|^2 \]

\[ \ll x^3. \]

Clearly, the lower bound in (3.6) contradicts the upper bound in (3.7). Similarly, the other possibilities for the error term

\[ E_1 = O \left( \frac{x}{(\log x)^c} \right) \quad \text{and} \quad E_2 = O \left( x e^{-c\sqrt{\log x}} \right), \]

contradict large sieve inequality. Therefore, the error term is of the form

\[ E(x) = O \left( \frac{x}{q \log x} \right) = O \left( \frac{x}{(\log x)^c} \right) = O \left( \frac{x}{q} \right), \]

where \( c > 0 \) is a constant.

\[ \blacksquare \]

4. PROOF OF THEOREM 1.1

The proof explored in this section is based on a new technique from harmonic analysis. This technique uses the Ramanujan sum \( c_q(n) \) and elementary analytic methods to derive an asymptotic formula for the Mobius autocorrelation function

\[ R(t) = \sum_{n \leq x} \mu(n)\mu(n + t). \] (4.1)

**Proof.** (Theorem 1.1) Without loss in generality, assume that \( x \geq 1 \) is a large integer, and let \( t = 1 \). Replace this identity, see [1, Section 8.3],

\[ c_n(1) = \sum_{1 \leq u < n, \ \gcd(n,u) = 1} e^{i2\pi u/n} = \mu(n) \] (4.2)

in (4.1) twice, and substitute the characteristic function

\[ \sum_{d \mid a \mid n} \mu(d) = \begin{cases} 1 & \text{if } \gcd(a, n) = 1, \\ 0 & \text{if } \gcd(a, n) \neq 1, \end{cases} \] (4.3)

of relatively prime numbers. These substitutions transform (4.1) into an exponential autocorrelation function:

\[ \sum_{n \leq x} \mu(n)\mu(n + 1) = \sum_{n \leq x} \sum_{1 \leq u < n, \ \gcd(n,u) = 1} e^{i2\pi u/n} \times \sum_{1 \leq v < n + 1, \ \gcd(n+1,v) = 1} e^{i2\pi v/n+1} \]

\[ = \sum_{n \leq x} \sum_{1 \leq u < n} e^{i2\pi u/n} \sum_{d_1 \mid n} \mu(d_1) \times \sum_{d_2 \mid n+1} e^{i2\pi v/n+1} \sum_{d_2 \mid n} \mu(d_2). \] (4.4)
Next, switch the order of summations

\[ R(1) = \sum_{n \leq x} \mu(n)\mu(n+1) \] (4.5)

\[ = \sum_{n \leq x} \sum_{n \mid d_1} \mu(d_1) \sum_{1 \leq u < n \atop d_1 \mid u} e^{i2\pi u/n} \times \sum_{d_2 \mid n+1} \mu(d_2) \sum_{1 \leq v < n+1 \atop d_2 \mid v} e^{i2\pi v/(n+1)} \]

\[ = \sum_{1 \leq d_1 < x} \sum_{1 \leq d_2 < x+1} \mu(d_1)\mu(d_2) \sum_{1 \leq u < n \atop d_1 \mid u} e^{i2\pi u/n} \sum_{1 \leq v < n+1 \atop d_2 \mid v} e^{i2\pi v/(n+1)}. \]

Proceed to substitute the change of variables

\[ u = d_1 r, \quad n = d_1 k; \quad (4.6) \]
\[ v = d_2 s, \quad n + 1 = d_2 m; \quad (4.7) \]

to simplify (4.5). Specifically,

\[ \sum_{n \leq x} \mu(n)\mu(n+1) = \sum_{1 \leq d_1 < x} \sum_{1 \leq d_2 < x+1} \mu(d_1)\mu(d_2) \sum_{1 \leq u < n \atop d_1 \mid u} e^{i2\pi r/k} \sum_{1 \leq s < m} e^{i2\pi s/m} \]

\[ = \sum_{1 \leq d_1 < x} \sum_{1 \leq d_2 < x+1} \mu(d_1)\mu(d_2) 1. \quad (4.8) \]

The last equality follows from

\[ \sum_{1 \leq r < k} e^{i2\pi r/k} = \sum_{1 \leq s < m} e^{i2\pi s/m} = -1 \quad (4.9) \]

for any integers \( k, m \geq 2 \).
Now, the conditions \( d_1 \mid n \) and \( d_2 \mid n + 1 \) imply that \( \text{lcm}(d_1, d_2) = d_1d_2 \). Rearrange the last finite sum in the equivalent form

\[
R(1) = \sum_{n \leq x} \mu(n)\mu(n + 1)
\]

\[
= \sum_{1 \leq d_1 < x} \mu(d_1)\mu(d_2) \sum_{1 \leq n \leq x \atop d_1 \mid n, d_2 \mid n + 1} 1
\]

\[
= \sum_{1 \leq d_1 < x} \mu(d_1)\mu(d_2) \left( \sum_{1 \leq n \leq x \atop d_1 \mid n, d_2 \mid n + 1} 1 - \frac{x}{d_1d_2} + \frac{x}{d_1d_2} \right)
\]

\[
= x \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x + 1, \gcd(d_1, d_2) = 1} \frac{\mu(d_1)d_2}{d_1d_2} + \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x + 1, \gcd(d_1, d_2) = 1} \mu(d_1)\mu(d_2) \left( \sum_{1 \leq n \leq x \atop d_1 \mid n, d_2 \mid n + 1} 1 - \frac{x}{d_1d_2} \right)
\]

\[
= R_0(x) + R_1(x).
\]

The first finite sum has the upper bound

\[
R_0(x) = x \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x + 1, \gcd(d_1, d_2) = 1} \frac{\mu(d_1)d_2}{d_1d_2}
\]

\[
= x \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x + 1} \frac{\mu(d_1)d_2}{d_1d_2}
\]

\[
= O \left( xe^{-c_1\sqrt{\log x}} \right),
\]

where \( d_1 < x \) and \( d_2 < x + 1 \) are relatively prime and independent variables, this follows from Theorem 2.1. The upper bound of the second sum is computed in Lemma 4.1. Summing these estimates yields

\[
\sum_{n \leq x} \mu(n)\mu(n + 1) = R_0(x) + R_1(x)
\]

\[
= O \left( xe^{-c_1\sqrt{\log x}} \right)
\]

as claimed.

**Lemma 4.1.** If \( x \) is a large number, then

\[
R_1(x) = \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x + 1 \atop \gcd(d_1, d_2) = 1} \mu(d_1)\mu(d_2) \left( \sum_{1 \leq n \leq x \atop d_1 \mid n, d_2 \mid n + 1} 1 - \frac{x}{d_1d_2} \right) = O \left( xe^{-c_2\sqrt{\log x}} \right),
\]
where $c_2 > 0$ is a constant.

**Proof.** Let $q = d_1 d_2 \leq x$. Taking absolute value and plugging in the asymptotic number of integers in the arithmetic progression, see Lemma 3.1, into the inner sum yield the second finite sum has the upper bound,

\[
|R_1(x)| \leq \left| \sum_{\substack{1 \leq d_1 < x \\ 1 \leq d_2 < x+1 \atop \gcd(d_1, d_2) = 1}} \mu(d_1) \mu(d_2) \left( \sum_{1 \leq n \leq x \atop d_1 | n, d_2 | n+1} 1 - \frac{x}{d_1 d_2} \right) \right| \tag{4.13}
\]

\[
\leq \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x+1} \left| \sum_{1 \leq n \leq x \atop d_1 | n, d_2 | n+1} 1 - \frac{x}{d_1 d_2} \right|
\]

\[
\ll \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x+1} \frac{x}{d_1 d_2} e^{-c \sqrt{\log x}}
\]

\[
\ll x e^{-c \sqrt{\log x}} \sum_{1 \leq d_1 < x \atop 1 \leq d_2 < x+1} \frac{1}{d_1 d_2}
\]

\[
= O \left( x e^{-c_2 \sqrt{\log x}} \right),
\]

where $c, c_2 > 0$ are constants, and the asymptotic estimate

\[
\sum_{1 \leq n \leq x} \frac{1}{n} = O \left( e^{\log \log x} \right),
\]

is used on the last line of expression (4.13) to simplify the upper bound. ■

**References**

[1] Apostol, Tom M. *Introduction to analytic number theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1976.

[2] Chowla, S. *The Riemann hypothesis and Hilbert’s tenth problem*. Mathematics and Its Applications, Vol. 4. Gordon and Breach Science Publishers, New York-London-Paris, 1965.

[3] Davenport, H. *Multiplicative number theory*. volume 74 of Graduate Texts in Mathematics. Springer-Verlag, New York, third edition, 2000.

[4] De Koninck, Jean-Marie; Luca, Florian. *Analytic number theory. Exploring the anatomy of integers*. Graduate Studies in Mathematics, 134. American Mathematical Society, Providence, RI, 2012.

[5] Elliott, P. D. T. A. *On the correlation of multiplicative and the sum of additive arithmetic functions*. Mem. Amer. Math. Soc., 112(538), 1994.
[6] Gallagher, P. X. *The large sieve*. Mathematika 14 (1967), 14-20. 

[7] Helfgott, Harald Andres. *Expansion, divisibility and parity: an explanation*. http://arxiv.org/abs/2201.00799. 

[8] Hardy, G. H.; Wright, E. M. *An introduction to the theory of numbers*. Sixth edition. Oxford University Press, Oxford, 2008. 

[9] Klurman, O.; Mangerel, A. P.; Teravainen, J. *On Elliott’s conjecture and applications*. http://arxiv.org/abs/2304.05344. 

[10] Matomaki, Kaisa, Radziwill, M., Tao, T. *An averaged form of Chowla’s conjecture*. http://arxiv.org/abs/1503.05121. 

[11] Montgomery, Hugh L.; Vaughan, Robert C. *Multiplicative number theory. I. Classical theory*. Cambridge University Press, Cambridge, 2007. 

[12] Ramare, Olivier. *Chowla’s Conjecture: From the Liouville Function to the Moebius Function*. Part of the Lecture Notes in Mathematics book series (LNM, volume 2213), 16 June 2018. 

[13] Sarnak, P. *Three Lectures on the Mobius Function Randomness and Dynamics*. Preprint 2010. https://www.math.ias.edu/files/wam/2011/PSMobius.pdf.