

ON BOMBIERI'S ASYMPTOTIC SIEVE

KEVIN FORD

ABSTRACT. If a sequence (a_n) of non-negative real numbers has “best possible” distribution in arithmetic progressions, Bombieri showed that one can deduce an asymptotic formula for the sum $\sum_{n \leq x} a_n \Lambda_k(n)$ for $k \geq 2$. By constructing appropriate sequences, we show that any weakening of the well-distribution property is not sufficient to deduce the same conclusion.

1. INTRODUCTION

Many of the most famous problems in number theory can be described in terms of estimating the number of primes in an integer sequence. More generally, given a sequence (a_n) of positive real numbers (e.g. the characteristic function of a set of natural numbers), one can ask for bounds on the sum

$$S_1(x) = \sum_{n \leq x} a_n \Lambda(n),$$

where Λ is the von Mangoldt function. Removing from the sequence those terms with n divisible by a prime $\leq z$ leaves behind only terms with n composed of at most $\lfloor \frac{\log x}{\log z} \rfloor$ prime factors. If $z > \sqrt{x}$, then only terms with n prime are left. Motivated by this simple fact, the modern *sieve* was created by V. Brun ([3], [4]) to attack such problems, in particular, the Twin Prime Conjecture and Goldbach's Conjecture. Estimating the number of “unsifted” elements is usually accomplished by means of a weighted form of inclusion-exclusion, its precision entirely determined by the regularity of the sequence on the arithmetic progressions $0 \pmod d$ for squarefree d (see the monographs [8] and [9] for more about sieve procedures). Writing

$$A_d(x) = \sum_{\substack{n \leq x \\ d|n}} a_n,$$

one postulates the existence of a multiplicative function g so that

$$A_d(x) = g(d)A(x) + r_d(x),$$

where $A(x)$ is an approximation to $A_1(x)$ and the “remainders” $r_d(x)$ are small in some average sense. A typical hypothesis is

$$R(\nu) : \quad \forall B > 0, \quad \sum_{d \leq x^\nu} |r_d(x)| \ll_{\nu, B} \frac{A(x)}{\log^B x}.$$

Received by the editors September 16, 2003 and, in revised form, December 1, 2003.

2000 *Mathematics Subject Classification*. Primary 11N35.

This research was supported by National Science Foundation grants DMS-0070618 and DMS-0301083.

One also needs mild growth conditions on $A(x)$ and regularity conditions on g . There is some flexibility in choosing these conditions (see e.g. [2], [5], [6], [8], [9]), and generally these are easy to verify in practice. We say that a sieve problem has *sifting density* or *dimension* κ if $g(p)$ is about κ/p on average over primes p . In the important special case $\kappa = 1$, one expects for many problems that

$$(1.1) \quad S_1(x) \sim HA(x), \quad H = \prod_p (1 - g(p))(1 - 1/p)^{-1}.$$

For example, for the twin prime problem, we take $a_n = \Lambda(n + 2)$, $A(x) = x$, $g(d) = \frac{1}{\phi(d)}$ for odd d and $g(d) = 0$ for even d . It is known that $R(\nu)$ holds for all $\nu < 1/2$ (the Bombieri-Vinogradov theorem), and it is conjectured that $R(\nu)$ holds for all $\nu < 1$.

That sieve methods cannot produce (1.1) was discovered by Selberg [10] in the 1940s. His example is $a_n = 1 + \lambda(n)$, where $\lambda(n) = (-1)^k$ if n is the product of k primes (not necessarily distinct). With $A(x) = x$ and $g(d) = 1/d$, $R(\nu)$ holds for all $\nu < 1$, but $a_n = 0$ for prime n and

$$S_1(x) = O(\sqrt{x}) = O(A(x)x^{-1/2}).$$

In a sense, sieve procedures cannot distinguish between numbers with an even number of prime factors and an odd number of prime factors, a property known as the “parity problem”. Bombieri ([1], [2]) clarified things further, showing essentially that knowledge of $R(\nu)$ for all $\nu < 1$ (and no other information about the sequence) implies an asymptotic formula for $\sum_{n \leq x} a_n f(n)$ if and only if f gives “equal weight” to numbers with an even number of prime factors and an odd number of prime factors. The generalized von Mangoldt functions

$$(1.2) \quad \Lambda_k(n) = \sum_{d|n} \mu(d) \log^k(n/d)$$

have this property for $k \geq 2$ (in fact, these functions together with convolutions of the type $\Lambda_{i_1} * \dots * \Lambda_{i_j}$ ($i_1 + \dots + i_j \geq 2$) form a kind of basis for all such f ; see [2] for details). In particular, Bombieri proved that if $R(\nu)$ for all $\nu < 1$, then

$$(1.3) \quad S_k(x) := \sum_{n \leq x} a_n \Lambda_k(n) \sim kHA(x)(\log x)^{k-1}.$$

A different proof of (1.3) was given by Friedlander and Iwaniec [5]. The required conditions on $A(x)$ and g differ in [2] and [5], but they are all trivially satisfied if $A(x) = x$ and $g(d) = 1/d$ (here $H = 1$).

The special case of (1.3) corresponding to $k = 2$ and $a_n = 1$ for all n was earlier proved by Selberg, and it served as a foundation for the first “elementary” proofs of the Prime Number Theorem.

It is natural to inquire what may be deduced from $R(\nu)$ for some *fixed* $\nu < 1$. For twin primes, Bombieri [1] deduced from $R(\nu)$ for $\nu < 1/2$ that

$$1 - c_k \leq \frac{S_k(x)}{kHA(x)(\log x)^{k-1}} \leq 1 + c_k,$$

where c_2, \dots are constants with $c_k \rightarrow 0$ as $k \rightarrow \infty$. We show that knowing $R(\nu)$ for any fixed $\nu < 1$ is not sufficient to deduce (1.3) for any k .

Theorem 1. Fix $\nu \in (0, 1)$. There is a sequence (a_n) which satisfies $R(\nu)$ with $A(x) = x$ and $g(d) = 1/d$, and for which (1.3) fails for every $k \geq 1$. Furthermore, we can specify the manner in which (1.3) fails, constructing (a_n) so that

$$T_k(x) := \frac{S_k(x)}{kx(\log x)^{k-1}}$$

satisfies either (i) $T_k(x) \sim \xi_k$ with $\xi_k < 1$ for every k ; or (ii) $T_k(x) \sim \xi_k$ with $\xi_k > 1$ for every k ; or (iii) for every k , $\limsup_{x \rightarrow \infty} T_k(x) > 1$ and $\liminf_{x \rightarrow \infty} T_k(x) < 1$.

By slightly modifying the construction of the sequence (a_n) , we can create sequences satisfying Theorem 1 for which $a_n \in \{0, 1, 2\}$ for every n .

Recently there was a major breakthrough on the parity problem by Friedlander and Iwaniec [6]. They proved $S_1(x) \sim HA(x)$ under two major assumptions. First, $R(\nu)$ holds for some $\nu > 2/3$. Second, the bilinear sum condition

$$(1.4) \quad \sum_m \left| \sum_{\substack{N < n \leq 2N \\ mn \leq x}} \gamma(n, C) \mu(mn) a_{mn} \right| \ll A(x) (\log x)^{-1996},$$

$$\gamma(n, C) = \sum_{d|n, d \leq C} \mu(d),$$

holds uniformly for $\Delta^{-1}x^{\nu/2} < N < \delta^{-1}\sqrt{x}$, $1 \leq C \leq x^{1-\nu}$, where δ, Δ are parameters depending on x in such a way that $\delta \rightarrow \infty$ and $\frac{\log \delta}{\log \Delta} \rightarrow 0$ as $x \rightarrow \infty$. In [7], they applied this successfully to give an asymptotic formula for the number of primes of the form $a^2 + b^4$ which are $\leq x$. The condition (1.4) strongly eliminates the possibility of the sequence having a ‘‘parity bias’’, meaning a tendency for $\mu(n)a_n$ to be of one sign.

The sequences used to prove Theorem 1 all exhibit a ‘‘global parity bias’’, meaning that

$$(1.5) \quad P(x) = \sum_{n \leq x} a_n \mu(n)$$

is large (or large infinitely often). In light of Selberg’s example and the theorem of Friedlander and Iwaniec, it is natural to inquire whether or not, for each $\nu < 1$, there are sequences (a_n) satisfying $R(\nu)$ and also

$$(1.6) \quad P(x) \ll_B x(\log x)^{-B} \quad (\forall B > 0),$$

but failing (1.3). We cannot yet answer this question entirely, but for all $\nu < 1$, we can construct sequences satisfying $R(\nu)$ and (1.6), but failing (1.3) for all $k \geq 2$. These sequences do satisfy (1.3) for $k = 1$.

Theorem 2. Fix $\nu \in (0, 1)$. There is a sequence (a_n) which satisfies $R(\nu)$ with $A(x) = x$ and $g(d) = 1/d$, satisfies (1.6) and for which (1.3) fails for all $k \geq 2$.

It is an interesting problem to examine the situation if (1.6) is replaced by a stronger condition (but one weaker than (1.4)). One possibility, suggested by C. Hooley, is to postulate that the parity bias in arithmetic progressions is small on average, something like

$$\sum_{d \leq x^\alpha} \left| \sum_{\substack{n \leq x \\ d|n}} \mu(n) a_n \right| \ll_B x(\log x)^{-B} \quad (\forall B > 0).$$

The sequences we construct for the proof of Theorem 1 do satisfy this condition with arbitrary but fixed $\alpha < 1 - \nu$. The case $\alpha + \nu > 1$ remains open.

2. OVERALL PLAN

The only analytic tool we require is the Prime Number Theorem with the de la Vallée Poussin error term. In fact, a much weaker error term would suffice.

Lemma 2.1. *For some positive constant c_0 ,*

$$\sum_{n \leq x} \Lambda(n) = x + O\left(xe^{-c_0\sqrt{\log x}}\right).$$

Assume without loss of generality that $\nu > 1/2$. Let M be an integer, and δ and ϖ real numbers satisfying

$$(2.1) \quad 0 < \delta \leq \frac{1}{3M^2}, \quad M\delta + \frac{1}{M} < \varpi < 1 - \nu.$$

Take x_0 sufficiently large and $c_1 \in (0, c_0)$ (both depending on ν, M and δ). For $j \geq 1$ put

$$(2.2) \quad x_{j+1} = x_j \left(1 + e^{-c_1\sqrt{\log x_j}}\right), \quad I_j = \mathbb{Z} \cap (x_j, x_{j+1}], \quad K_j = |I_j|.$$

In what follows, all constants implied by the O -symbol may depend on ν and M . Dependence on other variables will be indicated by subscripts to the O -symbol. The numbers a_n for $n \in I_j$ will satisfy three basic properties. First,

$$(2.3) \quad 0 \leq a_n \leq 2.$$

Second,

$$(2.4) \quad \sum_{\substack{n \in I_j \\ d|n}} a_n = \frac{K_j}{d} + O\left(\frac{K_j}{d}e^{-c_1\sqrt{\log x_j}}\right) \quad (1 \leq d \leq x_j^{1-\varpi}).$$

Third, for some positive constants θ_k ($k \geq 1$ for Theorem 1, $k \geq 2$ for Theorem 2) which depend on ν, M and δ , and some numbers $\sigma_j \in \{-1, 1\}$ (which we are free to choose), we have

$$(2.5) \quad \sum_{n \in I_j} a_n \Lambda_k(n) = (k + \sigma_j \theta_k) K_j (\log x_j)^{k-1} \left(1 + O_k\left(e^{-c_1\sqrt{\log x_j}}\right)\right).$$

For Theorem 2, we also require that

$$(2.6) \quad \sum_{n \in I_j} a_n \mu(n) = O\left(K_j e^{-c_1\sqrt{\log x_j}}\right).$$

Deducing Theorems 1 and 2 from (2.3)–(2.6) is straightforward. For $d \leq x^\nu$, (2.3) implies

$$A_d(x) = O\left(\frac{x^{\nu+\varpi}}{d}\right) + \sum_{x^{\nu+\varpi} \leq x_j \leq x} \sum_{\substack{n \in I_j \\ d|n}} a_n.$$

By (2.1), if $x_j \geq x^{\nu+\varpi}$, then $d < x^{(1-\varpi)(\nu+\varpi)} \leq x_j^{1-\varpi}$. Thus, by (2.4),

$$\begin{aligned} A_d(x) &= O\left(\frac{x^{\nu+\varpi}}{d}\right) + \sum_j \frac{K_j}{d} \left(1 + O\left(e^{-c_1\sqrt{(\nu+\varpi)\log x}}\right)\right) \\ &= \frac{x}{d} + O\left(\frac{x}{d}e^{-\frac{1}{2}c_1\sqrt{\log x}}\right) = g(d)A(x) + O\left(\frac{A(x)}{d}e^{-\frac{1}{2}c_1\sqrt{\log x}}\right). \end{aligned}$$

Summing on d gives $R(\nu)$. Similarly, (2.6) implies (1.6). From (1.2), we have $\log^k n = (1 * \Lambda_k)(n) \geq \Lambda_k(n)$. Thus, using (2.3) and (2.5), we obtain

$$\begin{aligned} \sum_{n \leq x} a_n \Lambda_k(n) &= O(x(\log x)^{k-2}) + \sum_{\frac{x}{\log^2 x} \leq x_j \leq x} \sum_{n \in I_j} a_n \Lambda_k(n) \\ &= O(x(\log x)^{k-2}) + (\log x)^{k-1} \left(1 + O\left(\frac{\log \log x}{\log x}\right)\right) \sum_{x_j \leq x} (k + \sigma_j \theta_k) K_j \\ &= (\log x)^{k-1} (xk + \theta_k \sum_{x_j \leq x} \sigma_j K_j) + O(x(\log x)^{k-3/2}). \end{aligned}$$

The three types of behavior for $T_k(x)$ in Theorem 1 are obtained by taking (respectively) (i) $\sigma_j = -1$ for all j ; (ii) $\sigma_j = 1$ for all j ; or (iii) $\sigma_j = -1$ if $2^{2^r} < x_j \leq 2^{2^{r+1}}$ for an even r and $\sigma_j = 1$ if $2^{2^r} < x_j \leq 2^{2^{r+1}}$ for an odd r .

It remains, therefore, to construct numbers a_n on each interval I_j satisfying (2.3)–(2.6) as appropriate for Theorems 1 and 2. The basic idea is to start with $a_n = 1$ for all n , then shift around some of the mass from the numbers a_n with n composed of “large” prime factors. This must be done very delicately in order to preserve (2.4), and this is the most complex part of the argument. We will work with smooth functions defined on numbers with a given number of prime factors. Let

$$\begin{aligned} T_r &= \{(u_1, \dots, u_r) : 0 \leq u_1 \leq \dots \leq u_r, u_1 + \dots + u_r = 1\}, \\ U_r &= \{(u_1, \dots, u_r) : u_i \geq 0 (1 \leq i \leq r), u_1 + \dots + u_r = 1\}. \end{aligned}$$

For positive numbers ε, B , let $\mathcal{F}_r(\varepsilon, B)$ be the set of functions $f(u_1, \dots, u_r)$ on U_r that are (i) symmetric in all variables, (ii) zero whenever $\min u_i \leq \varepsilon$, and (iii) f and all first order partial derivatives are at most B in absolute value on U_r . If $n = p_1 \cdots p_r$, the numbers p_i being primes with no assumptions on their relative sizes, then

$$f\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n}\right)$$

is well defined. With these assumptions, we may estimate in a standard way sums over f in terms of integrals.

Lemma 2.2. *Let $f \in \mathcal{F}_r(\varepsilon, B)$, $0 \leq y \leq x$ and x large in terms of ε, r, B . Then*

$$\begin{aligned} \sum_{\substack{p_1, \dots, p_r \\ x \leq n = p_1 \cdots p_r \leq x+y}} f\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n}\right) \\ = \frac{y}{\log x} \int_{U_r} \frac{f(u_1, \dots, u_r)}{u_1 \cdots u_r} + O_{\varepsilon, r, B} \left(\frac{y^2}{x \log x} + x e^{-c_0 \sqrt{\varepsilon^{r-1} \log x}}\right). \end{aligned}$$

When $r = 1$, the integral is $f(1)$.

Proof. Let F denote the sum in the lemma. In this proof, constants implied by the O -symbol may depend on ε, r, B . When $r = 1$, by Lemma 2.1,

$$F = f(1)(\pi(x + y) - \pi(x)) = \frac{y}{\log x} f(1) + O(xe^{-c_0\sqrt{\log x}}).$$

We now proceed by induction on r . Suppose $r \geq 2$ and fix $p_1 \in [x^\varepsilon, 2x^{1-\varepsilon}]$. Writing $n' = p_2 \cdots p_r$ and $v_j = \frac{\log p_j}{\log n'}$ for $2 \leq j \leq r$, we have

$$f\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n}\right) = \left(1 + O\left(\frac{y}{x \log x}\right)\right) g(v_2, \dots, v_r),$$

where

$$g(v_2, \dots, v_r) = f\left(\frac{\log p_1}{\log x}, \frac{\log(x/p_1)}{\log x} v_2, \dots, \frac{\log(x/p_1)}{\log x} v_r\right).$$

We have $g \in \mathcal{F}_{r-1}(\varepsilon, B)$, so by the induction hypothesis,

$$F = \sum_{p_1} \frac{y}{p_1 \log(x/p_1)} \int_{\mathbf{v} \in U_{r-1}} \frac{g(v_2, \dots, v_r)}{v_2 \cdots v_r} + O\left(\frac{y^2}{p_1 x \log x} + \frac{x}{p_1} e^{-c_0\sqrt{\varepsilon^{r-2} \log(x/p_1)}}\right).$$

Since $\log(x/p_1) \geq \varepsilon \log x - 1$, $\sum_{p_1} 1/p_1 \ll 1$ and thus the error terms above total

$$O\left(\frac{y^2}{x \log x} + xe^{-c_0\sqrt{\varepsilon^{r-1} \log x}}\right).$$

By Lemma 2.1 and partial summation, for a fixed v_2, \dots, v_r , we have

$$\begin{aligned} & \sum_{p_1} \frac{f\left(\frac{\log p_1}{\log x}, \frac{\log(x/p_1)}{\log x} v_2, \dots, \frac{\log(x/p_1)}{\log x} v_r\right)}{p_1 \log(x/p_1)} \\ &= \int_{x^\varepsilon}^{2x^{1-\varepsilon}} \frac{f\left(\frac{\log t}{\log x}, \frac{\log(x/t)}{\log x} v_2, \dots, \frac{\log(x/t)}{\log x} v_r\right)}{t \log t \log(x/t)} dt + O(e^{-c_0\sqrt{\varepsilon \log x}}) \\ &= \frac{1}{\log x} \int_\varepsilon^{1-\varepsilon/2} \frac{f(u, (1-u)v_2, \dots, (1-u)v_r)}{u(1-u)} du + O(e^{-c_0\sqrt{\varepsilon \log x}}). \end{aligned}$$

Therefore,

$$\begin{aligned} F &= \frac{y}{\log x} \int_{\substack{v_2 + \dots + v_r = 1 \\ 0 < u < 1}} \frac{f(u, (1-u)v_2, \dots, (1-u)v_r)}{u(1-u)v_2 \cdots v_r} \\ &+ O\left(\frac{y^2}{x \log x} + xe^{-c_0\sqrt{\varepsilon^{r-1} \log x}}\right). \end{aligned}$$

Making the change of variables $u_1 = u$, $u_j = (1-u)v_j$ ($2 \leq j \leq r$) gives the lemma. □

3. THE CONSTRUCTION ON I_j

To facilitate working with sets of numbers with prime factors in specific ranges, we adopt some special notation. A *partition* is a non-decreasing sequence of positive integers $\alpha = (\alpha_1, \dots, \alpha_r)$ (also thought of as a “multi-set”). Let $|\alpha| = r$ and $\Sigma(\alpha) = \alpha_1 + \dots + \alpha_r$. Let $\text{perm}(\alpha)$ be the number of permutations of the numbers in α , e.g., $\text{perm}(1, 1, 2, 3) = 12$. Let

$$\mathbf{1}_n = \overbrace{(1, \dots, 1)}^n$$

and let $P(m)$ be the set of all partitions of m (all α with $\Sigma(\alpha) = m$). Let \mathbf{E} denote the empty partition ($|\mathbf{E}| = 0$ and $\Sigma(\mathbf{E}) = 0$). The notation $\alpha \subseteq \beta$ means that each number in β occurs at least as many times as the number occurs in α , and $\alpha + \beta$ is the partition consisting of all the parts of α and of β , so in particular, $|\alpha + \beta| = |\alpha| + |\beta|$ and $\Sigma(\alpha + \beta) = \Sigma(\alpha) + \Sigma(\beta)$. Also, if $\alpha \subseteq \beta$, then $\beta - \alpha$ is defined by $\alpha + (\beta - \alpha) = \beta$.

For brevity, write $x = x_j, K = K_j, I = I_j$. For $1 \leq i \leq M$, let \mathcal{P}_i be the set of primes in the interval $[x^{i(1/M-\delta)}, x^{i(1/M+\delta)}]$. For each partition $\alpha = (\alpha_1, \dots, \alpha_r)$, let

$$\mathcal{D}_\alpha = \{p_1 \cdots p_r : p_i \in \mathcal{P}_{\alpha_i} (1 \leq i \leq r)\}, \quad \mathcal{C}_\alpha = \mathcal{D}_\alpha \cap I.$$

In particular, $\mathcal{D}_\mathbf{E} = \{1\}$ and $\mathcal{C}_{(M)}$ is the set of primes in I . Also, by (2.1), \mathcal{C}_α is empty unless $\alpha \in P(M)$.

Let $c_1 = c_0(2M)^{-M}$. We put $a_n = 1 + b_n$, where $|b_n| \leq 1$, and $b_n = 0$ unless n lies in some \mathcal{C}_α with $\alpha \in P(M)$. Thus, if $1 < d \leq x^{1-\varpi}$, then

$$\sum_{\substack{n \in I \\ d|n}} a_n = \frac{K}{d} + O(1)$$

unless $d \in \mathcal{D}_\beta$ for some $\beta = (\beta_1, \dots, \beta_s)$. In this case $x^{(1/M-\delta)(\beta_1+\dots+\beta_s)} \leq d \leq x^{1-\varpi}$, which by (2.1) implies $\beta_1+\dots+\beta_s \leq M-2$. Let $Q = \mathbf{E} \cup P(1) \cup \dots \cup P(M-2)$. To obtain (2.4), it suffices to prove that for each $\beta \in Q$ and each $d \in \mathcal{D}_\beta$,

$$(3.1) \quad \sum_{\substack{\alpha \in P(M) \\ \beta \subseteq \alpha}} \sum_{\substack{n \in \mathcal{C}_\alpha \\ d|n}} b_n = O\left(\frac{K}{d} e^{-c_1 \sqrt{\log x}}\right).$$

This system of inequalities has the trivial solution $b_n = 0$ for all n , but we need a solution with $|b_n| \gg 1$ on average in order to obtain (2.5).

For $1 \leq i \leq M$, let $J_i = [i(1/M - \delta), i(1/M + \delta)]$. For each $\alpha = (\alpha_1, \dots, \alpha_r) \in P(M)$, suppose $f_\alpha \in \mathcal{F}_r(\frac{1}{2M}, B)$ is supported on $T_r \cap (J_{\alpha_1} \times \dots \times J_{\alpha_r})$ and the symmetric regions in U_r . For $n \in \mathcal{C}_\alpha, n = p_1 \cdots p_r$, let

$$(3.2) \quad b_n = f_\alpha \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n} \right).$$

Suppose that $\beta = (\beta_1, \dots, \beta_s) \in Q$ with $\beta \subseteq \alpha$. Then $r \geq s + 1$. Let $d = p_1 \cdots p_s \in \mathcal{D}_\beta$ with $p_i \in \mathcal{P}_{\beta_i} (1 \leq i \leq s)$ and put $v_i = \frac{\log p_i}{\log x}$ for $1 \leq i \leq s$. We have

$$f_\alpha \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n} \right) = g \left(\frac{\log p_{s+1}}{\log(n/d)}, \dots, \frac{\log p_r}{\log(n/d)} \right) + O(e^{-c_1 \sqrt{\log x}}),$$

where

$$g(w_1, \dots, w_{r-s}) = f(v_1, \dots, v_s, \frac{\log(x/d)}{\log x} w_1, \dots, \frac{\log(x/d)}{\log x} w_{r-s}).$$

Since $g \in \mathcal{F}_{r-s}(\frac{1}{2M}, B)$, Lemma 2.2 implies that

$$\begin{aligned} \sum_{\substack{n \in \mathcal{C}_\alpha \\ d|n}} b_n &= \frac{K}{(r-s)! d \log(x/d)} \int_{\mathbf{w} \in U_{r-s}} \frac{g(w_1, \dots, w_{r-s})}{w_1 \cdots w_{r-s}} + O\left(\frac{x}{d} e^{-2c_1 \sqrt{\log x}}\right) \\ &= \frac{K}{(r-s)! d \log x} \int_{\mathbf{u} \in V_{r-s}(1-v_1-\dots-v_s)} \frac{f_\alpha(v_1, \dots, v_s, u_1, \dots, u_{r-s})}{u_1 \cdots u_{r-s}} \\ &\quad + O\left(\frac{K}{d} e^{-c_1 \sqrt{\log x}}\right), \end{aligned}$$

where $V_t(A) = \{(u_1, \dots, u_t) : u_i \geq 0 \forall i, \sum u_i = A\}$.

Therefore, to prove (3.1), it suffices to find functions f_α so that for all $\beta = (\beta_1, \dots, \beta_s) \in Q$ and $(v_1, \dots, v_s) \in J_{\beta_1} \times \dots \times J_{\beta_s}$, we have

$$(3.3) \quad \sum_{\mu \in P(M - \Sigma(\beta))} \frac{1}{|\mu|!} \int_{\mathbf{u} \in V_{|\mu|(1-v_1-\dots-v_s)}} \frac{f_{\beta+\mu}(v_1, \dots, v_s, u_1, \dots, u_{r-s})}{u_1 \cdots u_{r-s}} = 0.$$

When $M \geq 6$, $|Q| > |P(M)|$ (i.e., there are more equations than functions), but there is enough structure in the system (3.3) to find a nontrivial solution. In fact, once $f_{\mathbf{1}_M}$ is chosen, the other functions f_α are uniquely determined by (3.3), but we do not need to prove this. Suppose $\alpha = (\alpha_1, \dots, \alpha_r) \in P(M)$, ρ is a permutation of α and $\mathbf{v} \in J_{\rho_1} \times \dots \times J_{\rho_r}$. For some constant e_α , define

$$(3.4) \quad f_\alpha(\mathbf{v}) = e_\alpha v_1 \cdots v_r \int_{(3.5)} \frac{f_{\mathbf{1}_M}(\mathbf{w})}{\prod w_{ij}},$$

where the integration is over the set of $\mathbf{w} = \{w_{ij} : 1 \leq j \leq r, 1 \leq i \leq \rho_j\} \in J_1^M$ with

$$(3.5) \quad \sum_{i=1}^{\rho_j} w_{ij} = v_j \quad (1 \leq j \leq r).$$

For example, if $\alpha = (1, 1, 2, 3)$, $v_1 \in J_1$, $v_2 \in J_2$, $v_3 \in J_3$ and $v_4 \in J_1$, we have

$$f_\alpha(\mathbf{v}) = e_\alpha v_1 v_2 v_3 v_4 \int_{\substack{w_{11}=v_1 \\ w_{12}+w_{22}=v_2 \\ w_{13}+w_{23}+w_{33}=v_3 \\ w_{14}=v_4}} \frac{f_{\mathbf{1}_7}(w_{11}, w_{12}, w_{22}, w_{13}, w_{23}, w_{33}, w_{14})}{w_{11} w_{12} w_{22} w_{13} w_{23} w_{33} w_{14}}.$$

For consistency, set $e_{\mathbf{1}_M} = 1$.

Next we show that substituting (3.4) into (3.3) reduces the problem to solving a system of equations in the numbers e_α . Fix $\beta = (\beta_1, \dots, \beta_s) \in Q$ and $\mu = (\mu_1, \dots, \mu_{r-s}) \in P(M - \Sigma(\beta))$. Suppose that $v_j \in J_{\beta_j}$ ($1 \leq j \leq s$) and let ρ be a permutation of μ . Take \mathbf{u} so that $\sum v_i + \sum u_i = 1$ and $u_i \in J_{\rho_i}$ ($1 \leq i \leq r - s$). Because $f_{\beta+\mu}$ is symmetric in all variables, for each ρ the contribution to the integral in (3.3) is identical. In other words, the integral in (3.3) equals $\text{perm}(\mu)$ times the integral over those $\mathbf{u} \in J_{\mu_1} \times \dots \times J_{\mu_{r-s}}$. For such \mathbf{u} , (3.4) implies

$$f_{\beta+\mu}(\mathbf{v}, \mathbf{u}) = e_{\beta+\mu} v_1 \cdots v_s u_1 \cdots u_{r-s} \int_{(3.6)} \frac{f_{\mathbf{1}_M}(\mathbf{w}, \mathbf{z})}{\prod_{i,j} w_{ij} \prod_{i,j} z_{ij}},$$

where the integral is over the variables $w_{ij}, z_{ij} \in J_1$ satisfying

$$(3.6) \quad \sum_{i=1}^{\beta_j} w_{ij} = v_j \quad (1 \leq j \leq s); \quad \sum_{i=1}^{\mu_j} z_{ij} = u_j \quad (1 \leq j \leq r - s).$$

Thus, with β and \mathbf{v} fixed,

$$\int_{\substack{\mathbf{u} \in V_{|\mu|(1-v_1-\dots-v_s)} \\ u_j \in J_{\mu_j} (1 \leq j \leq r-s)}} \frac{f_{\beta+\mu}(\mathbf{v}, \mathbf{u})}{u_1 \cdots u_{r-s}} = e_{\beta+\mu} v_1 \cdots v_s \int_{(3.6)} \frac{f_{\mathbf{1}_M}(\mathbf{w}, \mathbf{z})}{\prod w_{ij} \prod z_{ij}}.$$

Since u_1, \dots, u_{r-s} are dependent variables in the integral on the right side, the left side is actually independent of $\boldsymbol{\mu}$. Thus, with (3.4), (3.3) follows from the system

$$(3.7) \quad \sum_{\boldsymbol{\mu} \in P(M-\Sigma(\boldsymbol{\beta}))} \frac{\text{perm}(\boldsymbol{\mu})}{|\boldsymbol{\mu}|!} e_{\boldsymbol{\beta}+\boldsymbol{\mu}} = 0 \quad (\boldsymbol{\beta} \in Q), \quad e_{\mathbf{1}_M} = 1.$$

As noted before, (3.7) has more equations than variables when $M > 6$, but there is a simple solution (again we do not need to prove uniqueness, but it is straightforward), namely

$$(3.8) \quad e_{\boldsymbol{\alpha}} = \frac{(-1)^{\Sigma(\boldsymbol{\alpha})+|\boldsymbol{\alpha}|}}{\alpha_1 \cdots \alpha_r}, \quad \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r).$$

With (3.8), $e_{\boldsymbol{\beta}+\boldsymbol{\mu}} = e_{\boldsymbol{\beta}}e_{\boldsymbol{\mu}}$ for all $\boldsymbol{\beta}, \boldsymbol{\mu}$, so (3.7) is equivalent to $\gamma_m = 0$ ($2 \leq m \leq M$), where

$$\gamma_m := \sum_{\substack{\boldsymbol{\mu} \in P(m) \\ \boldsymbol{\mu}=(\mu_1, \dots, \mu_r)}} \frac{(-1)^{|\boldsymbol{\mu}|} \text{perm}(\boldsymbol{\mu})}{|\boldsymbol{\mu}|! \mu_1 \cdots \mu_r} = \sum_{r=1}^m \frac{(-1)^r}{r!} \sum_{\substack{d_1+\dots+d_r=m \\ d_i \geq 1(1 \leq i \leq r)}} \frac{1}{d_1 \cdots d_r}.$$

This follows by considering the generating function

$$G(z) = \sum_{m=1}^{\infty} \gamma_m z^m.$$

Since $|\gamma_m| \leq \sum_{r=1}^m \binom{m-1}{r-1} = 2^{m-1}$, $G(z)$ has radius of convergence $\geq 1/2$. Thus, for $|z| \leq 1/3$,

$$\begin{aligned} G(z) &= \sum_{r=1}^{\infty} \frac{(-1)^r}{r!} \sum_{d_1, \dots, d_r \geq 1} \frac{z^{d_1+\dots+d_r}}{d_1 \cdots d_r} \\ &= \sum_{r=1}^{\infty} \frac{(-1)^r}{r!} \left(\sum_{d=1}^{\infty} \frac{z^d}{d} \right)^r \\ &= \sum_{r=1}^{\infty} \frac{(-1)^r}{r!} (-\log(1-z))^r = e^{\log(1-z)} - 1 = -z, \end{aligned}$$

which proves (3.7). As noted earlier, (3.7) implies (3.3), which implies (3.1), which implies (2.4).

Modulo the choice of function $f_{\mathbf{1}_M}$, we have constructed our numbers b_n . The following theorem sums up the properties we are interested in.

Theorem 3. Fix M, ϖ, δ so that (2.1) is satisfied and also $\delta \leq (2M)^{-M}$. Let B be large depending on M, δ . Let $f_{\mathbf{1}_M} \in \mathcal{F}_M(\frac{1}{2M}, B)$ with $|f_{\mathbf{1}_M}(\mathbf{u})| \leq 1$ for all $\mathbf{u} \in U_M$. For every $\boldsymbol{\alpha} \in P(M)$, define $e_{\boldsymbol{\alpha}}$ by (3.8), define $f_{\boldsymbol{\alpha}}$ by (3.4), b_n by (3.2), and put $a_n = 1 + b_n$. Then, for each interval $I = I_j$, (2.3) and (2.4) are satisfied, plus we have

$$(3.9) \quad \sum_{n \in I} a_n \Lambda_k(n) = K(\log x)^{k-1} [k + (-1)^{M+1} Z_k + O(e^{-c_1 \sqrt{\log x}})] \quad (k \geq 1)$$

and

$$(3.10) \quad \sum_{n \in I} a_n \mu(n) = \frac{K}{\log x} [(-1)^M Z_0 + O(e^{-c_1 \sqrt{\log x}})],$$

where

$$Z_k := \int_{\mathbf{u} \in U_M} \frac{u_1^k f_{\mathbf{1}_M}(\mathbf{u})}{u_1 \cdots u_M} = \frac{1}{M} \int_{\mathbf{u} \in U_M} \frac{u_1^k + \cdots + u_M^k}{u_1 \cdots u_M} f_{\mathbf{1}_M}(\mathbf{u}).$$

Formula (3.10) also holds with $\mu(n)$ replaced by $\lambda(n)$.

Proof. We have already seen that (2.4) is satisfied. Let $\alpha \in P(M)$, $\alpha \neq \mathbf{1}_M$. By (3.8), $|e_\alpha| \leq 1$, so by (3.4),

$$|f_\alpha| \leq (1/M - \delta)^{-M} \delta^{M-|\alpha|} \leq \delta(2M)^M \leq 1.$$

Next, by (1.2) and Lemma 2.2, for each $\alpha = (\alpha_1, \dots, \alpha_r) \in P(M)$, we have

$$\begin{aligned} \sum_{n \in \mathcal{C}_\alpha} b_n \Lambda_k(n) &= \frac{1}{r!} \sum_{\substack{p_1, \dots, p_r \\ n=p_1 \cdots p_r \in \mathcal{C}_\alpha}} f_\alpha \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_r}{\log n} \right) \\ &\quad \times \sum_{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\}} (-1)^{r-\varepsilon_1-\dots-\varepsilon_r} \log^k(p_1^{\varepsilon_1} \cdots p_r^{\varepsilon_r}) \\ &= \frac{(-1)^r \text{perm}(\alpha)}{r!} K(\log x)^{k-1} \left[O(e^{-c_1 \sqrt{\log x}}) \right. \\ &\quad \left. + \int_{\substack{\mathbf{v} \in U_r \\ v_i \in J_{\alpha_i} (1 \leq i \leq r)}} \frac{f_\alpha(\mathbf{v})}{v_1 \cdots v_r} \sum_{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} \left(\sum_{j=1}^r \varepsilon_j v_j \right)^k \right]. \end{aligned}$$

By (3.4) and the fact that $f_{\mathbf{1}_M}$ is symmetric in all variables, we obtain

$$\begin{aligned} \sum_{n \in \mathcal{C}_\alpha} b_n \Lambda_k(n) &= \frac{\text{perm}(\alpha) e_\alpha (-1)^{|\alpha|}}{|\alpha|!} K(\log x)^{k-1} \left[O(e^{-c_1 \sqrt{\log x}}) \right. \\ &\quad \left. + \int_{\mathbf{w} \in U_M} \frac{f_{\mathbf{1}_M}(\mathbf{w})}{\prod w_{i,j}} \sum_{\boldsymbol{\varepsilon}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} \left(\sum_{j=1}^{|\alpha|} \varepsilon_j (w_{j,1} + \dots + w_{j,\alpha_j}) \right)^k \right] \\ &= \frac{\text{perm}(\alpha) e_\alpha (-1)^{|\alpha|}}{|\alpha|!} K(\log x)^{k-1} \left[O(e^{-c_1 \sqrt{\log x}}) \right. \\ &\quad \left. + \int_{\mathbf{u} \in U_M} \frac{f_{\mathbf{1}_M}(\mathbf{u})}{u_1 \cdots u_M} \sum_{\substack{N=1 \\ \varepsilon_1 \alpha_1 + \dots + \varepsilon_r \alpha_r = N}}^M \left(\sum_{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} \right) (u_1 + \dots + u_N)^k \right]. \end{aligned}$$

Summing on $\alpha \in P(M)$ and using (3.8) gives

$$\begin{aligned} (3.11) \quad \sum_{n \in I} b_n \Lambda_k(n) &= K(\log x)^{k-1} \left[O(e^{-c_1 \sqrt{\log x}}) \right. \\ &\quad \left. + (-1)^M \sum_{N=1}^M W(M, N) \int_{\mathbf{u} \in U_M} \frac{f_{\mathbf{1}_M}(\mathbf{u})}{u_1 \cdots u_M} (u_1 + \dots + u_N)^k \right], \end{aligned}$$

where

$$(3.12) \quad W(M, N) = \sum_{r=1}^M \frac{1}{r!} \sum_{\substack{d_1 + \dots + d_r = M \\ d_i \geq 1 \forall i}} \frac{1}{d_1 \cdots d_r} \sum_{\substack{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\} \\ \varepsilon_1 d_1 + \dots + \varepsilon_r d_r = N}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r}.$$

By examining the generating function, we next prove that

$$(3.13) \quad \begin{aligned} W(M, N) &= 0 \quad (2 \leq N \leq M), \\ W(M, 1) &= -1, W(M, 0) = 1 \quad (M \geq 1). \end{aligned}$$

For $\max(|x|, |y|) < 1/3$, we have

$$\begin{aligned} &\sum_{M=1}^{\infty} \sum_{N=0}^M W(M, N)x^M y^N \\ &= \sum_{r=1}^{\infty} \frac{1}{r!} \sum_{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} \sum_{d_1, \dots, d_r \geq 1} \left(\frac{x^{d_1} y^{\varepsilon_1 d_1}}{d_1}\right) \dots \left(\frac{x^{d_r} y^{\varepsilon_r d_r}}{d_r}\right) \\ &= \sum_{r=1}^{\infty} \frac{1}{r!} \sum_{\varepsilon_1, \dots, \varepsilon_r \in \{0,1\}} (-1)^{\varepsilon_1 + \dots + \varepsilon_r} (-\log(1 - xy^{\varepsilon_1})) \dots (-\log(1 - xy^{\varepsilon_r})) \\ &= \sum_{r=1}^{\infty} \frac{1}{r!} (\log(1 - xy) - \log(1 - x))^r \\ &= -1 + \frac{1 - xy}{1 - x} = (1 - y)(x + x^2 + x^3 + \dots). \end{aligned}$$

This proves (3.13), and, together with (3.11), completes the proof of (3.9). For the sum of $\mu(n)b_n$, we obtain a similar expression corresponding to the “ $N = 0$ ” term. Thus

$$\sum_{n \in I} b_n \mu(n) = \frac{(-1)^M K}{\log x} \left(W(M, 0)I_0 + O(e^{c_1 \sqrt{\log x}}) \right).$$

The asymptotic (3.10) now follows from (3.13). Finally, $b_n = 0$ if n has a prime factor $< x^{1/M-\delta}$. Hence, when $b_n \mu(n) \neq b_n \lambda(n)$, n is divisible by the square of a prime $\geq x^{1/M-\delta}$. The number of such $n \leq x$ is $\ll x^{1-1/M+\delta}$ and this proves the final claim. \square

Proof of Theorems 1, 2. Define

$$\ell(v_1, \dots, v_M; \xi) = \max(0, \xi^{-4}(\xi^2 - v_1^2 - \dots - v_M^2)^2),$$

which is nonzero only when $|v_i| \leq \xi$ for each i . To prove Theorem 1, take in Theorem 3,

$$f_{\mathbf{1}_M}(\mathbf{u}) = (-1)^{M+1} \sigma_j \ell(u_1 - 1/M, \dots, u_M - 1/M; \delta).$$

For $\mathbf{u} \in U_M$, $u_1 + \dots + u_M = 1$ and thus $Z_1 = Z_0/M$. To prove Theorem 2, we must exhibit a function $f_{\mathbf{1}_M}$ so that $Z_0 = 0$ and $Z_k \neq 0$ for $k \geq 2$. Let M be even and put $\mathbf{w} = (\frac{1}{M}, \dots, \frac{1}{M})$. Let V be the set of vectors in \mathbb{R}^M with exactly $M/2$ components equal to $\delta/2$ and $M/2$ components equal to $-\delta/2$. We will take

$$f_{\mathbf{1}_M}(\mathbf{u}) = u_1 \dots u_M \left[\ell(\mathbf{u} - \mathbf{w}; \delta^3) - \binom{M}{M/2}^{-1} \sum_{\mathbf{v} \in V} \ell(\mathbf{u} - \mathbf{w} - \mathbf{v}; \delta^3) \right].$$

Letting

$$J = \int_{v_1 + \dots + v_M = 0} \ell(\mathbf{v}; \delta^3),$$

it follows that

$$\begin{aligned} Z_k &\leq M\left(\frac{1}{M} + \delta^3\right)^k J - \left[\frac{M}{2}\left(\frac{1}{M} - \frac{\delta}{2} - \delta^3\right)^k + \frac{M}{2}\left(\frac{1}{M} + \frac{\delta}{2} - \delta^3\right)^k\right] J \\ &= \frac{J}{2M^{k-1}} \left[2(1 + \lambda)^k - (1 - \varepsilon - \lambda)^k - (1 + \varepsilon - \lambda)^k\right], \end{aligned}$$

where $\varepsilon = \frac{\delta}{2M}$ and $\lambda = \frac{\delta^3}{M}$. Since x^k has convex derivative for $x > 0$, we have

$$\begin{aligned} &(1 + \varepsilon - \lambda)^k - 2(1 + \lambda)^k + (1 - \varepsilon - \lambda)^k \\ &\geq (\varepsilon - 2\lambda)k(1 + \varepsilon/2)^{k-1} - (\varepsilon + 2\lambda)k(1 + \lambda)^{k-1} \\ &\geq k(1 + \lambda)^{k-2} [(\varepsilon - 2\lambda)(1 + \varepsilon/2) - (\varepsilon + 2\lambda)(1 + \lambda)] \\ &= k(1 + \lambda)^{k-1} \left(\frac{\delta^2}{8M^2} + O\left(\frac{\delta^3}{M}\right) \right). \end{aligned}$$

This proves $Z_k < 0$ for $k \geq 2$ if δ is small enough, and completes the proof. \square

ACKNOWLEDGEMENT

The author thanks John Friedlander for helpful conversations and Denka Kutza-rova for help constructing the function f_{1M} for Theorem 2.

REFERENCES

- [1] E. Bombieri, *On twin almost-primes*, Acta Arith. **28** (1975), 177–193, 457–461. MR53:300
- [2] E. Bombieri, *The asymptotic sieve*, Mem. Acad. Naz. dei XL, **1/2** (1976), 243–269. MR58:10799
- [3] V. Brun, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Archiv for Math. og Naturvid. B **34** (1915), no. 8, 19 pp.
- [4] V. Brun, *Le crible d’Eratosthène et le théorème de Goldbach*, Skr. Norske Vid.-Akad. Kristiania I. **1920**, no. 3, 36 pp.
- [5] J. Friedlander and H. Iwaniec, *Bombieri’s sieve*, in Analytic Number Theory, Proc. Halberstam Conf., Allerton Park, Illinois, June 1995, ed. B. C. Berndt et al. (1996), pp. 411–430, Birkhäuser (Boston). MR97m:11117
- [6] J. Friedlander and H. Iwaniec, *Asymptotic sieve for primes*, Annals of Math. (2) **148** (1998), 1041–1065. MR2000c:11150b
- [7] J. Friedlander and H. Iwaniec, *The polynomial $X^2 + Y^4$ captures its primes*, Annals of Math. (2) **148** (1998), 945–1040. MR2000c:11150a
- [8] G. Greaves, *Sieves in Number Theory*, Springer-Verlag, Ergebnisse der Mathematik und ihrer Grenzgebiete vol. 43 (2001). MR2002i:11092
- [9] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London (1974). MR54:12689
- [10] A. Selberg, *The general sieve method and its place in prime-number theory*, Proc. Internat. Congress of Math., Cambridge, Mass. (1950), Amer. Math. Soc. (Providence, RI) (1952) **1**, 286–292. MR13:438d

DEPARTMENT OF MATHEMATICS, 1409 WEST GREEN SREET, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, ILLINOIS 61801