

THE AREA METHOD AND APPLICATIONS

T. AGAMA

ABSTRACT. In this paper, we develop a general method for estimating correlations of the forms

$$\sum_{n \leq x} G(n)G(x-n)$$

and

$$\sum_{n \leq x} G(n)G(n+l)$$

for a fixed $1 \leq l \leq x$ and where $G : \mathbb{N} \rightarrow \mathbb{R}^+$. To distinguish between the two types of correlation, we call the first correlation the **type 2** correlation and the second the **type 1** correlation. As an application, we estimate the lower bound for the **type 2** correlation of the master function

$$\sum_{n \leq x} \Upsilon(n)\Upsilon(n+l_0) \geq (1+o(1))\frac{x}{2C(l_0)} \log \log^2 x$$

provided that $\Upsilon(n)\Upsilon(n+l_0) > 0$. We also use this method to provide a first proof of the twin prime conjecture showing that

$$\sum_{n \leq x} \Lambda(n)\Lambda(n+2) \geq (1+o(1))\frac{x}{2C(2)}$$

for some $C := C(2) > 0$.

1. INTRODUCTION

The correlations of arithmetic functions lie at the heart of several central problems in additive and multiplicative number theory. Sums of the form

$$\sum_{n \leq x} G(n)G(n+l) \quad \text{and} \quad \sum_{n \leq x} G(n)G(x-n),$$

where $G : \mathbb{N} \rightarrow \mathbb{R}^+$, encode delicate information about the distribution of primes, almost primes, divisor-type functions, and other multiplicative sequences. In many settings, the main difficulty is not the formal definition of such sums but rather the extraction of a usable lower bound or asymptotic estimate. Classical tools such as the circle method, sieve methods, and summation formulas remain indispensable in this direction; see, for example, [2, 3, 4]. The purpose of this paper is to introduce a geometric device, which we call the *area method*, that converts certain one-dimensional correlation sums into structured double sums that are more amenable to estimation.

Date: June 27, 2026.

2010 Mathematics Subject Classification. Primary 11N37; Secondary 11P32, 11N25.

Key words and phrases. master function; correlation; two-point.

The guiding idea is simple but effective. We distinguish between two basic correlation types. The first, which we call the **type 1** correlation, is the shifted correlation

$$\sum_{n \leq x} G(n)G(n+l),$$

for a fixed shift l . The second, which we call the **type 2** correlation, is the complementary correlation

$$\sum_{n \leq x} G(n)G(x-n).$$

The area method begins with a decomposition of a geometric region into triangles, trapezia, rectangles, and squares. This geometric decomposition produces an algebraic identity that reorganizes bilinear expressions into double sums over triangular domains. Once this rearrangement is available, standard analytic tools, especially partial summation and mean-value estimates for arithmetic functions, can be applied in a unified way. In this sense, the method provides a bridge between a geometric picture and the analytic structure of correlation sums.

A key feature of the method is that it produces a general lower-bound principle for positive arithmetic functions. Roughly speaking, if the correlation $\sum_{n \leq x} f(n)f(n+l_0)$ is nontrivial, then it can be bounded from below by a multiple of a cumulative double sum built from f . The resulting constants depend on the shift l_0 and on the arithmetic structure of the function under consideration. This dependence is important: it allows the same formal argument to treat a variety of examples, while still reflecting the different local behaviors of the underlying sequence. In particular, functions with sparse support, such as the von Mangoldt function Λ , may behave differently from denser sequences such as the divisor function $d(n)$, the totient function $\phi(n)$, or the squarefree indicator $\mu^2(n)$. The paper shows how the same geometric mechanism yields explicit lower bounds in each of these cases, with the relevant average-order estimates taken from [2, 3, 4]. The master function Υ , introduced in [1], is also treated as a further illustration of the flexibility of the method.

The paper has several goals. First, it develops the fundamental identity underlying the area method and shows how the identity can be used to control the correlations of type 1 and type 2. Second, it applies this framework to obtain lower bounds for a number of standard arithmetic functions, including Λ , d , d_l , ϕ , and μ^2 . Third, it applies the same method to the master function Υ , which yields estimates for its shifted self-correlation and thus connects the geometric method to the study of integers with exactly two prime factors. Finally, the paper applies the type 2 framework to complementary correlations of prime-related and divisor-type sequences, including a Goldbach-type representation problem.

1.1. Organization of the paper. The organization of the paper is as follows. In Section 2, we develop the area method and prove the basic identity that drives the entire argument. Section 3 applies the method to the shifted correlation of the von Mangoldt function and to the associated twin-prime-type pair correlation. Section 4 uses the same principle to derive lower bounds for other correlations of type 1, including those involving $d(n)$, $d_l(n)$, $\phi(n)$, and $\mu^2(n)$. Section 5 is devoted to the master function Υ and its two-point correlations. Section 6 studies the complementary correlation associated with representations of even

numbers as sums of two primes. Section 7 extends the type 2 method to divisor-type and totient-type sequences. The final section discusses the global distributional consequences for integers with $\Omega(n) = 2$.

2. THE AREA METHOD

In this section, we introduce and develop a fundamental method for solving problems related to correlations of arithmetic functions. This method is fundamental in the sense that it uses the properties of four main geometric shapes, namely the triangle, trapezium, rectangle, and square. The basic identity that we will derive is an outgrowth of exploiting the areas of these shapes and putting them together in a unified manner.

Theorem 2.1. *Let $\{r_j\}_{j=1}^n$ and $\{h_j\}_{j=1}^n$ be any sequence of real numbers, and let r and h be any real numbers satisfying*

$$\sum_{j=1}^n r_j = r \quad \text{and} \quad \sum_{j=1}^n h_j = h$$

and

$$(r^2 + h^2)^{1/2} = \sum_{j=1}^n (r_j^2 + h_j^2)^{1/2}.$$

We have

$$\sum_{j=2}^n r_j h_j = \sum_{j=2}^n h_j \left(\sum_{i=1}^j r_i + \sum_{i=1}^{j-1} r_i \right) - 2 \sum_{j=1}^{n-1} r_j \sum_{k=1}^{n-j} h_{j+k}.$$

Proof. Consider a right-angled triangle $\triangle ABC$ in a plane with height h and base r . Next, let us partition the height of the triangle into n parts, not necessarily equal. Now, we link those partitions along the height to the hypotenuse with the aid of a parallel line. At the point of contact of each line with the hypotenuse, we drop a vertical line to the next line connecting the last point of the previous partition, thus forming another right-angled triangle $\triangle A_1 B_1 C_1$ with base and height r_1 and h_1 respectively. We note that this triangle is covered by the triangle $\triangle ABC$ with hypotenuse constituting a proportion of the hypotenuse of triangle $\triangle ABC$. We continue this process until we obtain n right-angled triangles $\triangle A_j B_j C_j$, each with base and height r_j and h_j for $j = 1, 2, \dots, n$. This construction satisfies

$$h = \sum_{j=1}^n h_j \quad \text{and} \quad r = \sum_{j=1}^n r_j$$

and

$$(r^2 + h^2)^{1/2} = \sum_{j=1}^n (r_j^2 + h_j^2)^{1/2}.$$

Now, we deform the original triangle $\triangle ABC$ by removing the smaller triangles $\triangle A_j B_j C_j$ for $j = 1, 2, \dots, n$. Essentially, we are left with rectangles and squares piled on top of each other, each end poking out a bit further than the one just above. We observe that the total

area of this portrait is

$$\begin{aligned}\mathcal{A}_1 &= r_1 h_2 + (r_1 + r_2) h_3 + \cdots + (r_1 + r_2 + \cdots + r_{n-2}) h_{n-1} + (r_1 + r_2 + \cdots + r_{n-1}) h_n \\ &= r_1 (h_2 + h_3 + \cdots + h_n) + r_2 (h_3 + h_4 + \cdots + h_n) + \cdots + r_{n-2} (h_{n-1} + h_n) + r_{n-1} h_n \\ &= \sum_{j=1}^{n-1} r_j \sum_{k=1}^{n-j} h_{j+k}.\end{aligned}$$

On the other hand, we observe that the area of this portrait is the same as the difference between the area of triangle $\triangle ABC$ and the sum of the areas of triangles $\triangle A_j B_j C_j$ for $j = 1, 2, \dots, n$. We deduce

$$\mathcal{A}_1 = \frac{1}{2} r h - \frac{1}{2} \sum_{j=1}^n r_j h_j.$$

This completes the first part of the argument. For the second part, along the hypotenuse we construct small pieces of triangle, each of base and height (r_i, h_i) ($i = 1, 2, \dots, n$) so that the trapezoid and the one triangle formed by partitioning become rectangles and squares. We observe that this construction satisfies the relation

$$(r^2 + h^2)^{1/2} = \sum_{i=1}^n (r_i^2 + h_i^2)^{1/2}.$$

Now, we compute the area of the triangle in two different ways. By direct strategy, we find that the area of the triangle, denoted by \mathcal{A} , is

$$\mathcal{A} = 1/2 \left(\sum_{i=1}^n r_i \right) \left(\sum_{i=1}^n h_i \right).$$

On the other hand, we compute the area of the triangle by computing the area of each trapezium and the one remaining triangle and sum them together. That is,

$$\mathcal{A} = h_n/2 \left(\sum_{i=1}^n r_i + \sum_{i=1}^{n-1} r_i \right) + h_{n-1}/2 \left(\sum_{i=1}^{n-1} r_i + \sum_{i=1}^{n-2} r_i \right) + \cdots + 1/2 r_1 h_1.$$

By comparing the deduced area of the triangle in the second argument and comparing this to the deduced area of the triangle in the first argument, the result follows immediately. \square

We state a result for a general lower bound for any two-point correlation that captures all real arithmetic functions.

Theorem 2.2. *Let $f : \mathbb{N} \rightarrow \mathbb{R}^+$. If*

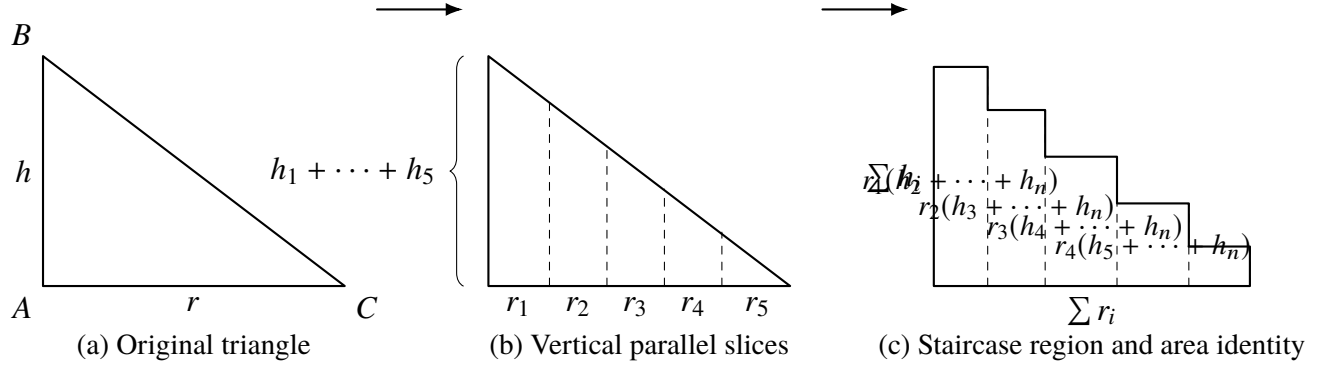
$$\sum_{n \leq x} f(n) f(n + l_0) > 0$$

then there exists some $C := C(l_0) > 0$ such that

$$\sum_{n \leq x} f(n) f(n + l_0) \geq \frac{1}{C(l_0)x} \sum_{2 \leq n \leq x} f(n) \sum_{m \leq n-1} f(m).$$

Proof. By Theorem 2.1, we obtain the identity by taking $f(j) = r_j = h_j$

$$\sum_{n \leq x-1} \sum_{j \leq x-n} f(n) f(n + j) = \sum_{2 \leq n \leq x} f(n) \sum_{m \leq n-1} f(m).$$



Key identity:
$$\sum_{j=1}^{n-1} r_j \sum_{k=1}^{n-j} h_{j+k} = \frac{1}{2} r h - \frac{1}{2} \sum_{j=1}^n r_j h_j.$$

FIGURE 1. Geometric mechanism behind the area method. The triangle is first partitioned by vertical parallel slices, and the remaining staircase region produces the double-sum identity.

It follows that

$$\begin{aligned}
\sum_{n \leq x-1} \sum_{j \leq x-n} f(n) f(n+j) &\leq \sum_{n \leq x-1} \sum_{j < x} f(n) f(n+j) \\
&= \sum_{n \leq x} f(n) f(n+1) + \sum_{n \leq x} f(n) f(n+2) \\
&\quad + \cdots + \sum_{n \leq x} f(n) f(n+l_0) + \cdots + \sum_{n \leq x} f(n) f(n+x) \\
&\leq |\mathcal{M}(l_0)| \sum_{n \leq x} f(n) f(n+l_0) \\
&\quad + |\mathcal{N}(l_0)| \sum_{n \leq x} f(n) f(n+l_0) \\
&\quad + \cdots + \sum_{n \leq x} f(n) f(n+l_0) + \cdots + |\mathcal{R}(l_0)| \sum_{n \leq x} f(n) f(n+l_0) \\
&= \left(|\mathcal{M}(l_0)| + |\mathcal{N}(l_0)| + \cdots + 1 \right. \\
&\quad \left. + \cdots + |\mathcal{R}(l_0)| \right) \sum_{n \leq x} f(n) f(n+l_0) \\
&\leq C(l_0) x \sum_{n < x} f(n) f(n+l_0)
\end{aligned}$$

where

$$\max\{|\mathcal{M}(l_0)|, |\mathcal{N}(l_0)|, \dots, |\mathcal{R}(l_0)|\} = C(l_0).$$

Inverting this inequality, the result follows immediately. \square

The nature of the implicit constant $C(l_0)$ could also depend on the structure of the arithmetic function. The Von Mangoldt function $\Lambda(\cdot)$, contrary to many classes of arithmetic functions, has a relatively small constant. This behaviour stems from the fact that the

Von-Mangoldt function is defined on the prime powers. Thus, one would expect most terms in sums of forms

$$\sum_{n \leq x-1} \sum_{\substack{j \leq x-n \\ j \neq x-2n}} \Lambda(n)\Lambda(n+j)$$

to fall off when j is odd for any prime power $n = p^k$ such that $j + p^k \neq 2^s$.

Theorem 2.3. *Let $f : \mathbb{N} \rightarrow \mathbb{R}^+$. Suppose that there exists some $C := C(x)$ with $x > C(x) > 0$ such that*

$$\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ j \neq x-2n}} f(n)f(n+j) = \frac{C(x)}{x} \sum_{n \leq x} \sum_{j \leq x-n} f(n)f(n+j).$$

For all $x \geq 2$, we have

$$\sum_{n \leq \frac{x}{2}} f(n)f(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} f(n) \sum_{m \leq n-1} f(m),$$

where $x - \mathcal{D}(x) = C(x)$.

Proof. By Theorem 2.1, we obtain

$$\sum_{n \leq x} f^2(n) = f^2(1) + \sum_{2 \leq n \leq x} f(n) \left(\sum_{m \leq n-1} f(m) + \sum_{m \leq n} f(m) \right) - 2 \sum_{n \leq x-1} f(n) \sum_{s \leq x-n} f(n+s)$$

for $f : \mathbb{N} \rightarrow \mathbb{R}^+$ by taking $r_j = h_j = f(j)$. Rearranging the terms in this identity, we obtain

$$\sum_{n \leq x-1} \sum_{j \leq x-n} f(n)f(n+j) = \sum_{2 \leq n \leq x} f(n) \sum_{m \leq n-1} f(m).$$

Let $n + j = x - n$, then we deduce $x - 2n = j$. It follows that $j \leq x - 2$ if and only if $1 \leq n < \frac{x}{2}$. We can rewrite for the sum on the left-hand side

$$\begin{aligned} \sum_{n \leq x-1} \sum_{j \leq x-n} f(n)f(n+j) &= \sum_{n \leq x-1} \sum_{x-2n=j} f(n)f(n+j) + \sum_{\substack{n \leq x-1 \\ x-2n \neq j}} \sum_{j \leq x-n} f(n)f(n+j) \\ &= \sum_{n < \frac{x}{2}} f(n)f(x-n) + \frac{C(x)}{x} \sum_{n \leq x-1} \sum_{j \leq x-n} f(n)f(n+j) \end{aligned}$$

where $0 < \frac{C(x)}{x} < 1$. We deduce

$$\frac{\mathcal{D}(x)}{x} \sum_{n \leq x-1} \sum_{j \leq x-n} f(n)f(n+j) = \sum_{n < \frac{x}{2}} f(n)f(x-n)$$

where $0 < \frac{\mathcal{D}(x)}{x} = 1 - \frac{C(x)}{x} < 1$. Using Theorem 2.1, we obtain

$$\sum_{n < \frac{x}{2}} f(n)f(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} f(n) \sum_{m \leq n-1} f(m)$$

and the result follows immediately. \square

3. APPLICATION TO THE TWIN PRIME CONJECTURE

Theorem 3.1. *There exists an absolute constant $C > 0$ such that*

$$\sum_{n \leq x} \Lambda(n)\Lambda(n+2) \geq (1 + o(1)) \frac{x}{2C}.$$

Proof. Using Theorem 2.2, we get

$$\sum_{n \leq x} \Lambda(n)\Lambda(n+2) \geq \frac{1}{C(2)x} \sum_{2 \leq n \leq x} \Lambda(n) \sum_{m \leq n-1} \Lambda(m).$$

Using the prime number theorem [2] of the form

$$\sum_{n \leq x} \Lambda(n) = (1 + o(1))x$$

establishes the lower bound. □

The lower bound in Theorem 3.1 solves the twin prime conjecture. The proposed method not only solves the twin prime conjecture, but is good in terms of its generality. It can be used to obtain lower bounds for a general class of correlated sums of the form

$$\sum_{n \leq x} f(n)f(n+k)$$

for a uniform $1 \leq k \leq x$.

4. APPLICATION TO OTHER CORRELATED SUMS OF TYPE 1

In this section, we apply Theorem 2.2 to provide lower estimates of other correlated sums, but with the price of an implicit constant depending on the range of shift.

Corollary 4.1. *For a fixed $l_0 > 0$, there exist some $C := C(l_0) > 0$ such that*

$$\sum_{n \leq x} d(n)d(n+l_0) \geq (1 + o(1)) \frac{x \log^2 x}{2C(l_0)}.$$

Proof. Using Theorem 2.2 and the crude estimate

$$\sum_{n \leq x} d(n) = (1 + o(1))x \log x$$

in [3] gives the lower bound. □

Corollary 4.2. *For a fixed $k_0 > 0$ and for $l \geq 2$, there exists a constant $C := C(k) > 0$ such that*

$$\sum_{n \leq x} d_l(n)d_l(n+k) \geq (1 + o(1)) \left(\frac{1}{(l-1)!} \right) \left(1 - \frac{1}{2(l-1)!} \right) \frac{x \log^{2(l-1)} x}{C(k)}.$$

Proof. We recall the weaker estimate for the l th divisor function

$$\sum_{n \leq x} d_l(n) = (1 + o(1)) \frac{1}{(l-1)!} x \log^{l-1} x$$

in [3] where

$$d_l(n) = \sum_{n_1 \cdot n_2 \cdots n_l = n} 1.$$

Using Theorem 2.2 and using partial summation gives the lower bound. □

Corollary 4.3. For a fixed $l_0 > 0$, there exists a constant $C := C(l_0) > 0$ such that

$$\sum_{n \leq x} \phi(n)\phi(n + l_0) \geq (1 + o(1)) \frac{9}{2\pi^4} \frac{x^3}{C(l_0)}.$$

Proof. Using Theorem 2.2 and the estimate

$$\sum_{n \leq x} \phi(n) = (1 + o(1)) \frac{3}{\pi^2} x^2$$

in [4] gives the lower bound. □

Corollary 4.4. For a fixed $l_0 > 0$, there exists a constant $C := C(l_0) > 0$ such that

$$\sum_{n \leq x} \mu^2(n)\mu^2(n + l_0) \geq (1 + o(1)) \frac{18}{\pi^4} \frac{x}{C(l_0)}.$$

Proof. Using Theorem 2.2 and the estimate

$$\sum_{n \leq x} \mu^2(n) = (1 + o(1)) \frac{6}{\pi^2} x$$

in [4] gives the lower bound. □

5. APPLICATION TO LOWER BOUND FOR TWO-POINT CORRELATION OF THE MASTER FUNCTION OF TYPE 1 AND TYPE 2

In this section, we apply the area method developed to establish a lower bound for the two-point **type 1** correlation and an estimate for the **type 2** correlation of the master function. Here, p, q denotes the prime numbers and $\Omega(\cdot)$ counts the number of prime factors with multiplicity.

Definition 5.1.

$$\Upsilon(n) := \begin{cases} \log p & \text{if } n = p^2 \\ \log(pq) & \text{if } n = pq, \quad p \neq q \\ 0 & \Omega(n) > 2. \end{cases}$$

We call Υ the *master function*.

Lemma 5.2. Let Υ denote the master function. We have

$$\sum_{n \leq x} \Upsilon(n) = x \log \log x + O(x).$$

Proof. For a proof, see, e.g., [1]. □

Theorem 5.3. We have

$$\sum_{n \leq x} \Upsilon(n)\Upsilon(n + l_0) \geq (1 + o(1)) \frac{x}{2C(l_0)} \log \log^2 x$$

provided that $\Upsilon(n)\Upsilon(n + l_0) > 0$.

Proof. Applying Theorem 2.2 and using Lemma 5.2, we get

$$\begin{aligned}\sum_{n \leq x} \Upsilon(n) \Upsilon(n + l_0) &\geq \frac{1}{xC(l_0)} \sum_{2 \leq n \leq x} \Upsilon(n) \sum_{m \leq n-1} \Upsilon(m) \\ &= \frac{1}{xC(l_0)} (1 + o(1)) \sum_{2 \leq n \leq x} \Upsilon(n) n \log \log n.\end{aligned}$$

By partial summation, we deduce

$$\begin{aligned}\sum_{2 \leq n \leq x} \Upsilon(n) n \log \log n &= x \log \log x \sum_{n \leq x} \Upsilon(n) - \int_2^x (1 + o(1)) t (\log \log t) \left(\log \log t + \frac{1}{\log t} \right) dt \\ &= (1 + o(1)) x^2 \log \log^2 x - (1 + o(1)) \int_2^x t (\log \log t) \left(\log \log t + \frac{1}{\log t} \right) dt \\ &= (1 + o(1)) \frac{x^2}{2} \log \log^2 x.\end{aligned}$$

The claimed lower bound follows immediately. \square

Theorem 5.4. *Under the assumption*

$$\frac{\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ j \neq x-2n}} \Upsilon(n) \Upsilon(n+j)}{\sum_{n \leq x} \sum_{j \leq x-n} \Upsilon(n) \Upsilon(n+j)} < 1$$

we have

$$\sum_{n \leq \frac{x}{2}} \Upsilon(n) \Upsilon(x-n) = (1 + o(1)) \frac{x}{2} \mathcal{D}(x) \log \log^2 x$$

where $\mathcal{D} := \mathcal{D}(x) > 0$.

Proof. The result follows by applying the area method. \square

6. APPLICATION TO ESTIMATES FOR THE NUMBER OF REPRESENTATIONS OF AN EVEN NUMBER AS A SUM OF TWO PRIMES

In this section, we apply the area method developed 2.3 to obtain a weaker estimate for the number of representations of an even number as a sum of two primes under the assumption that the binary Goldbach conjecture is true.

Theorem 6.1. *Assume that the Goldbach conjecture is true. For any even $x \geq 6$, we have*

$$\sum_{n \leq \frac{x}{2}} \Lambda(n) \Lambda(x-n) = (1 + o(1)) \frac{x}{2} \mathcal{D}(x)$$

where $\mathcal{D} := \mathcal{D}(x) > 0$.

Proof. Under the assumption that the Goldbach conjecture is true, we deduce

$$\frac{\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ j \neq x-2n}} \Lambda(n)\Lambda(n+j)}{\sum_{n \leq x} \sum_{j \leq x-n} \Lambda(n)\Lambda(n+j)} < 1.$$

Applying the area method, there exists some $\mathcal{D} := \mathcal{D}(x) > 0$ with $\mathcal{D}(x) < x$ such that

$$\sum_{n \leq \frac{x}{2}} \Lambda(n)\Lambda(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} \Lambda(n) \sum_{m \leq n-1} \Lambda(m).$$

Using the prime number theorem [2] in the form

$$\sum_{n \leq x} \Lambda(n) = (1 + o(1))x$$

we obtain

$$\sum_{2 \leq n \leq x} \Lambda(n) \sum_{m \leq n-1} \Lambda(m) = (1 + o(1)) \frac{x^2}{2}.$$

The result follows immediately. \square

7. APPLICATION TO OTHER CORRELATED SUMS OF TYPE 2

In this section, we apply the area method 2.3 to obtain estimates for various correlated sums of **type 2**.

Theorem 7.1. *We have*

$$\sum_{n \leq \frac{x}{2}} d(n)d(x-n) = \mathcal{D}(1 + o(1)) \frac{x \log^2 x}{2}$$

where $\mathcal{D} := \mathcal{D}(x) > 0$.

Proof. Since the divisor function is non-vanishing on the integers, we deduce

$$\frac{\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ x-2n \neq j}} d(n)d(n+j)}{\sum_{n \leq x} \sum_{j \leq x-n} d(n)d(n+j)} < 1.$$

Using the area method (Theorem 2.3), there exist some $\mathcal{D} := \mathcal{D}(x)$ with $0 < \mathcal{D}(x) < x$ such that

$$\sum_{n \leq \frac{x}{2}} d(n)d(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} d(n) \sum_{m \leq n-1} d(m).$$

Using the weaker estimate

$$\sum_{n \leq x} d(n) = (1 + o(1))x \log x$$

in [3], we obtain

$$\sum_{2 \leq n \leq x} d(n) \sum_{m \leq n-1} d(m) = (1 + o(1)) \frac{x^2 \log^2 x}{2}.$$

The estimate is obtained using the area method. \square

Theorem 7.2. *We have*

$$\sum_{n \leq \frac{x}{2}} \phi(n)\phi(x-n) = (1 + o(1))\mathcal{D} \frac{9}{2\pi^4} x^3$$

where $\mathcal{D} := \mathcal{D}(x) > 0$.

Proof. Since $\phi(n)$ is non-vanishing on the integers, we deduce

$$\frac{\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ x-2n \neq j}} \phi(n)\phi(n+j)}{\sum_{n \leq x} \sum_{j \leq x-n} \phi(n)\phi(n+j)} < 1.$$

By the area method (Theorem 2.3), there exist some $\mathcal{D} := \mathcal{D}(x) > 0$ with $\mathcal{D}(x) < x$ such that

$$\sum_{n \leq \frac{x}{2}} \phi(n)\phi(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} \phi(n) \sum_{m \leq n-1} \phi(m).$$

Using the estimate

$$\sum_{n \leq x} \phi(n) = (1 + o(1)) \frac{3}{\pi^2} x^2$$

in [3], we obtain

$$\sum_{2 \leq n \leq x} \phi(n) \sum_{m \leq n-1} \phi(m) = (1 + o(1)) \frac{9}{2\pi^4} x^4.$$

The claimed estimate is deduced using the area method. \square

Theorem 7.3. *We have*

$$\sum_{n \leq \frac{x}{2}} d_l(n)d_l(x-n) = (1 + o(1))\mathcal{D} \left(\frac{1}{(l-1)!} \right) \left(1 - \frac{1}{2(l-1)!} \right) x \log^{2(l-1)} x$$

where $\mathcal{D} := \mathcal{D}(x) > 0$ and where

$$d_l(n) := \sum_{n_1 n_2 \cdots n_l = n} 1.$$

Proof. We observe

$$\frac{\sum_{n \leq x} \sum_{\substack{j \leq x-n \\ j \neq x-2n}} d_l(n)d_l(n+j)}{\sum_{n \leq x} \sum_{j \leq x-n} d_l(n)d_l(n+j)} < 1.$$

Using the area method (Theorem 2.3), there exists some $\mathcal{D} := \mathcal{D}(x) > 0$ with $\mathcal{D}(x) < x$ such that

$$\sum_{n \leq \frac{x}{2}} d_l(n)d_l(x-n) = \frac{\mathcal{D}(x)}{x} \sum_{2 \leq n \leq x} d_l(n) \sum_{m \leq n-1} d_l(m).$$

Using the estimate

$$\sum_{n \leq x} d_l(n) = (1 + o(1)) \frac{1}{(l-1)!} x \log^{l-1} x,$$

[3], we deduce

$$\sum_{2 \leq n \leq x} d_l(n) \sum_{m \leq n-1} d_l(m) = (1 + o(1)) \left(\frac{1}{(l-1)!} \right) \left(1 - \frac{1}{2(l-1)!} \right) x^2 \log^{2(l-1)} x.$$

The claimed estimate is deduced using the area method. \square

8. APPLICATION TO THE GLOBAL DISTRIBUTION OF INTEGERS WITH TWO PRIME FACTORS WITH MULTIPLICITY

The lower bounds of the correlations of the arithmetic functions determine their local and global distribution. Theorem 5.3 gives

$$\sum_{n \leq x} \Upsilon(n) \Upsilon(n + l_0) \geq (1 + o(1)) \frac{x}{2C(l_0)} \log \log^2 x,$$

provided that $\sum_{n \leq x} \Upsilon(n) \Upsilon(n + l_0) > 0$. Thus, for some shift in the range $[1, x]$ the correlation can be made arbitrarily large by taking the right hand side arbitrarily large. This implies that there are infinitely many pairs of the form $(n, n + l_0)$ such that n and $n + l_0$ each have exactly two prime factors.

1

REFERENCES

1. T. Agama *The master function and applications*, arXiv preprint arXiv:1707.05679, 2017.
2. H.L. Montgomery and R.C. Vaughan, *Multiplicative number theory I: Classical theory*. vol.97, Cambridge university press, 2006.
3. M.B. Nathanson, *Graduate Texts in Mathematics*, New York, NY: Springer New York, 2000.
4. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, vol. 163, American Mathematical Soc., 2015.

DEPARTMENT OF MATHEMATICS, AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCE, GHANA
E-mail address: theophilus@aims.edu.gh/emperordagama@yahoo.com

1