SIEVE METHODS LECTURE NOTES, SPRING 2023

KEVIN FORD

1. Basic sieve methods and applications

A *sieve* is a technique for bounding the size of a set after the elements with "undesirable properties" (usually of a number theoretic nature) have been removed. The undesirable properties could be divisibility by a prime from a given set, other multiplicative constraints (divisibility by a perfect square for example) or inclusion in a set of residue classes. Inclusion-exclusion yields an exact formula, however for k properties this produces 2^k summands which is usually too much to effectively deal with. A *sieve* is a procedure to estimate the number of "desirable" elements of the set using $k^{O(1)}$ summands. While inexact, oftentimes the sieve is capable of estimating the size very accurately.

The original sieve is, of course, the *Sieve of Eratosthenes*, the familiar process of creating a table of prime numbers by systematically removing those integers divisible by small primes (but keeping the primes themselves). The modern sieve was created by Viggo Brun in the period 1915-1922 as a way of attacking famous unsolved problems such as Golbach's Conjecture and the Twin Prime problem (both, so far, unsuccessfully). Sieve methods have since found enormous application in number theory, often used as tools in many other types of problems, e.g. in studying Diophantine equations.

1.1. **Notational conventions.** $\tau(n)$ is the number of positive divisors of n
$\omega(n)$ is the number of distinct prime factors of n
$\Omega(n)$ is the number of prime factors of n counted with multiplicity
$\mu(n)$ is the Möbius's function; $\mu(n) = (-1)^{\omega(n)}$ if n is squarefree and $\mu(n) = 0$ otherwise.
$P^+(n)$ is the largest prime factor of n; $P^+(1) = 0$ by convention
$P^-(n)$ is the smallest prime factor of n; $P^-(1) = \infty$ by convention
$\mathscr{P}(z)$ is the set of positive squarefree integers composed only of primes $\leqslant z$
$\Lambda(n)$ denotes the von Mangoldt function
$\mathbb{1}_X$ is the indicator function of the statement X or of the set X
the symbol p, with or without subscripts, always denotes a prime
$\pi(x; q, a)$ is the number of primes $p \leqslant x$ in the progression $a \mod q$.
$\mathbb{P}$ denotes probability and $\mathbb{E}$ expectation
$\log_k x$ is the k-th iterate of the logarithm of x

1.2. **General sieve setup.** A *sieve problem* is a probability space $(\mathcal{A}, \mathcal{F}, \mathbb{P})$, together with a "total mass" quantity M and events $\mathcal{A}_p$, one for each prime p. The "sifting function" is

$$S(\mathcal{A}, z) = M \cdot \mathbb{P}(\text{not } \mathcal{A}_p, \forall p \leqslant z).$$

Oftentimes $\mathcal{A}$ will be a finite set of integers, with $\mathbb{P}(n \in \mathcal{A}) = 1/|\mathcal{A}|$ for each $n \in \mathcal{A}$ (the uniform probability measure), $M = |\mathcal{A}|$ and $\mathcal{A}_p$ is the event that $p|n$, that is, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. This is the standard "small sieve" problem, where $S(\mathcal{A}, z)$ is the number of $n \in \mathcal{A}$ with $p \nmid n$ for all $p \leqslant z$; these are called the "unsifted numbers". In the above definitions, M can be anything, but in practice it has some arithmetical meaning.

Some specific examples:

- **Eratosthenes sieve for primes.** $\mathcal{A} = [1, x] \cap \mathbb{N}$, uniform probabilities on $\mathcal{A}$, $M = |\mathcal{A}| = \lfloor x \rfloor$, and $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. Then $S(\mathcal{A}, \sqrt{x}) = \pi(x) - \pi(\sqrt{x}) + 1$, as the unsifted elements are the primes in $(\sqrt{x}, x]$ together with the number 1.
- **Twin primes.** $\mathcal{A} = [1, x] \cap \mathbb{N}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n(n+2)\}$. The unsifted numbers are numbers n such that $n(n+2)$ has no prime factor $\leqslant z$. In particular, $S(\mathcal{A}, \sqrt{x+2})$ counts $k \in (\sqrt{x+2}, x]$ for which *both* k and $k+2$ are prime.

  Equivalently, $\mathcal{A}_p$ is the set of n that avoid the residue classes $0 \mod p$ and $-2 \mod p$.
- **Twin primes, weighted version.** $\mathcal{A} = [2, x] \cap \mathbb{N}$, and

$$\mathbb{P}(n \in \mathcal{A}) = \frac{\Lambda(n+2)}{M}, \quad M = \sum_{2 \leqslant n \leqslant x} \Lambda(n+2),$$

  $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. Here $M = \sum_{2 \leqslant n \leqslant x} \Lambda(n+2) \sim x$ by the Prime Number Theorem, and

$$\begin{aligned} S(\mathcal{A}, \sqrt{x}) &= \sum_{\substack{2 \leqslant n \leqslant x \\ P^-(n) > \sqrt{x}}} \Lambda(n+2) = \sum_{\sqrt{x} < p \leqslant x} \Lambda(p+2) \\ &= \sum_{\substack{\sqrt{x} < p \leqslant x \\ p+2 \text{ prime}}} \log(p+2) + O(x^{1/2}), \end{aligned}$$

  the error term coming from terms $p + 2 = q^b$ where q is prime and $b \geqslant 2$.
- **Prime tuples**. Let $a_1, \cdots, a_k \in \mathbb{N}$ and $b_1, \ldots, b_k \in \mathbb{Z}$. Put $\mathcal{A} = [1, x] \cap \mathbb{N}$, and[1]

$$\mathcal{A}_p = \{n \in \mathcal{A} : p|(a_1 n + b_1) \cdots (a_k n + b_k)\}.$$

  For an appropriate $c > 0$, which depends on $a_1, a_2, \ldots, a_k, b_k$, $S(\mathcal{A}, c\sqrt{x})$ counts those $n \leqslant x$ for which $a_1 n + b_1, \ldots, a_k n + b_k$ are simultaneously prime.
- **Prime values of a polynomial.** Let $f : \mathbb{Z} \to \mathbb{Z}$ be an irreducible polynomial of degree $h \geqslant 1$, put $\mathcal{A} = [1, x] \cap \mathbb{Z}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|f(n)\}$. Let C be large enough so that $|f(n)| \leqslant Cn^h$ for all $n \geqslant 1$. Then, as before, $S(\mathcal{A}, \sqrt{Cx^h})$ captures values of n for which $f(n)$ is prime.
- **Goldbach's problem.** Let N be an even, positive integer, put $\mathcal{A} = \{1, 2, \ldots, N-1\}$, $M = |\mathcal{A}| = N - 1$, and $\mathcal{A}_p = \{k \in \mathcal{A} : p|k(N-k)\}$. Then $S(\mathcal{A}, \sqrt{N})$ counts numbers $k \in (\sqrt{N}, N]$ for which both k and $N - k$ are prime. In particular, $S(\mathcal{A}, \sqrt{N}) > 0$ implies that N is the sum of two primes. If one shows this for all $N \geqslant 4$, one deduces Goldbach's Conjecture.
- **Primes in an arithmetic progression.** Fix coprime positive integers a and q, let

$$\mathcal{A} = \{1 \leqslant n \leqslant x : n \equiv a \pmod{q}\},$$

[1]Unless otherwise specified, from now on whenever $\mathcal{A}$ is a finite set, the probability measure on $\mathcal{A}$ will be the uniform measure, and M will be the number of elements of $\mathcal{A}$.

$\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. Then $S(\mathcal{A}, \sqrt{x})$ captures primes in $(\sqrt{x}, x]$ that are in the arithmetic progression $a \mod q$.

- **Sums of two squares.** $\mathcal{A} = [1, x] \cap \mathbb{Z}$,

$$\mathcal{A}_p = \begin{cases} \{n \in \mathcal{A} : p^e \| n \text{ for some odd } e\} & p \equiv 3 \pmod 4 \\ \emptyset & \text{otherwise.} \end{cases}$$

  Then $S(\mathcal{A}, x)$ counts integers $n \leqslant x$, for which we don't have $p^e \| n$ for any prime $p \equiv 3 \pmod 4$ and odd exponent e. That is, $S(\mathcal{A}, x)$ is the number of integers $n \leqslant x$ which are the sum of two squares.
- **Sieve by multiple residue classes.** Let M, N be two integers, $\mathcal{A} = \{N + 1, M + 2, \dots, N + M\}$ and for each prime p let $\mathcal{I}_p$ be some subset (possibly empty) of the residue classes modulo p. Put

$$\mathcal{A}_p = \mathcal{A} \cap \mathcal{I}_p.$$

  Here $S(\mathcal{A}, z)$ counts the integers in $(N, M+N]$ avoiding all the residue classes $\mathcal{I}_p$ for primes $p \leqslant z$. If $|\mathcal{I}_p|$ is bounded or bounded on average, then this is a very general sieving problem of "small sieve" type, whereas if $|\mathcal{I}_p|$ is unbounded on average, the problem falls under the umbrella of the "large sieve". The case of prime values of a polynomial, see above, is a special case with

$$\mathcal{I}_p = \{n \in \mathbb{Z}/p\mathbb{Z} : f(n) \equiv 0 \pmod p\}.$$

- **Multivariate polynomial sieve.** Let $F(\mathbf{x}) : \mathbb{Z}^k \to \mathbb{Z}$ be a multivariate polynomial of $\mathbf{x} = (x_1, \dots, x_k)$, take any finite $\mathcal{A} \in \mathbb{Z}^k$, and $\mathcal{A}_p = \{\mathbf{x} \in \mathcal{A} : p|F(\mathbf{x})\}$. Then $S(\mathcal{A}, z)$ counts $\mathbf{x} \in \mathcal{A}$ for which $F(\mathbf{x})$ has no prime factor $p \leqslant z$.
- **The square-free sieve.** Let $\mathcal{A}$ be a finite set of integers and for each prime p let $\mathcal{A}_p = \{n \in \mathcal{A} : p^2|n\}$. Then $S(\mathcal{A}, z)$, with an appropriately large z, will count the elements of $\mathcal{A}$ which are squarefree. A famous application is for squarefree values of a polynomial, e.g. $\mathcal{A} = \{f(n) : 1 \leqslant n \leqslant x\}$, where f is an irreducible polynomial.

  One can similarly set up a k-free sieve problem.
- **Elliptic curve sieve.** Fix and elliptic curve E over $\mathbb{Q}$ Let $\mathcal{A}$ be the set of primes $q \leqslant x$. Let $\mathcal{A}_p = \{q \in \mathcal{A} : p|\#E/\mathbb{F}_q\}$, where $E/\mathbb{F}_q$ is the reduction of E modulo q. It is known that

$$\#E/\mathbb{F}_q \leqslant q + 1 + 2\sqrt{q} \leqslant 3q,$$

  and thus $S(\mathcal{A}, 2\sqrt{x})$ counts those q for which $\#E/\mathbb{F}_q$ is prime.

If d is a square-free integer, define

$$\mathcal{A}_d = \bigcap_{p|d} \mathcal{A}_p, \qquad A_d = M \cdot \mathbb{P}(\mathcal{A}_d). \tag{1.1}$$

In particular, $\mathcal{A}_1 = \mathcal{A}$ and $A_1 = M$. In the case where $\mathcal{A}$ is a finite set of integers with uniform measure, $M = |\mathcal{A}|$, and $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$, A_d counts the number of $n \in \mathcal{A}$ divisible by d.

In this notation, inclusion-exclusion gives

$$S(\mathcal{A}, z) = \sum_{d \in \mathscr{P}(z)} \mu(d) A_d. \tag{1.2}$$

1.3. **Legendre's sieve for primes (Legendre, 1808).** Let $\mathcal{A} = [1, x] \cap \mathbb{Z}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. Then $S(\mathcal{A}, z)$ counts the positive integers $n \leqslant x$ with no prime factor $\leqslant z$; in particular, this includes all of the primes between z and x. Also, for each aquarefree d, $\mathcal{A}_d = \{n \in \mathcal{A} : d|n\}$ and thus

$$A_d = |\mathcal{A}_d| = \lfloor x/d \rfloor = x/d + O(1).$$

By (1.2),

$$\begin{aligned}\pi(x) - \pi(z) \leqslant S(\mathcal{A}, z) &= \sum_{d \in \mathscr{P}(z)} \mu(d) \left(\frac{x}{d} + O(1) \right) \\ &= x \sum_{d \in \mathscr{P}(z)} \frac{\mu(d)}{d} + O\big(2^{\pi(z)}\big) \\ &= x \prod_{p \leqslant z} \left(1 - \frac{1}{p} \right) + O\big(2^{\pi(z)}\big).\end{aligned}$$

Taking $z = \log x$, and using the crude bound $\pi(z) \leqslant z$ together with Mertens' bound, we conclude that

$$\pi(x) \leqslant z + \frac{x}{\log z} \big(e^{-\gamma} + o(1) \big) + O(2^{\log x}) = O\left(\frac{x}{\log_2 x} \right).$$

Remark: We can do better by observing that in fact $A_d = 0$ for $d > x$ (cf. Hooley [**94**]), and thus restricting the sums to such d.

1.4. **The role of independence.** If the events $\mathcal{A}_p$ are independent, then the sieving problem is trivial, for then

$$S(\mathcal{A}, z) = M \prod_{p \leqslant z} (1 - \mathbb{P}\mathcal{A}_p). \tag{1.3}$$

In practice, the events are not independent, but are close to being so, especially for small primes. For example, in Legendre's sieve for primes, $A_d = \lfloor x/d \rfloor$ is very close to x/d. Even with this strong approximation, however, the accumulation of error terms (coming from k-correlations) becomes unwieldy when z is much larger than $\log x$.

1.5. **Main goals.** We will see that the right side of (1.3) is a good approximation to $S(\mathcal{A}, z)$ under very general conditions. We will accomplish this by a judicious pruning of the summands in (1.2).

To set things up, we adopt some additional notation. Take X to be an approximation of M (this can be anything, but in practice it is very close to M). We assume that there is a function g satisfying

$$0 \leqslant g(p) < 1 \qquad \text{(all primes } p\text{)}, \tag{g}$$

which, when extended to a multiplicative function by $g(d) = \prod_{p|d} g(p)$, gives $A_d \approx Xg(d)$ for squarefree d; that is, we require that the "remainders"

$$r_d := A_d - Xg(d) \tag{r}$$

to be "small on average". In the case where the events $\mathcal{A}_p$ are independent, taking $g(p) = \mathbb{P}\mathcal{A}_p$ yields $r_d = 0$ for all d and we recover (1.3). In practice, however, we will not take $g(p) = \mathbb{P}\mathcal{A}_p$ but

something very close which is convenient for calculations. We also adopt the short-hand

$$V(z) = \prod_{p \leqslant z} (1 - g(p)). \tag{V}$$

In this notation, the right side of (1.3) is about $XV(z)$.

Broadly speaking, if $pg(p)$ is bounded on average and r_d is $O(1)$ on average over "small" d, we will be able to prove the following:

$$\begin{aligned} &\text{(asymptotic)} && S(\mathcal{A}, z) \sim XV(z) \quad (z = X^{o(1)}); \\ &\text{(upper bound)} && S(\mathcal{A}, z) \ll XV(z) \quad (z \leqslant X); \\ &\text{(lower bound)} && S(\mathcal{A}, z) \gg XV(z) \quad (z \leqslant X^c), \end{aligned} \tag{1.4}$$

where the constant $c > 0$ depends on the nature of the sieve problem $\mathcal{A}$.

In plain language, we can prove the expected asymptotic formula for small z, an upper bound of the expected order for all z, and a lower bound of the expected order if z is at most a small power of X. The upper bound in (1.4) is amazing in its generality, and it has enormous utility as an auxilliary counting device in many problems.

1.6. **Sifting density (dimension), and level of distribution.** For most sieving problems, we have $g(p) \approx \kappa/p$ on average over p for some fixed κ (we'll make this precise below). In this case κ is referred to as the *sifting density* or *sifting dimension*. In the literature, the *linear sieve* refers to dimension 1. Various sieving procedures have been optimized for sieves of a particular density, e.g. the Rosser-Iwaniec theory of the linear sieve, and Iwaniec's theory of the half-dimensional sieve.

Roughly speaking, the *level of distribution* of a sieve problem $\mathcal{A}$ is the largest D for which

$$\sum_{d \leqslant D, d \in \mathscr{P}(z)} |r_d| \leqslant \varepsilon XV(z),$$

where $\varepsilon > 0$, X and z are given, and ε is small. The larger the level of distribution, the better quality of the bounds we can prove in (1.4).

1.7. **More examples.** There are limitations of the sieve, which are illustrated by the following examples.

1.7.1. *Eratosthenes sieve.* The asymptotic in (1.4) cannot be expected to hold for z being a fixed power of X. Take $\mathcal{A} = [1, x] \cap \mathbb{Z}$, $X = x$, and set $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. As before,

$$A_d = \#\{n \leqslant x : d|n\} = \lfloor x/d \rfloor = x/d + O(1)$$

Thus, taking $g(d) = 1/d$, we see that the error terms $r_d = O(1)$. Also,

$$S(\mathcal{A}, \sqrt{x}) = \pi(x) - \pi(\sqrt{x}) + 1 \sim \frac{x}{\log x} \qquad (x \to \infty)$$

by the Prime Number Theorem. However, Mertens' theorem gives

$$XV(\sqrt{x}) = x \prod_{p \leqslant \sqrt{x}} \left(1 - \frac{1}{p}\right) \sim \frac{2e^{-\gamma} x}{\log x},$$

with $2e^{-\gamma} = 1.122\ldots$. The discrepancy between $S(\mathcal{A}, \sqrt{x})$ and $XV(\sqrt{x})$ stems from the large amount of dependence among the events $\mathcal{A}_p$ for large primes p; e.g. $\mathcal{A}_p \cap \mathcal{A}_{p'} \cap \mathcal{A}_{p''} = \emptyset$ if $p > p' > p'' > x^{1/3}$. In fact, for fixed $c > 0$, one has $S(\mathcal{A}, x^c) \sim w(c)XV(z)$, where $w(\cdot)$ is the

Buchstab function, which satisfies $w(c) \neq 1$ for almost all $c \in (0,1]$. Thus, in general the condition $z = X^{o(1)}$ is necessary in order to conclude the asymptotic in (1.4).

1.7.2. *Twin primes.* . Take $\mathcal{A} = [1,x] \cap \mathbb{Z}$, $X = x$, $\mathcal{A}_p = \{n \in \mathcal{A} : p | n(n+2)\}$. Here $S(\mathcal{A}, \sqrt{x+2})$ counts the number of twin prime pairs between $\sqrt{x+2}$ and x. Hardy and Littlewood [**78**, Conjecture B] conjectured that the count of such pairs is $\sim Cx/\log^2 x$ where

$$C = 2\prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right) \approx 1.32.$$

For a heuristic explanation of this formula, see Section 1.8 below. By breaking up $[1,x]$ into subintervals of length d, we easily derive

$$A_d = \#\{n \leqslant x : d | n(n+2)\} = x\frac{\rho(d)}{d} + O(\rho(d)),$$

where

$$\rho(d) = \#\{0 \leqslant k \leqslant d-1 : k(k+2) \equiv 0 \pmod d\}.$$

By the Chinese remainder theorem, ρ is multiplicative, $\rho(2) = 1$ and $\rho(p) = 2$ for $p > 2$. Thus, this is a sieve problem of dimension 2 (or sifting density 2). Putting $g(d) = \rho(d)/d$ and applying Mertens, we get that

$$XV(\sqrt{x+2}) = \frac{x}{2}\prod_{3\leqslant p\leqslant z}\left(1 - \frac{2}{p}\right) \sim \frac{(4e^{-2\gamma})Cx}{\log^2 x} \qquad (x \to \infty). \tag{1.5}$$

As in subsection 1.7.1 above, $XV(z)$ differs by a constant multiplicative factor from what is expected to be true.

Sieve methods deliver an upper bound of the expected order

$$\#\{n \leqslant x : n \text{ and } n+2 \text{ are both prime}\} \ll xV(\sqrt{x}) \asymp \frac{x}{\log^2 x}.$$

However, sieve methods only deliver a lower bound for somewhat smaller z. For example, we will show in section 3 that

$$S(\mathcal{A}, x^{1/8}) \gg XV(x^{1/8}) \asymp \frac{x}{\log^2 x}.$$

From the last estimate, we conclude that there are $\gg x/\log^2 x$ values of $k \leqslant x$ for which each of k and $k+2$ has at most 7 prime factors. This is a typical conclusion from a lower bound sieve. A better conclusion is possible using the "linear variant" of the twin prime sieving problem, where $\mathcal{A} = \{p+2 : p \leqslant x\}$ a set of shifted primes, $\mathcal{A}_p = \{n \in \mathcal{A} : p | n\}$. In this set-up,

$$A_d = \pi(x; d, -2) \sim \frac{\mathrm{li}(x)}{\phi(d)} \qquad (2 \nmid d)$$

by the prime number theorem in arithmetic progressions. For the application to the sieve, we need a *uniform* estimate on $\pi(x; d, -2)$, at least on average over d. Here we take $X = \mathrm{li}(x)$, ad $g(p) = \frac{1}{p-1}$ for odd p. Since $g(p) \approx 1/p$, this is a sieve problem of dimension 1. We will show, in section 3, that

$$S(\mathcal{A}, x^{1/7}) \gg X\,V(x^{1/7}) \gg \frac{x}{\log^2 x},$$

and hence there are $\gg x/\log^2 x$ primes $p \leqslant x$ such that $p+2$ has at most 6 prime factors.

1.7.3. *Prime k-tuples.* Let $a_1, \cdots, a_k \in \mathbb{N}$ and $b_1, \ldots, b_k \in \mathbb{Z}$. Put $\mathcal{A} = [1, x] \cap \mathbb{N}$, and

$$\mathcal{A}_p = \{n \in \mathcal{A} : p | (a_1 n + b_1) \cdots (a_k n + b_k)\}.$$

An easy counting yields

$$A_d = x\frac{\rho(d)}{d} + O(\rho(d)),$$

where

$$\rho(d) = \#\{0 \leqslant n < d : (a_1 n + b_1) \cdots (a_k n + b_k) \equiv 0 \pmod{d}\}.$$

By the Chinese Remainder Theorem, $\rho(d)$ is multiplicative. We say that the collection of linear forms $a_1 n + b_1, \ldots, a_k n + b_k$ is *admissible* if $\rho(p) < p$ for all primes p. In the contrary case that $\rho(p) = p$ for some p, for every n one of the forms $a_j n + b_j$ is divisible by p and hence there are only finitely many n making all of the $a_j n + b_j$ simultaneously prime. Some examples:

Admissible	Non-admissible
$(n, n + 2k)$; $k \in \mathbb{N}$	$(n, n + 1)$
$(n, 2n + 1)$	$(n, 3n + 1)$
$(n, n + 2, n + 6)$	$(n, n + 2, n + 4)$

We set $g(p) = \rho(p)/p$. Note that if

$$p \nmid \prod_{i=1}^{k} a_i \prod_{1 \leqslant i < j \leqslant k} (a_i b_j - a_j b_i),$$

then $\rho(p) = k$ (see Theorem 2.5 below). In particular, $\rho(p) = k$ for all sufficiently large p, and hence this is a sieve problem of density k (or dimension k). At present, sieve methods can deliver bounds of the form

$$\#\{n \leqslant x : a_j n + b_j \text{ are prime for all } j\} \leqslant C\frac{x}{(\log x)^k},$$

where C is an explicit function of $a_1, b_1, \ldots, a_k, b_k$. See Theorem 2.5 below for a precise statement. If $k \geqslant 2$ we still do not know how to show that the left side goes to ∞ as $x \to \infty$.

1.7.4. *Selberg's examples.* [**135**]. The values of c for which sieve method deliver lower bounds (1.4) are invariable smaller than we would like. There is a fundamental barrier at work which explains this, known as the "parity barrier". Roughly speaking, the small sieve works with inputs X, the function g and estimates for the remainders r_d. However, even if the level of distribution is very large, say $x^{1-o(1)}$, the sieve fundamentally cannot distinguish between numbers with an odd number of prime factors from those with an even number of prime factors. Consider two sequences defined as follows. Recall the Liouville function $\lambda(n) = (-1)^{\Omega(n)}$ (the completely multiplicative function which is -1 at all primes). Define

$$\mathcal{A}^+ = \{1 \leqslant n \leqslant x : \lambda(n) = 1\}, \qquad \mathcal{A}^- = \{1 \leqslant n \leqslant x : \lambda(n) = -1\},$$

respectively, and $\mathcal{A}_p^{\pm} = \{n \in \mathcal{A}^{\pm} : p | n\}$ for each p.

The prime number theorem (with classical error term) implies

$$\sum_{n \leqslant x} \lambda(n) = O(xe^{-c'\sqrt{\log x}}), \text{ some } c' > 0.$$

Therefore,

$$\#\{n\in\mathcal{A}^{\pm}: d|n\} = \sum_{m\leqslant x/d}\frac{1\pm\lambda(dm)}{2} = \frac{1}{2}\Big\lfloor\frac{x}{d}\Big\rfloor \pm \frac{\lambda(d)}{2}\sum_{m\leqslant x/d}\lambda(m)$$
$$= \frac{x}{2d} + O\left(\frac{x}{d}e^{-c\sqrt{\log(x/d)}}\right).$$

In particular, taking $d=1$, we wee that $|\mathcal{A}^{\pm}|\sim\frac{x}{2}$. Take $X=\frac{x}{2}$ and $g(d)=\frac{1}{d}$ for all d. For both sequences, the level of distribution is $x^{1-o(1)}$ (pretty much the best range that one can hope for), and for both sieve problems we have

$$V(\sqrt{x})\sim\frac{2e^{-\gamma}}{\log x}.$$

However, numbers $n\leqslant x$ that have no prime factor $\leqslant\sqrt{x}$ are prime or 1, thus

$$S(\mathcal{A}^+,\sqrt{x})=1, \qquad \text{while} \qquad S(\mathcal{A}^-,\sqrt{x})\sim\frac{x}{\log x}=\frac{2X}{\log x}.$$

Thus, although the sieve inputs (quantity X and bounds on A_d) are identical in both problems, the truth is very different. In order to distinguish primes from integers with 2 prime factors, say, further inputs into the sieve are required. Sieve procedures which include as hypotheses bilinear sum bounds for $\mathcal{A}$ have proven to be very successful in detecting primes, e.g. Friedlander-Iwaniec [**60**] and Harman [**82**].

1.8. **The prime k–tuples conjecture.** Much of sieve theory has been driven by attempts to prove special cases of the general *Prime k-tuples Conjecture*. The setup is a finite collection $f_1,\dots,f_k$ of nonconstant irreducible (over $\mathbb{Q}$) polynomials in m variables, with integer coefficients. For each prime p, let

$$\rho(p)=\#\{\mathbf{x}=(x_1,\dots,x_m) \mod p : f_1(\mathbf{x})\cdots f_k(\mathbf{x})\equiv 0 \pmod p\}.$$

Definition 1. *The set $(f_1,\dots,f_k)$ is* admissible *if the f_i are distinct and $\rho(p)<p^m$ for all p.*

Conjecture 1.1 (General Prime k-tuples Conjecture)**.** *Let $f_1,\dots,f_k$ be an admissible collection of polynomials from $\mathbb{Z}^m$ to $\mathbb{Z}$. Then*

$$\#\{0\leqslant x_i<x\ (1\leqslant i\leqslant m): f_1(\mathbf{x}),\dots,f_k(\mathbf{x}) \text{ are all prime }\}\sim\frac{\mathfrak{S}}{\prod_{i=1}^k deg(f_i)}\frac{x^m}{(\log x)^k}, \tag{1.6}$$

where $deg(f_i)$ is the total degree of f_i, and

$$\mathfrak{S}=\mathfrak{S}(f_1,\dots,f_k)=\prod_p\left(1-\frac{\rho(p)}{p^m}\right)\left(1-\frac{1}{p}\right)^{-k}$$

is the so-called singular series *associated with $f_1,\dots,f_k$.*

One can show that $\rho(p)/p^{m-1}$ is k on average and this implies that the infinite product converges (details below in the case of univariate polynomials).

This conjecture subsumes a large number of conjectures that have been made over time, in various degrees of generality. We mention here the conjectures of Bunyakowsky [**16**], Dickson [**28**], Hardy-Littlewood [**78**], Schinzel [**131**], and Bateman-Horn [**11**].

Some special cases.

- **Primes in an arithmetic progression.** $m = 1$, $k = 1$, $f_1(n) = qn + a$, where $(a, q) = 1$. Dirichlet proved [**29**] in 1837 that there are infinitely many n with $f_1(n)$ prime; the Prime Number Theorem for arithmetic progressions (de la Vallée Poussin, 1896) gives the full asymptotic (1.6).
- **Twin primes.** $m = 1$, $k = 2$, $f_1(n) = n$, $f_2(n) = n + 2$. Here $\mathfrak{S} = C := 2\prod_{p>2}(1 - \frac{1}{(p-1)^2}) = 1.32\ldots$ is the "twin prime constant".
- **"Sexy" primes.** $m = 1$, $k = 2$, $f_1(n) = n$, $f_2(n) = n + 6$. Here $\mathfrak{S} = 2C$ since $\rho(2) = \rho(3) = 1$ and $\rho(p) = 2$ for $p > 3$. That is, conjecturally there are about twice as many "sexy primes" as twin primes below x.
- **Sophie Germain primes.** $m = 1$, $k = 2$, $f_1(n) = n$, $f_2(n) = 2n + 1$.
- **Primes of form $n^2 + 1$.** $m = 1$, $k = 1$, $f_1(n) = n^2 + 1$.
- **k-term arithmetic progressions of primes.** $m = 2$, forms $n_1, n_1 + n_2, n_1 + 2n_2, \ldots, n_1 + (k-1)n_2$. When $k = 3$, the asymptotic in Conjecture 1.1 was essentially proved by Vinogradov in 1937. Balog [**5**] extended this to include many other collections of forms with $m \geqslant 2$. Green and Tao [**71**] showed that there are infinitely many k-term arithmetic progressions of primes for any $k \geqslant 4$, and this was extended by Green and Tao and by Green, Tao and Ziegler [**72, 73, 74, 75**] in 2010–12 to prove the full asymptotic in (1.6). Their theory extends to many other collections of linear forms when $m \geqslant 2$.
- **Primes of the form $x^2 + y^2$.** $m = 2$, $k = 1$, $f_1(x, y) = x^2 + y^2$. Fermat showed that every prime $p \equiv 1 \pmod 4$ is the sum of two squares.
- **Primes of the form $x^2 + y^4$.** $m = 2$, $k = 1$, $f_1(x, y) = x^2 + y^4$. The infinitude of such primes, and in fact the full asymptotic (1.6), is a celebrated theorem of Friedlander and Iwaniec [**59**] in 1998.

The only case of (1.6) which is known when $m = 1$ is the case when $k = 1$ and f_1 is linear (the prime number theorem for arithmetic progressions).

There is a relatively easy heuristic for (1.6). According to the Prime Number Theorem, a randomly chosen integer near x has a likelihood of about $\frac{1}{\log x}$ of being prime. Assuming that the $f_i(\mathbf{x})$ behave randomly, the likelihood of $f_i(\mathbf{x})$ being prime should be about $\frac{1}{\log f_i(\mathbf{x})} \sim \frac{1}{\deg(f_i)\log x}$ if each $x_i \in [1, x]$. This leads to the prediction that

$$\#\{0 \leqslant x_i < x \ (1 \leqslant i \leqslant m) : f_1(\mathbf{x}), \ldots, f_k(\mathbf{x}) \text{ are all prime } \} \sim \frac{1}{\prod_{i=1}^k \deg(f_i)} \frac{x^m}{(\log x)^k}.$$

This matches (1.6) except for the singular series factor. Implicit in the "randomness" hypothesis is the assumption that for any prime p, the likelihood that each of $f_i(\mathbf{x})$ is coprime to p is about $(1 - 1/p)^k$. This is not correct, however, and if $\mathbf{x}$ is chosen randomly modulo p, then the likelihood that each of the $f_i(\mathbf{x})$ is coprime to p is exactly $1 - \rho(p)/p^m$. Thus, in our heuristic we should insert a correction factor

$$\left(1 - \frac{\rho(p)}{p^m}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

Doing this for all p produces a correction factor equal to $\mathfrak{S}(f_1, \ldots, f_k)$, and leads to the more precise prediction (1.6).

Recently, Banks, Ford and Tao [**9**] showed that when the polynomials f_i are all linear, then the asymptotic formula (1.6) can be heuristically derived by sieving up to x^{1/e^γ}, suggested by Polya in

the case of the prime number theorem. That is, with $X = x^m$ and $z = x^{1/\mathrm{e}^\gamma}$,

$$XV(z) \sim \frac{\mathfrak{S} x^m}{(\log x)^k}.$$

This is easy to show: indeed,

$$\begin{aligned} V(z) &= \prod_{p\leqslant z} \left(1 - \frac{\rho(p)}{p^m}\right) \\ &= \prod_{p\leqslant z} \left(1 - \frac{\rho(p)}{p^m}\right) (1-1/p)^{-k}(1-1/p)^k \\ &\sim \mathfrak{S} \left(\frac{e^{-\gamma}}{\log z}\right)^k = \frac{\mathfrak{S}}{(\log x)^k} \quad (x\to\infty). \end{aligned}$$

1.9. **The small sieve.** Viggo Brun's fundamental idea was to replace the huge sum on the right side of (1.2) with a sum over a much smaller range of d, at the expense of replacing the equality in (1.2) with an inequality. In general, a *sieve* (or *small sieve*) is a sequence $\boldsymbol{\lambda} = (\lambda_d)$ supported on squarefree integers d which replaces $\mu(d)$ in (1.2). We will suppose that either $\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ = (\lambda_d^+)$ satisfies

$$(\lambda^+) \qquad \lambda_1^+ = 1, \quad \sum_{d|m} \lambda_d^+ \geqslant 0 \qquad (m>1)$$

or that $\boldsymbol{\lambda} = \boldsymbol{\lambda}^- = (\lambda_d^-)$ satisfies

$$(\lambda^-) \qquad \lambda_1^- = 1, \quad \sum_{d|m} \lambda_d^- \leqslant 0 \qquad (m>1).$$

We call $\boldsymbol{\lambda}^+$ an *upper bound sieve* and $\boldsymbol{\lambda}^-$ a *lower bound sieve*, which comes from the following easy result:

Lemma 1.2. *For $z \geqslant 2$ let $\mathscr{P}(z)$ denote the set of squarefree positive integers, divisible only by primes $p \leqslant z$. Assume (λ^+) and (λ^-). Then, for any sieve problem and and $z \geqslant 2$,*

$$\sum_{d\in\mathscr{P}(z)} \lambda_d^- A_d \leqslant S(\mathcal{A}, z) \leqslant \sum_{d\in\mathscr{P}(z)} \lambda_d^+ A_d. \tag{1.7}$$

Further, for any multiplicative function g such that $0 \leqslant g(p) < 1$ for each $p \leqslant z$, we have

$$\sum_{d\in\mathscr{P}(z)} \lambda_d^- g(d) \leqslant \prod_{p\leqslant z}(1-g(p)) \leqslant \sum_{d\in\mathscr{P}(z)} \lambda_d^+ g(d). \tag{1.8}$$

Proof. For a random $\omega \in \mathcal{A}$, let

$$m_\omega = \prod_{\substack{p\leqslant z\\ \omega\in\mathcal{A}_p}} p.$$

Then, by (λ^+),

$$\begin{aligned} S(\mathcal{A}, z) &= M\mathbb{P}(m_\omega = 1) = M\cdot \mathbb{E}\mathbb{1}_{m_\omega=1} \\ &\leqslant M\cdot \mathbb{E}\sum_{d|m_\omega} \lambda_d^+ = \sum_{d\in\mathscr{P}(z)} \lambda_d^+ A_d, \end{aligned}$$

and similarly by (λ^-),

$$S(\mathcal{A},z) \geqslant M\cdot \mathbb{E}\sum_{d|m_\omega}\lambda_d^- = \sum_{d\in\mathscr{P}(z)}\lambda_d^- A_d.$$

The claim (1.8) is a special case, where $M=$, our probability space has independent events $\mathcal{A}_p$ such that $\mathbb{P}\mathcal{A}_p = g(p)$, so that $S(\mathcal{A},z)=\prod_{p\leqslant z}(1-g(p))$ and $A_d=g(d)$. Such a probability space is easy to construct explicitly: take $m=\pi(z)$, p_k the k-th prime, $\mathcal{A}=\{0,1\}^m$, the probability defined by

$$\mathbb{P}((v_1,\dots,v_m)) = \prod_{v_k=0} g(p_k)\prod_{v_k=1}(1-g(p_k)),$$

and $\mathcal{A}_{p_k}=\{(v_1,\dots,v_m)\in\mathcal{A} : v_k=0\}$. □

The major goal of sieve methods is to construct good sieve parameters $\boldsymbol{\lambda}^\pm$, with *small support* inside $\mathscr{P}(z)$ (usually this means the support has $z^{O(1)}$ elements in $\mathscr{P}(z)$) and with the sums in (1.7) mimicking $S(\mathcal{A},z)$ as closely as possible. With this notation, (g) and (r), we have

$$\sum_{d\in\mathscr{P}(z)}\lambda_d A_d = X\sum_{d\in\mathscr{P}(z)} g(d)\lambda_d + \sum_{d\in\mathscr{P}(z)}\lambda_d r_d. \tag{1.9}$$

As λ_d is a replacement for $\mu(d)$, it is reasonable to suppose (if we have constructed our sieve well) that

$$\sum_{d\in\mathscr{P}(z)} g(d)\lambda_d \approx \sum_{d\in\mathscr{P}(z)} g(d)\mu(d) = \prod_{p\leqslant z}(1-g(p)) = V(z).$$

1.10. **Legendre's sieve, general version.**

Theorem 1.3. *Consider a sieve problem, assume (g), (r) and adopt notation (V). Then*

$$S(\mathcal{A},z) = XV(z) + O\Big(\sum_{d\in\mathscr{P}(z)}|r_d|\Big). \tag{1.10}$$

Proof. Trivially, the weights $\lambda_d^\pm=\mu(d)$ satisfy (λ^+) and (λ^-) with equality. By (1.7),

$$S(\mathcal{A},z) = X\sum_{d\in\mathscr{P}(z)}\mu(d)g(d) + \sum_{d\in\mathscr{P}(z)}\mu(d)r_d = XV(z)+O\Big(\sum_{d\in\mathscr{P}(z)}|r_d|\Big).$$ □

Although this sieve suffers from the large number, $2^{\pi(z)}$, of remainder summands, it is useful in situations where the "densities" $g(p)$ are rather small on average, and so the product on the right hand side of (1.10) captures the true behavior of set of interest for relatively small z. A prominent example of the use of Legendre's sieve was given by Hooley [**90**], who deduced Artin's primitive root conjecture from the Generalized Riemann Hypothesis for Dedekind zeta functions of certain number fields. Details will be given later in Section 2.2.6.

1.11. **Brun's pure sieve.** Brun's sieve is based on a simple truncated version of inclusion-exclusion, due to Brun (1915).

Lemma 1.4 (Inclusion-exclusion)**.** *Let u be a non-negative integer. Then, for any $k\in\mathbb{N}$,*

$$\mathbb{1}_{u=0} = \sum_{r=0}^{\infty}(-1)^r\binom{u}{r} = \sum_{r=0}^{k}(-1)^r\binom{u}{r} + (-1)^{k+1}\binom{u-1}{k},$$

where we set $\binom{-1}{j} = 0$ *for all* j.

Proof. We may assume that $u \geqslant 1$, as the statement is trivial when $u = 0$. The first equality is trivial from the binomial theorem. For the second, when $u \geqslant 1$ we have

$$\sum_{r=k+1}^{\infty} (-1)^r \binom{u}{r} = \sum_{r=k+1}^{\infty} (-1)^r \left[\binom{u-1}{r-1} + \binom{u-1}{r} \right] = (-1)^{k+1} \binom{u-1}{k}. \qquad \square$$

Theorem 1.5 (Brun's pure sieve)**.** *Let* k *be a nonnegative integer and define*

$$\lambda_d = \begin{cases} \mu(d) & \text{if } \omega(d) \leqslant k \\ 0 & \text{otherwise.} \end{cases}$$

Then $\boldsymbol{\lambda}$ *satisfies* (λ^+) *if* k *is even and* (λ^-) *if* k *is odd. Thus, if* k_e *is even and* k_o *is odd, then for any sieve problem* $\mathcal{A}$ *and any* $z \geqslant 2$ *we have*

$$\sum_{\substack{d \in \mathscr{P}(z) \\ \omega(d) \leqslant k_o}} \mu(d) A_d \leqslant S(\mathcal{A}, z) \leqslant \sum_{\substack{d \in \mathscr{P}(z) \\ \omega(d) \leqslant k_e}} \mu(d) A_d. \tag{1.11}$$

In particular, taking k_e, k_o *to be consecutive integers we see that for any non-negative integer* k*,*

$$\left| S(\mathcal{A}, z) - \sum_{\substack{d \in \mathscr{P}(z) \\ \omega(d) \leqslant k}} \mu(d) A_d \right| \leqslant \sum_{\substack{d \in \mathscr{P}(z) \\ \omega(d) = k+1}} \mu(d) A_d.$$

Proof. Let $k \geqslant 0$ and let m be a positive, squarefree integer. Then, by Lemma 1.4,

$$\begin{aligned} \mathbb{1}_{m=1} = \mathbb{1}_{\omega(m)=0} &= \sum_{r=0}^{k} (-1)^r \binom{\omega(m)}{r} + (-1)^{k+1} Y, \qquad Y = \binom{\omega(m)-1}{k} \geqslant 0, \\ &= \sum_{r=0}^{k} (-1)^r \sum_{\substack{d \mid m \\ \omega(d) = r}} 1 + (-1)^{k+1} Y \\ &= \sum_{d \mid m \omega(d) \leqslant k} \mu(d) + (-1)^{k+1} Y \\ &= \sum_{d \mid m} \lambda_d + (-1)^{k+1} Y, \end{aligned}$$

from which follows (λ^+) if k is even and (λ^-) if k is odd. The final claim (1.11) follows from (1.7) in Lemma 1.2. $\square$

The inequalities in Theorem 1.5 were later rediscovered by Bonferroni, and are often referred to as the "Bonferroni inequalities" in probability theory.

Lemma 1.6. *Let* $z \geqslant 2$ *and* $0 \leqslant g(p) \leqslant 1$ *for each prime* $p \leqslant z$*. Let* k *be a non-negative integer. Extend* f *multiplicatively by defining* $g(d) = \prod_{p \mid d} g(p)$ *for any squarefree* d *with* $d \in \mathscr{P}(z)$*. Then*

$$\sum_{\substack{d \in \mathscr{P}(z) \\ \omega(d) \leqslant k}} \mu(d) g(d) = \prod_{p \leqslant z} (1 - g(p)) + (-1)^k W,$$

where

$$0 \leqslant W \leqslant \sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)=k+1}} g(d) \leqslant \frac{1}{(k+1)!}\Big(\sum_{p\leqslant z} g(p)\Big)^{k+1}.$$

Proof. Again, consider the sieve problem where $\mathbb{P}(\mathcal{A}_p) = g(p)$ for each $p \leqslant z$, and the events $\mathcal{A}_p$ are independent. By Theorem 1.5 and (1.8),

$$\prod_{p\leqslant z}(1-g(p)) = \sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)\leqslant k}} \mu(d)g(d) + (-1)^{k+1}W, \qquad 0 \leqslant W \leqslant \sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)=k+1}} g(d).$$

The final inequality for W comes from expanding the $(k+1)$-fold sum (this is an old Erdős trick). □

Theorem 1.7. *Assume* (g), (r) *and* $V(z) > 0$, *where* $V(z)$ *is given by* (V). *Then*

$$S(\mathcal{A}, z) = XV(z) + O(XV(z)^{3/2}) + O\Big(\sum_{\substack{d\leqslant z^{4\log(1/V(z))+1}\\ d\in\mathscr{P}(z)}} |r_d|\Big).$$

Proof. Apply Theorem 1.5 with $k = \Big\lfloor 4\log\frac{1}{V(z)}\Big\rfloor$, followed by an application of Lemma 1.6. This gives

$$\begin{aligned} S(\mathcal{A}, z) &= \sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)\leqslant k}} \mu(d)A_d + O\Big(\sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)=k+1}} A_d\Big), \quad A_d = Xg(d) + r_d \\ &= X\sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)\leqslant k}} \mu(d)g(d) + O\Big(X\sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)=k+1}} g(d)\Big) + O\Big(\sum_{\substack{d\in\mathscr{P}(z)\\ \omega(d)\leqslant k+1}} |r_d|\Big) \\ &= XV(z) + O\Big(X\frac{1}{(k+1)!}\Big(\sum_{p\leqslant z} g(p)\Big)^{k+1}\Big) + O\Big(\sum_{\substack{d\in\mathscr{P}(z)\\ d\leqslant z^{k+1}}} |r_d|\Big). \end{aligned}$$

Lastly, by our definition of k,

$$\begin{aligned} \frac{1}{(k+1)!}\Big(\sum_{p\leqslant z} g(p)\Big)^{k+1} &\leqslant \frac{1}{(k+1)!}\Big(\sum_{p\leqslant z} -\log(1-g(p))\Big)^{k+1} \\ &= \frac{1}{(k+1)!}\left(\log\frac{1}{V(z)}\right)^{k+1} \\ &\leqslant \left(\frac{e\log\frac{1}{V(z)}}{k+1}\right)^{k+1} \\ &\leqslant (e/4)^{k+1} \ll V(z)^{4\log 4-4} \leqslant V(z)^{3/2}. \end{aligned} \tag{1.12}$$

This completes the proof. □

Remarks. We have imposed virtually no hypotheses on the sieve problem in this theorem. In applications, one has typically $V(z) \asymp (\log X)^{-\kappa}$ for some fixed κ, and thus we have applied (1.11) with $k \approx 4\kappa\log_2 X$. Typically in the sum in identity (1.2) and the sums in Brun's inequalities

(1.11), the summands corresponding to $d > X$ are negligible (or identically zero). Hardy and Ramanujan [**80**] in 1917 showed that most integers $d \leqslant X$ have about $\log_2 X$ prime factors, and thus it is natural to choose k somewhat larger than $\log_2 X$ in order to capture the bulk of the sum.

1.11.1. *Example: twin primes.* As before, let $\mathcal{A} = [1, x] \cap \mathbb{Z}$, $\mathcal{A}_p = \{k \in \mathcal{A} : p | k(k+2)\}$. Take $X = x$,

$$g(p) = \frac{\rho(p)}{p}, \quad \rho(2) = 1, \quad \rho(p) = 2 \ \ (p > 2),$$

and extend ρ to a multiplicative function on $\mathscr{P}(z)$. For squarefree d, A_d counts integers lying in $\rho(d)$ residue classes modulo d, hence

$$\lfloor x/d \rfloor \rho(d) \leqslant A_d \leqslant \lceil x/d \rceil \rho(d).$$

This implies that $A_d = xg(d) + r_d$ where

$$|r_d| \leqslant \rho(d) \leqslant \tau(d).$$

By Mertens' estimate (cf., (1.5)), $V(z) \sim c(\log z)^{-2}$ for some constant c. Take

$$z = x^{\frac{1}{16 \log_2 x}},$$

so that in the summation of the error terms r_d we have

$$z^{4 \log(1/V(z))+1} = z^{8 \log_2 z + O(1)} \ll x^{1/2+o(1)}.$$

Thus,

$$\sum_{\substack{d \in \mathscr{P}(z) \\ d \leqslant z^{4 \log(1/V(z))+1}}} |r_d| \leqslant \sum_{d \leqslant z^{4 \log(1/V(z))+1}} \tau(d) \ll x^{1/2+o(1)}.$$

By Theorem 1.5,

$$S(\mathcal{A}, z) \sim XV(z) \sim \frac{cx}{\log^2 z} \sim 256cx \left(\frac{\log_2 x}{\log x} \right)^2.$$

As $S(\mathcal{A}, z) \geqslant \#\{z < k \leqslant x : k, k+2 \text{ both prime}\}$, we see that

$$\#\{k \leqslant x : k, k+2 \text{ both prime}\} \ll x \left(\frac{\log_2 x}{\log x} \right)^2,$$

which misses the conjectured order by a factor $(\log_2 x)^2$. Applying partial summation gives an immediate corollary.

Corollary 1.8 (Brun [**19**], 1919)**.** *We have*

$$\sum_{p : p, p+2 \ prime} \frac{1}{p} < \infty.$$

Remarks. Applying Theorem 1.5 to the sieve of Eratosthenes yields $\pi(x) \ll x \frac{\log_2 x}{\log x}$, missing the true order of $\pi(x)$ by a factor $\log_2 x$.

Brun later [**20**] gave a much more complicated version of his sieve, where the simple truncation (1.11) is replaced by sieves where one considers the summation over integers d with restricted prime factors of various sizes. A greatly simplified version of this idea was found by Hooley, and which will be the subject of the following two sections.

2. The Brun-Hooley sieve: upper bounds

In this section and the next, we present a variation of Brun's pure sieve due to C. Hooley [**95**], which is very simple, from a combinatorial viewpoint, and yet powerful enough to deliver quickly each of the desired sieve bounds in (1.4). Some general estimates using this method were given by Ford and Halberstam [**55**].

2.1. **Upper bounds.** The fundamental idea of the upper bound sieve is to partition the primes blow z into sets $\mathcal{P}_1, \dots, \mathcal{P}_t$ and apply the Brun sieve bounds on each set $\mathcal{P}_i$ separately.

Lemma 2.1. *Partition the primes $\leqslant z$ as $\mathcal{P}_1 \cup \cdots \cup \mathcal{P}_t$. For each i, let $(\lambda_d^{(i)})$ be an upper bound sieve. For any $d \in \mathscr{P}(z)$, let*

$$\lambda_d^+ = \prod_{i=1}^t \lambda_{d_i}^{(i)},$$

where $d_1, \dots, d_t$ are defined uniquely by

$$d = d_1 \cdots d_t, \qquad (\forall i,\ p|d_i \Rightarrow p \in \mathcal{P}_i). \tag{2.1}$$

Then λ^+ is an upper bound sieve, i.e., satisfies (λ^+).

Proof. Clearly $\lambda^+ = 1$. Also, for any $m > 1$, m has a unique decomposition as $m = m_1 \cdots m_t$, where for each i, all of the prime factors of m_i are in $\mathcal{P}_i$. Then, by (λ^+),

$$\sum_{d|m} \lambda^+ = \prod_{i=1}^t \Big(\sum_{d_i|m_i} \lambda_{d_i}^{(i)} \Big) \geqslant 0,$$

as required. □

Taking each $\lambda_d^{(i)}$ to be an upper bound Brun sieve (from Theorem 1.5) with $k = k_i$ even, we arrive at

Lemma 2.2 (The Brun-Hooley upper bound sieve)**.** *Let $\mathcal{P}_1 \cup \cdot \cup \mathcal{P}_t$ be any partition of the primes $\leqslant z$, and let $k_1, \dots, k_t$ be arbitrary non-negative even integers. Then the sequence $\boldsymbol{\lambda}$ given by*

$$\lambda_d^+ = \begin{cases} \mu(d) & \text{if } \omega(d_j) \leqslant k_j \ \ (1 \leqslant j \leqslant t) \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

is an upper bound sieve satisfying (λ^+). Here d uniquel decomposes as in (2.1).

The number theoretic motivation for this comes from looking at the statistical distribution of prime factors of integers $\leqslant X$. Not only does a typical integer have about $\log\log X$ prime factors, but the prime factors themselves are typically uniformly distributed on a $\log\log$-scale; that is, there are about $\log\log t$ prime factors $\leqslant t$, uniformly for $t \leqslant X$. With Lemma 2.2, we can restrict the number of large prime factors of the summands d while still retaining almost all significant summands.

Suppose that we have a sieve problem, $z \geqslant 2$ and assume (g) and (r). Partition the primes $\leqslant z$ as $\mathcal{P}_1 \cup \cdots \mathcal{P}_t$. Define

$$P_i = \prod_{p \in \mathcal{P}_i} p \qquad (1 \leqslant i \leqslant t).$$

From Lemma 2.2 and Lemma 1.2 we immediately get

$$\begin{aligned} S(\mathcal{A}, z) &\leqslant \sum_{\substack{d_1 | P_1 \\ \omega(d_1) \leqslant k_1}} \cdots \sum_{\substack{d_t | P_t \\ \omega(d_t) \leqslant k_t}} \mu(d_1) \cdots \mu(d_t) A_{d_1 \cdots d_t} \\ &= X U_1 \cdots U_t + R, \end{aligned}$$

where

$$U_j = \sum_{\substack{d_j | P_j \\ \omega(d_j) \leqslant k_j}} \mu(d_j) g(d_j), \qquad R = \sum_{\substack{d_1 | P_1 \\ \omega(d_1) \leqslant k_1}} \cdots \sum_{\substack{d_t | P_t \\ \omega(d_t) \leqslant k_t}} \mu(d_1 \cdots d_t) r_{d_1 \cdots d_t}. \tag{2.3}$$

Define

$$V_j = \prod_{p \in \mathcal{P}_j} (1 - g(p)), \quad L_j = \log \frac{1}{V_j}. \tag{2.4}$$

Notice that $V_1 \cdots V_t = V(z)$. Lemma 1.6 plus $g(p) \leqslant -\log(1 - g(p))$ implies

$$V_j \leqslant U_j \leqslant V_j + \frac{L_j^{k_j+1}}{(k_j+1)!} = V_j \left(1 + e^{L_j} \frac{L_j^{k_j+1}}{(k_j+1)!} \right) \leqslant V_j \exp\left\{ e^{L_j} \frac{L_j^{k_j+1}}{(k_j+1)!} \right\}.$$

The last inequality appears wasteful, but in practice the quantity in the exponential is small. When $k_j = 0$ we can omit the factor e^{L_j} since

$$U_j = 1 = V_j V_j^{-1} = V_j e^{L_j}.$$

This leads to small improvements in numerical constants. Multiplying these inequalities together for all j, we have

$$V(z) \leqslant U_1 \cdots U_t \leqslant e^E V(z), \tag{2.5}$$

where

$$E := \sum_{j=1}^{t} e^{L_j \mathbb{1}_{k_j > 0}} \frac{L_j^{k_j+1}}{(k_j+1)!}. \tag{2.6}$$

We conclude from (1.7) and (1.9) the following.

Theorem 2.3. *Let $\mathcal{A}$ be a sieve problem and assume (g), (r) and (V). Partition the set of primes in $[1, z]$ into $\mathcal{P}_1 \cup \cdots \cup \mathcal{P}_t$, let $P_i = \prod_{p \in \mathcal{P}_i} p$ and let $k_1, \dots, k_t$ be non-negative even integers. Then*

$$S(\mathcal{A}, z) \leqslant X V(z) e^E + R',$$

where E is given by (2.6) *and*

$$R' := \sum_{\substack{\omega((d, P_j)) \leqslant k_j \ (1 \leqslant j \leqslant t) \\ d \in \mathscr{P}(z)}} |r_d|. \tag{2.7}$$

In particular, defining λ_d^+ as in Lemma 2.2, we have

$$\sum_d \lambda_d^+ g(d) \leqslant e^E V(z). \tag{2.8}$$

To make further progress and produce an upper bound of general utility, we make a very mild assumption on the function g, which is satisfied in a large number of cases. It is sometimes called the Iwaniec condition. We assume that for some $\kappa \geqslant 0$ and $B > 0$ that

$$\prod_{y \leqslant p \leqslant w} (1 - g(p))^{-1} \leqslant \left(\frac{\log w}{\log y}\right)^{\kappa} \exp\left(\frac{B}{\log y}\right) \quad (2 \leqslant y \leqslant w \leqslant z) \tag{Ω}$$

Roughly speaking, this states that $g(p) \lesssim \kappa/p$ on average over p. Although B, κ are not uniquely defined by (Ω), the smallest admissible value of κ (with B remaining bounded) is sometimes referred to as the "dimension" or "sifting density". It is easy to show (cf., Exercise 2.1 below), that condition (Ω) is implied by the condition

$$g(p) \leqslant \min(\kappa/p, 1 - \delta). \tag{Ω_0}$$

Although (Ω_0) is easy to verify, for many problems we have (Ω) holding with a smaller value of κ than the global bound (Ω_0), and this produces quantitatively better sieve bounds.

We can now obtain a general purpose upper bound of type in (1.4). In sums of remainders r_d we recall that r_d is supported only on squarefree d.

Theorem 2.4. *Let $\mathcal{A}$ be a sieve problem, let $z \geqslant 2$ and assume (g), (r) and (V). Let $\kappa_0 > 0$ and $B_0 \geqslant 0$. Assume (Ω) with $0 \leqslant \kappa \leqslant \kappa_0$ and $0 \leqslant B \leqslant B_0$. Then for any $z \geqslant 2$ we have*

$$S(\mathcal{A}, z) \leqslant O_{\kappa_0, B_0}(XV(z)) + \sum_{\substack{d \leqslant z \\ \mu^2(d)=1}} |r_d|,$$

the implied constant depending only on κ_0, B_0.

Proof. Before invoking (Ω), we will work out a general procedure for producing upper bounds from Theorem 2.3. We suppose that

$$\begin{aligned} z_{t+1} = 2 < z_t < z_{t-1} < \cdots < z_1 = z \\ \mathcal{P}_j = \{p \text{ prime} : z_{j+1} < p \leqslant z_j\} \qquad (1 \leqslant j \leqslant t-1), \\ \mathcal{P}_t = \{p \text{ prime} : z_{t+1} \leqslant p \leqslant z_t\}. \end{aligned} \tag{2.9}$$

This way, $\mathcal{P}_1, \dots, \mathcal{P}_t$ partition the primes in $[2, z]$. The primes $> z$ play no role in $S(\mathcal{A}, z)$ and may be ignored (or assigned to $\mathcal{P}_1$, for example). In the sum (2.7) defining R', we have

$$d_1 \cdots d_t \leqslant z_1^{k_1} \cdots z_t^{k_t}.$$

A convenient choice is

$$z_j = z^{1/4^{j-1}} \qquad (1 \leqslant j \leqslant t), \quad z_{t+1} = 2,$$

where t is chosen maximally so that $z_t \geqslant 2$. Then define $\mathcal{P}_j$ by (2.9). We also take

$$k_1 = 0, \quad k_j = 2^{j-1} \ \ (j \geqslant 2).$$

In (2.7), we have

$$d \leqslant z_1^{k_1} \cdots z_t^{k_t} \leqslant z. \tag{2.10}$$

Invoking (Ω) and taking logarithms, we have

$$L_j = \sum_{z_{j+1} < p \leqslant z_j} -\log(1 - g(p)) \leqslant \kappa \log 4 + \frac{B}{\log z_{j+1}} \leqslant \kappa_0 \log 4 + \frac{B_0}{\log 2} =: L.$$

Therefore, in the notation of Theorem 2.3,

$$E \leqslant \sum_{j=1}^{t} \mathrm{e}^{L_j} \frac{L_j^{k_j}}{k_j!} \leqslant \mathrm{e}^{L} \sum_{j=1}^{\infty} \frac{L^{2^{j-1}+1}}{(2^{j-1}+1)!} \leqslant \mathrm{e}^{2L} \leqslant \mathrm{e}^{3\kappa_0+3B_0}. \tag{2.11}$$

The theorem now follows from Theorem 2.3. □

2.2. **Applications of the upper bound sieve.** Theorem (2.4) is very easy to apply in practice. It suffices to verify that the sifting function $g(p)$ has regular behavior ((Ω) or (Ω_0)) and that the remainders r_d are small on average in some reasonable range.

2.2.1. *Primes.* Take the Eratosthenes sieve $\mathcal{A} = [1, x] \cap \mathbb{N}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$, $X = x$, $g(p) = 1/p$, $z = X^{1/2}$. Then, by Theorem 2.4,

$$\pi(x) \leqslant S(\mathcal{A}, z) + z \ll XV(z) + O(z) \ll \frac{x}{\log x},$$

matching Chebyshev's bound.

2.2.2. *Prime k-tuples.*

Theorem 2.5. *Let $a_1, \dots, a_k$ be non-zero integers and $b_1, \dots, b_k$ integers, with $(a_j, b_j) = 1$ for all j, and the forms $a_j n + b_j$ are distinct. Let $\rho(d)$ be the number of solutions modulo d of*

$$(a_1 n + b_1) \cdots (a_k n + b_k) \equiv 0 \pmod{d},$$

and assume that $\rho(p) < p$ for all primes p (that is, the linear forms $a_i n + b_i$, $1 \leqslant i \leqslant k$, are admissible). Also let

$$\Delta = \left| \prod_{i=1}^{k} a_i \prod_{i<j} (a_i b_j - a_j b_i) \right|.$$

Then $\rho(p) \leqslant k$ for all p, $\rho(p) \neq k$ if and only if $p|\Delta$, and furthermore,

$$\begin{aligned} \#\{n \leqslant x : a_i n + b_i \text{ prime }, 1 \leqslant i \leqslant k\} &\ll_k \frac{\mathfrak{S} x}{\log^k x} \\ &\ll_k \frac{x}{\log^k x} \prod_{p|\Delta} \left(1 - \frac{1}{p}\right)^{\rho(p)-k} \\ &\ll_k \frac{x}{\log^k x} \left(\frac{\Delta}{\phi(\Delta)}\right)^k, \end{aligned}$$

where

$$\mathfrak{S} = \prod_{p} \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}$$

and the implied constants may depend on k alone.

Remarks. If $(a_j, b_j) > 1$ for some j then $\rho(p) = p$ for $p|(a_j, b_j)$. Hence, $\rho(p) < p$ for all p implies that $(a_j, b_j) = 1$ for every j. The other hypotheses then imply that $\Delta \neq 0$.

Proof. Define our sieve problem as follows. Let $\mathcal{A} = [1, x] \cap \mathbb{Z}$, with uniform probability and

$$\mathcal{A}_p = \{n \in \mathcal{A} : p|(a_1 n + b_1) \cdots (a_k n + b_k)\}$$

for each prime p. Let $X = x$ and put $z = X^{1/2}$. By the Chinese remainder theorem, ρ is multiplicative. Now let p be prime and consider $\rho(p)$. Since $(a_i, b_i) = 1$ for all i, the congruence $a_i n + b_i \equiv 0 \pmod{p}$ has a solution r_i if and only if $p \nmid a_i$, and then r_i is unique modulo p. Hence, if $p|a_i$ for some i, then $\rho(p) < k$. If $p \nmid a_1 \cdots a_k$, then $\rho(p)$ is the number of distinct values among $r_1, \dots, r_k$. It is clear that $r_i = r_j$ if and only if $p|(a_i b_j - a_j b_i)$. Thus, if $p|\Delta$ then $\rho(p) < k$, and if $p \nmid \Delta$ then $r_1, \dots, r_k$ all exist and are distinct, so $\rho(p) = k$. This proves the first two claims.

Setting $g(p) = \rho(p)/p$, we see that

$$g(p) \leqslant \min\left(1 - \frac{1}{p}, \frac{k}{p}\right) \leqslant \min\left(1 - \frac{1}{k+1}, \frac{k}{p}\right).$$

Thus, (Ω_0) holds with $\delta = 1/(k+1)$ and $\kappa = k$. Then (Ω) holds with $\kappa = k$ and some B depending only on k by Exercise 2.1. Moreover, $\lfloor x/d \rfloor \rho(d) \leqslant A_d \leqslant \lceil x/d \rceil \rho(d)$, and thus

$$A_d = x\frac{\rho(d)}{d} + r_d, \qquad |r_d| \leqslant \rho(d).$$

Hence,

$$\sum_{d \leqslant z} |r_d| \leqslant \sum_{d \leqslant x^{1/2}} \rho(d) \leqslant x^{1/2} \sum_{d \leqslant x^{1/2}} \frac{\rho(d)}{d} \leqslant x^{1/2} \prod_{p \leqslant z} \left(1 + \frac{k}{p}\right) \ll_k x^{1/2} (\log x)^k$$

by Mertens' theorem. By Theorem 2.4 plus the easy bound

$$V(z) \geqslant \prod_{p \leqslant 2k} (1/p) \prod_{2k < p \leqslant z} (1 - k/p) \gg_k \prod_{2k < p \leqslant z} (1 - k/p) \gg_k \frac{1}{\log^k z} \gg_k \frac{1}{\log^k x},$$

we have

$$S(\mathcal{A}, z) \ll_k xV(z) + x^{1/2}(\log x)^k \ll xV(z).$$

Also, $S(\mathcal{A}, z)$ is the number of $n \leqslant x$ for which $(a_1 n + b_1) \cdots (a_k n + b_k)$ has no prime factor $< z$. This includes all $n \leqslant x$ for which each of the forms $a_i n + b_i$ is a prime larger than z. For each i, there are at most z values of n such that $a_i n + b_i$ is a prime that is $\leqslant z$. Thus,

$$\#\{n \leqslant x : a_i n + b_i \text{ prime }, 1 \leqslant i \leqslant k\} \leqslant kz + S(\mathcal{A}, z) \ll_k xV(z).$$

Next,

$$\begin{aligned} V(z) &= \prod_{p\leqslant z}\left(1-\frac{\rho(p)}{p}\right) \\ &\ll_k \frac{1}{(\log z)^k}\prod_{p\leqslant z}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-k} \\ &= \frac{\mathfrak{S}}{(\log z)^k}\prod_{p> z}\left(1-\frac{\rho(p)}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^{k} \\ &\ll_k \frac{\mathfrak{S}}{(\log x)^k}\prod_{p> 2k}\left(1-\frac{k}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^{k} \\ &\ll_k \frac{\mathfrak{S}}{(\log x)^k}. \end{aligned}$$

Finally, since $\rho(p)=k$ if and only if $p\nmid\Delta$,

$$\mathfrak{S}=\prod_{p\mid\Delta}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-k}\prod_{p\nmid\Delta}\left(1-\frac{k}{p}\right)\left(1-\frac{1}{p}\right)^{-k}.$$

For $p>k$, $(1-k/p)(1-1/p)^{-k}=1+O_k(1/p^2)$, and also $1-h/p\leqslant(1-1/p)^h$ for a non-negative integer h. Thus,

$$\mathfrak{S}\ll_k \prod_{p\mid\Delta}\left(1-\frac{\rho(p)}{p}\right)\left(1-\frac{1}{p}\right)^{-k}\leqslant\prod_{p\mid\Delta}\left(1-\frac{1}{p}\right)^{\rho(p)-k}\leqslant\left(\frac{\Delta}{\phi(\Delta)}\right)^k,$$

where we used $\rho(p)\geqslant 0$ in the last step. □

Remarks. Examining the proof, we see that $\mathfrak{S}\gg_k 1$, that is , $\mathfrak{S}$ is bounded away from zero. On the other hand, $\mathfrak{S}$ is not bounded above, as it is possible to have $\rho(p)=0$ for all primes $p\leqslant M$ for some M, e.g. when, for all i, a_i is divisible by all the primes $\leqslant M$.

Corollary 2.6. *We have, for positive m,*

$$\begin{aligned} \#\{p\leqslant x: p,p+2m \text{ both prime}\} &\ll \frac{2m}{\phi(2m)}\frac{x}{\log^2 x}, \\ \#\{p\leqslant x: p,2p+1 \text{ both prime}\} &\ll \frac{x}{\log^2 x}, \\ \#\{p,q \text{ prime}: p+q=2m\} &\ll \frac{\sigma(m)}{m}\frac{m}{\log^2 m}. \end{aligned}$$

Proof. We apply Theorem 2.5 with $k=2$ in each case. The count of generalized twin primes has the forms $n, n+2m$. Here $\Delta=2m$, $\rho(p)=1$ for $p|2m$ and $\rho(p)=2$ otherwise, and we get the upper bound

$$\#\{p\leqslant x: p,p+2m \text{ both prime}\}\ll\frac{x}{\log^2 x}\prod_{p\mid 2m}(1-1/p)^{-1}\ll\frac{2}{\phi(2m)}\frac{x}{\log^2 x}.$$

In fact,

$$\mathfrak{S} \asymp \prod_{p|2m} \left(1 - \frac{1}{p}\right)^{-1} = \frac{2m}{\phi(2m)}.$$

The count of Sophie Germain primes is easier: here we use the forms $n, 2n+1$ and $\Delta = 2$. For the application to Goldbach's Conjecture, take $x = 2m$ and the pair of forms $n, 2m-n$, which gives $\Delta = 2m$. Again, we have $\rho(p) = 1$ for $p|2m$ and $\rho(p) = 2$ for $p \nmid 2m$. The final estimate comes from the elementary bound $\frac{m}{\phi(m)} \ll \frac{\sigma(m)}{m}$ (the bound is often presented in this way rather than using $2m/\phi(2m)$). □

2.2.3. *The Brun-Titchmarsh inequality.* Denote by $\pi(x; q, a)$ the number of primes $p \leqslant x$ satisfying $p \equiv a \pmod q$.

Theorem 2.7 (Brun-Titchmarsh inequality)**.** *There is a constant C so that uniformly for $1 \leqslant a \leqslant q < y \leqslant x$ and $(a, q) = 1$,*

$$\pi(x; q, a) - \pi(x - y; q, a) \leqslant C \frac{y}{\phi(q) \log(y/q)}.$$

Proof. WLOG assume that $y > 2q$, for otherwise trivially the left side is $\leqslant 2$ while the right side is $\geqslant C/\log(y/q) \geqslant C/\log 2$. Write

$$\{x - y < n \leqslant x : n \equiv a \pmod q\} = \{q(h+1) + a, \dots, q(h+m) + a\}.$$

Apply Theorem 2.5 with $k = 1$ and the single form $qn + (qh + a)$, where $1 \leqslant n \leqslant m = x$. Here $\Delta = q$ and thus

$$\pi(x; q, a) - \pi(x - y; q, a) \ll \frac{q}{\phi(q)} \frac{m}{\log m}.$$

Since $m \asymp y/q$, the theorem follows. □

Remarks. The case $a = q = 1$, that is, showing that $\pi(x) - \pi(x - y) \ll y/\log y$, was asserted by Hardy and Littlewood [**78**, p. 69] without proof, the authors claiming that it followed from the method of Brun [**20**]. Using the Eratosthenes sieve, Hardy and Littlewood [**78**, Theorem G] showed the weaker bound $\pi(x) - \pi(x - y) \ll \frac{y}{\log_2 y}$. The case $x = y$ with an arbitrary a, q was shown by Titchmarsh [**142**]. The constant $C = 2$ is admissible by work of Montgomery and Vaughan [**111**]. The utility of this bound is its uniformity in x, y, a, q. If $\pi(x; q, a) := \#\{p \leqslant x : p \equiv a \pmod q\}$, the prime number theorem for arithmetic progressions implies that

$$\pi(x; q, a) \sim \frac{x}{\phi(q) \log x}$$

for every individual q, a. Issues of uniformity are crucially important, and are intimately connected to questions of the existence of zeros of L-functions lying off the critical line (especially so-called Siegel zeros, real zeros lying very close to 1). If one assumes the Generalized Riemann Hypothesis, then the above asymptotic holds uniformly for $q \leqslant \sqrt{x}$ or so, and in a smaller range if one considers primes in "short" intervals $(x - y, x]$. On the other hand, the Brun-Titchmarsh inequality gives an upper bound of the correct order even for very large moduli $q = x^{1-\varepsilon}$ and for very short intervals $y/q = x^{\varepsilon}$ ($\varepsilon > 0$ fixed). Estimates for individual $q > \sqrt{x}$ are inaccessible by L-function methods.

In the special case $a = q = 1$ we have $\pi(x) - \pi(x - y) \ll y/\log y$. Contigent on the truth of the generalized prime k-tuples conjecture, Conjecture 1.1, this estimate is best possible; that is, for any

$y \geqslant 3$ there are intervals of length y containing $\gg y/\log y$ primes. To see this, let $q_1, \dots, q_k$ denote the primes that are in $(y, 2y]$. Since these are all odd, $k < y$. We claim that the set of linear forms $n+q_1, \dots, n+q_k$ is admissible in the sense of Conjecture 1.1. Here $\rho(p)$ is the number of solutions of the congruence

$$(n+q_1)\cdots(n+q_k) \equiv 0 \pmod{p},$$

which is equal to the number of residue classes modulo p occupied by the numbers $q_1, \dots, q_k$. For any prime $p \leqslant y$, the numbers q_i avoid the residue class $0 \mod p$, thus $\rho(p) < p$. For prime $p > y$, we clearly have $\rho(p) \leqslant k < y \leqslant p$. This proves that the forms $n+q_1, \dots, n+q_k$ are admissible. The conclusion of Conjecture 1.1 implies that for infinitely many n, the numbers $n+q_1, \dots, n+q_k$ are simultaneously prime, hence there are infinitely many intervals of length y having at least k primes. Finally, by the prime number theorem,

$$k = \pi(2y) - \pi(y) \gg \frac{y}{\log y}.$$

2.2.4. *Romanoff's Theorem.*

Theorem 2.8 (Romanoff [**127**], 1934)**.** *A positive proportion*[2] *of positive integers can be expressed as $p + 2^k$ where p is prime and k is a non-negative integer.*

Proof. We start with a common device for lower-bounding the size of a set defined by solutions of some equation. Let

$$r(n) = \#\{(p,k) : k \geqslant 1, n = p + 2^k\}$$

and $B(x) = \#\{n \leqslant x : r(n) > 0\}$. We need to show that $B(x) \gg x$ for large x. Assume that $x \geqslant 100$. By Cauchy-Schwarz,

$$\begin{aligned}\Big(\sum_{n\leqslant x} r(n)\Big)^2 &= \Big(\sum_{n\leqslant x} r(n)\mathbb{1}_{r(n)>0}\Big)^2 \\ &\leqslant \sum_{\substack{n\leqslant x\\ r(n)>0}} 1 \sum_{n\leqslant x} r(n)^2 \\ &= B(x)\sum_{n\leqslant x} r(n)^2.\end{aligned}$$

On the other hand,

$$\sum_{n\leqslant x} r(n) \geqslant \#\{(p,k) : p \leqslant x/2, 2^k \leqslant x/2\} \gg \pi(x/2)\log x \gg x$$

and, since $n \leqslant x$ implies that $p \leqslant x$ and $2^k \leqslant x$,

$$\begin{aligned}\sum_{n\leqslant x} r(n)^2 &= \sum_{n\leqslant x} \#\{(p_1,p_2,k_1,k_2) : p_1 + 2^{k_1} = p_2 + 2^{k_2} = n\} \\ &\leqslant D := \#\{(p_1,p_2,k_1,k_2) : p_j \leqslant x, 2^{k_j} \leqslant x \text{ for } j = 1,2; p_1 + 2^{k_1} = p_2 + 2^{k_2}\}.\end{aligned}$$

It follows that

$$B(x) \gg \frac{x^2}{D}. \tag{2.12}$$

[2]If A is a set of natural numbers with counting function $A(x) = \#\{a \leqslant x : a \in A\}$ satisfying $\liminf_{x\to\infty} A(x)/x > 0$, we say that A contains a positive proportion of all positive integers.

There are $O(x)$ quadruples (p_1, p_2, k_1, k_2) with $p_1 = p_2$ and $k_1 = k_2$. Thus,

$$D \ll x + \#\{(p_1, p_2, k_1, k_2) : p_1 < p_2 < x, 2^{k_j} \leqslant x \text{ for } j = 1, 2; p_2 = p_1 + 2^{k_1} - 2^{k_2}\}.$$

Now fix $k_1 > k_2 \geqslant 1$. Now apply Corollary 2.6 with $2m = 2^{k_1} - 2^{k_2}$. With k_1 and k_2 fixed, the number of possible pairs p_1, p_2 with $p_1 \leqslant x$ is

$$\ll \frac{2^{k_1} - 2^{k_2}}{\phi(2^{k_1} - 2^{k_2})} \frac{x}{\log^2 x}$$

and hence

$$D \ll x + \frac{x}{\log^2 x} \sum_{0 < k_2 < k_1 \leqslant \frac{\log x}{\log 2}} \frac{2^{k_1} - 2^{k_2}}{\phi(2^{k_1} - 2^{k_2})}.$$

Factoring out 2^{k_2} we see that

$$\frac{2^{k_1} - 2^{k_2}}{\phi(2^{k_1} - 2^{k_2})} \ll \frac{2^l - 1}{\phi(2^l - 1)}, \quad l = k_1 - k_2.$$

Also,

$$\frac{2^l - 1}{\phi(2^l - 1)} \ll \frac{\sigma(2^l - 1)}{2^l - 1} = \sum_{d | 2^l - 1} \frac{1}{d}.$$

For each l there are $O(\log x)$ pairs (k_1, k_2) with $k_1 - k_2 = l$ and thus

$$\begin{aligned} D &\ll x + \frac{x}{\log x} \sum_{1 \leqslant l \leqslant \frac{\log x}{\log 2}} \sum_{d | 2^l - 1} \frac{1}{d} \\ &\ll x + \frac{x}{\log x} \sum_{\substack{d \leqslant x \\ d \text{ odd}}} \frac{1}{d} \sum_{\substack{1 \leqslant l \leqslant \frac{\log x}{\log 2} \\ d | 2^l - 1}} 1. \end{aligned}$$

If $s(d)$ denotes the order of 2 modulo d, then the inner sum counts $l \leqslant \frac{\log x}{\log 2}$ which are divisible by $s(d)$ and is therefore $O(\frac{\log x}{s(d)})$. We arrive at

$$D \ll x + x \sum_{\substack{d \leqslant x \\ d \text{ odd}}} \frac{1}{d s(d)}. \tag{2.13}$$

Following Erdős, we show that the sum on d converges, when extended to a sum over all odd positive integers. To this end, define

$$t_k = \sum_{\substack{d \text{ odd} \\ s(d) \leqslant k}} \frac{1}{d},$$

which is finite since $s(d) \gg \log d$. Partial summation gives

$$\sum_d \frac{1}{d s(d)} = \sum_k \frac{t_k - t_{k-1}}{k} = \sum_k \frac{t_k}{k(k+1)}. \tag{2.14}$$

Letting

$$N_k = \prod_{i \leqslant k} (2^i - 1),$$

we see that $s(d) \leqslant k$ implies that $d | N_k$. Also $N_k \leqslant 2^{1+2+\cdots+k} < 2^{k^2}$, and so

$$t_k \leqslant \sum_{d|N_k} \frac{1}{d} = \frac{\sigma(N_k)}{N_k} \ll \log_2 N_k \ll \log k.$$

Thus, the sum on the right side of (2.14) converges, as desired. Consequently, by (2.13), $D \ll x$. Recalling (2.12), the proof is complete. □

Remark 1. Evidently, almost all even integers are not of the form $p + 2^k$. Erdős [**36**] showed that a positive proportion of *odd* numbers are not of the form $p + 2^k$. This paper introduced the notion of covering systems of congruences.

2.2.5. *Sums of primes: Schnirelmann's theorem.* In the 1933, Schnirelmann [**133**] showed that there is a constant k so that every positive integer $\geqslant 2$ is the sum of at most k primes. This was a huge step in the direction of Goldbach's Conjecture. His proof combined the sieve upper bound with a very elementary argument.

Given a set $\mathscr{A} \subset \mathbb{N} \cup \{0\}$, let $A(N)$ be the counting function of $\mathscr{A}$ (excluding zero), that is, $A(N) = \#\{n \in \mathscr{A} : 1 \leqslant n \leqslant N\}$. The *Schnirelmann density* $\sigma(\mathscr{A})$ of $\mathscr{A}$ is defined as

$$\sigma(\mathscr{A}) = \inf_{N \geqslant 1} \frac{A(N)}{N}.$$

Examples:

- $\sigma(\mathscr{A}) = 1$ if and only if $\mathscr{A} \supset \mathbb{N}$
- if $1 \notin \mathscr{A}$ then $\sigma(\mathscr{A}) = 0$.
- $\sigma(\{1, 3, 5, 7, 9, 11, \ldots\}) = 1/2$.
- $\sigma(\{1, 4, 9, 16, 25, \ldots\}) = 0$.

Given any two sets $\mathscr{A}, \mathscr{B}$ of integers, define the *sumset* $\mathscr{A} + \mathscr{B}$ by

$$\mathscr{A} + \mathscr{B} = \{a + b : a \in \mathscr{A}, b \in \mathscr{B}\}.$$

We denote the k-fold subset $k\mathscr{A}$ inductively as $k\mathscr{A} = (k-1)\mathscr{A} + \mathscr{A}$, so that $k\mathscr{A}$ is the set of integers that can be written as the sum of k (not necessarily distinct) elements of $\mathscr{A}$.

Theorem 2.9 (Shnirelmann [**133**], 1933)**.** (a) *Suppose that* $0 \in \mathscr{B}$ *and* $1 \in \mathscr{A}$*. Then*

$$\sigma(\mathscr{A} + \mathscr{B}) \geqslant \sigma(\mathscr{A}) + \sigma(\mathscr{B}) - \sigma(\mathscr{A})\sigma(\mathscr{B}).$$

(b) *If* $0 \in \mathscr{A}, 0 \in \mathscr{B}$ *and* $\sigma(\mathscr{A}) + \sigma(\mathscr{B}) \geqslant 1$ *then* $\mathscr{A} + \mathscr{B} = \mathbb{N} \cup \{0\}$.

The conclusion of part (a) is conveniently written as

$$1 - \sigma(\mathscr{A} + \mathscr{B}) \leqslant (1 - \sigma(\mathscr{A}))(1 - \sigma(\mathscr{B})).$$

In general, the conclusion is best possible, for example if

$$\mathscr{A} = \{1, 5, 6, 7, 8, 9, \ldots\}, \quad \mathscr{B} = \{0, 1, 4, 5, 6, 7, 8, \ldots\}$$

then $\mathscr{A} + \mathscr{B} = \{1, 2, 5, 6, 7, 8, 9, \ldots\}$, $\sigma(\mathscr{A}) = \frac{1}{4}$, $\sigma(\mathscr{B}) = \frac{1}{3}$ and $\sigma(\mathscr{A} + \mathscr{B}) = \frac{1}{2}$.

Corollary 2.10 (Schnirelmann [**133**])**.** *Suppose that* $0 \in \mathscr{A}$ *and* $\sigma(k\mathscr{A}) > 0$ *for some* $k \in \mathbb{N}$*. Then for some* $h \in \mathbb{N}$*,* $h\mathscr{A} = \mathbb{N} \cup \{0\}$.

Proof that Theorem 2.9 implies Corollary 2.10. Let $\mathscr{B} = k\mathscr{A}$. Obviously $0 \in \mathscr{B}$. Since $\sigma(\mathscr{B}) > 0$, we have $1 \in \mathscr{B}$. Hence, $\{0,1\} \subset r\mathscr{B}$ for all $r \in \mathbb{N}$. By Theorem 2.9 (a), for any $r \in \mathbb{N}$ we have

$$1 - \sigma((r+1)\mathscr{B}) \leqslant (1 - \sigma(\mathscr{B}))(1 - \sigma(r\mathscr{B})).$$

By induction,

$$1 - \sigma(r\mathscr{B}) \leqslant (1 - \sigma(\mathscr{B}))^r$$

Hence, for some r, $\sigma(r\mathscr{B}) \geqslant 1/2$. Then, by Theorem 2.9 (b), we conclude that $\sigma(2r\mathscr{B}) = 1$, that is,

$$2rk\mathscr{A} = 2r\mathscr{B} = \mathbb{N} \cup \{0\},$$

as required. □

Proof of Theorem 2.9. We begin with (a). Let $N \in \mathbb{N}$ and enumerate the elements of $\mathscr{A} \cap [1, N]$ as $1 = a_1 < a_2 < \ldots < a_r \leqslant N$. These integers are separated by gaps of $g_i = a_{i+1} - a_i - 1$ numbers not in $\mathscr{A}$, plus a final gap of $g_r = N - a_r$ integers from $a_r + 1$ to N. Let $\mathscr{C} = \mathscr{A} + \mathscr{B}$. Then $\mathscr{C} \cap [1, N]$ contains $\mathscr{A} \cap [1, N]$, plus additional numbers $a_i + b$, where $b \in \mathscr{B}, 1 \leqslant b \leqslant g_i$. The number of such elements b is $\geqslant g_i\sigma(\mathscr{B})$. It follows that

$$\begin{aligned} C(N) &\geqslant A(N) + \sum_{i=1}^{r} \sigma(\mathscr{B}) g_i \\ &= A(N) + \sigma(\mathscr{B})(N - A(N)) \\ &= A(N)(1 - \sigma(\mathscr{B})) + N\sigma(\mathscr{B}) \\ &\geqslant N\Big(\sigma(\mathscr{A})(1 - \sigma(\mathscr{B})) + \sigma(\mathscr{B})\Big). \end{aligned}$$

As this is true for every N, $\sigma(\mathscr{C}) \geqslant \sigma(\mathscr{A}) + \sigma(\mathscr{B}) - \sigma(\mathscr{A})\sigma(\mathscr{B})$.

To prove (b), first observe that $1 \in \mathscr{A} \cup \mathscr{B}$, so that $1 \in \mathscr{A} + \mathscr{B}$. Suppose that $\mathscr{A} + \mathscr{B} \neq \mathbb{N} \cup \{0\}$ and that $n \notin \mathscr{A} + \mathscr{B}$. Then $n \geqslant 2$, $n \notin \mathscr{A}$ and $n \notin \mathscr{B}$. Thus,

$$A(n-1) + B(n-1) = A(n) + B(n) \geqslant n(\sigma(\mathscr{A}) + \sigma(\mathscr{B})) \geqslant n.$$

Now let

$$\mathscr{C} = \{a \in \mathscr{A} : 1 \leqslant a \leqslant n-1\}, \qquad \mathscr{D} = \{n - b : 1 \leqslant b \leqslant n-1, b \in \mathscr{B}\}.$$

Then $|\mathscr{C}| + |\mathscr{D}| \geqslant n$. But $\mathscr{C}, \mathscr{D} \subset [1, n-1]$, hence by the pigeonhole principle, $\mathscr{C}$ and $\mathscr{D}$ must have a common element, thus $n \in \mathscr{A} + \mathscr{B}$, a contradiction. Therefore, $\mathscr{A} + \mathscr{B} = \mathbb{N} \cup \{0\}$. □

We now apply these results to primes. Let $\mathscr{P}$ denote the set of primes and set $\mathscr{P}_0 = \mathscr{P} \cup \{0, 1\}$. Now $\sigma(\mathscr{P}_0) = 0$ by the prime number theorem. However, by Exercise 2.5 below, we have $\sigma(2\mathscr{P}_0) > 0$. Hence, by Corollary 2.10, there is an h so that $h\mathscr{P}_0 = \mathbb{N} \cup \{0\}$. We also have $h \geqslant 3$, since $27 \notin 2\mathscr{P}_0$. Thus, for any positive integer $n \geqslant 4$, we have

$$n - 2 = p_1 + \cdots + p_r + s \cdot 1$$

where $p_1, \ldots, p_r$ are primes, $r \geqslant 0$, $s \geqslant 0$ and $r + s \leqslant h$. Now $2 + s = 2k + 3l$ where $k \geqslant 0, l \geqslant 0$ and $k + l \leqslant s + 1$. Thus, n is the sum of at most $r + s + 1 = h + 1$ primes. We conclude that

Theorem 2.11 (Schnirelmann [**133**], 1933)**.** *For some $h \in \mathbb{N}$, every positive integer $\geqslant 2$ is the sum of at most h primes.*

In 1937, I. M. Vinogradov (see [**145**]) proved that every sufficiently large odd integer is the sum of three primes. The proof used a version of the circle method and did not utilize sieve ideas. The precise measure of "sufficiently large" was shown by various authors, finally culminating in the work of Helfgott, who proved that all off integers $\geqslant 7$ are the sum of three primes [**88**].

2.2.6. *Primitive roots of primes.* Are there infinitely many primes p for which the fraction $1/p$ has period $p-1$ in base 10? Equivalently, is 10 a primitive root of p for infinitely many primes p? This question was addressed by Gauss in his Disquisitiones Arithmeticae (1801). In 1927, E. Artin conjectured an asymptotic formula for the number of primes $p \leqslant x$ for which a given squarefree integer a is a primitive root; his formula was later discovered not to hold up against numerical data by D. H. Lehmer, and a revised conjecture was later proposed by H. Heilbronn and others. In 1967, C. Hooley deduced the revised conjecture from the Generalized Riemann Hypothesis for certain Dedekind zeta functions $\zeta_K(s)$ [**90**]; see also [**94**, Ch. 3]. The theorem makes use of sieve methods. Here we will specialize to the case $a = 2$. For background on Dedekind zeta functions, see [**98**, Sec. 5.10].

Theorem 2.12 (Hooley [**90**])**.** *Assume the Generalized Riemann Hypothesis for $\zeta_K(s)$ for the number fields $K = \mathbb{Q}(\sqrt[k]{2}, e^{2\pi i/k})$, k running over all squarefree integers. Then*

$$\#\{p \leqslant x : 2 \text{ is a primitive root of } p\} \sim C\pi(x), \quad C = \prod_{q \text{ prime}} \left(1 - \frac{1}{q(q-1)}\right).$$

Proof. We set this up as a general sieve problem as follows. Let $\mathcal{A}$ be the set of primes $p \leqslant x$, and for any prime q, let

$$\mathcal{A}_q = \Big\{p \in \mathcal{A} : q | p-1 \text{ and } 2^{(p-1)/q} \equiv 1 \pmod p\Big\}.$$

Clearly, 2 is a primitive root of p (order $p-1$) if and only if $p \notin \mathcal{A}_q$ for all primes q (and it suffices to check for primes $q \leqslant p$). We need good, *uniform* bounds for A_d. Bounds for individual d, but with poor uniformity, are available unconditionally by the analog of the prime number theorem for the number fields K (the prime ideal theorem of Landau), but the uniformity is very poor, especially for $x^{o(1)} < d \leqslant \sqrt{x}$, and not good enough for our application.

Lemma 2.13 (Hooley [**90**])**.** *Assume the Generalized Riemann Hypothesis for the Dedekind zeta functions $\zeta_K(s)$ for all of the number fields $K = \mathbb{Q}(\sqrt[k]{2}, e^{2\pi i/k})$, k running over all squarefree integers. Then, uniformly for squarefree $d \leqslant x$,*

$$A_d = \frac{\operatorname{li}(x)}{d\phi(d)} + O(\sqrt{x}\log x).$$

Define the cutoffs

$$z_1 = \frac{1}{6}\log x, \quad z_2 = \sqrt{x}(\log x)^{-2}, \quad z_3 = \sqrt{x}\log x.$$

Evidently,

$$S(\mathcal{A}, z_1) \geqslant \#\{p \leqslant x : 2 \text{ is a primitive root of } p\} \geqslant S(\mathcal{A}, z_1) - S_1 - S_2 - T, \tag{2.15}$$

where

$$S_1 = \sum_{z_1 < q \leqslant z_2} A_q, \quad S_2 = \sum_{z_2 < q \leqslant z_3} A_q, \quad T = \#\{p \leqslant x : p \in \mathcal{A}_q \text{ for some } q > z_3\}.$$

Using the Legendre sieve (Theorem 1.3) with $X = \mathrm{li}(x)$ and with $g(d) = \frac{1}{d\phi(d)}$, Lemma 2.13 gives

$$\begin{aligned} S(\mathcal{A}, z_1) &= \mathrm{li}(x) \prod_{q \leqslant z_1} \left(1 - \frac{1}{q(q-1)}\right) + O(\sqrt{x}(\log x) 2^{\pi(z_1)}) \\ &= \left(C + O\left(\frac{1}{\log x}\right)\right) \mathrm{li}(x) = \left(C + O\left(\frac{1}{\log x}\right)\right) \pi(x). \end{aligned}$$

For S_1, we again use Lemma 2.13 and obtain

$$\begin{aligned} S_1 &= \sum_{z_1 < q \leqslant z_2} \left(\frac{\mathrm{li}(x)}{q(q-1)} + O(\sqrt{x} \log x)\right) \\ &\ll \frac{\mathrm{li}(x)}{\log x} + \sqrt{x}(\log x)\pi(z_2) \ll \frac{\pi(x)}{\log x}. \end{aligned}$$

For S_2, the bounds from Lemma 2.13 are too poor, but we do better by observing that $A_q \leqslant \pi(x; q, 1)$ and using Theorem 2.7 (the Brun-Titchmarsh inequality):

$$\begin{aligned} S_2 &\leqslant \sum_{z_2 < q \leqslant z_3} \pi(x; q, 1) \ll \sum_{z_2 < q \leqslant z_3} \frac{x}{q \log x} \ll \frac{x}{\log^2 x} \sum_{z_2 < q \leqslant z_3} \frac{\log q}{q} \\ &\ll \frac{x}{\log^2 x} \log \frac{z_3}{z_2} \ll \frac{x \log_2 x}{\log^2 x} \ll \pi(x) \frac{\log_2 x}{\log x}. \end{aligned}$$

Finally, if $p \in \mathcal{A}_q$ for some $q > z_3$, then $p | (2^m - 1)$ for some positive integer $m \leqslant \sqrt{x}/\log x$. Thus, $p | M$, where

$$M = \prod_{m \leqslant \sqrt{x}/\log x} (2^m - 1).$$

Thus,

$$T \leqslant \#\{p | M : p \text{ prime}\} \leqslant \frac{\log M}{\log 2} \ll \left(\frac{\sqrt{x}}{\log x}\right)^2 = \frac{x}{\log^2 x} \ll \frac{\pi(x)}{\log x}.$$

Combining the estimates for $S(\mathcal{A}, z_1)$, S_1, S_2 and T with (2.15) proves the theorem. □

In 1986, Heath-Brown [**87**] showed that for all but at most two primes r,

$$\{p : r \text{ is a primitive root of } p\}$$

is infinite. We cannot say which two prime might be exceptional, or even give a bound on the exceptional primes. We will revisit this type of result later.

2.3. **Exercises.**

Exercise 2.1. *Assume condition* (Ω_0). *Show that* (Ω) *holds with the same value of* κ *and with* B *depending only on* δ, κ.

Exercise 2.2. *For each prime, let* $\mathcal{I}_p$ *denote a set of residue classes modulo* p. *suppose that* $|\mathcal{I}_p| \leqslant \kappa$ *for all* p. *Show that*

$$\#\{x < n \leqslant x + y : \forall p \leqslant z, n \mod p \notin \mathcal{I}_p\} \ll_\kappa y \prod_{p \leqslant z} \left(1 - \frac{|\mathcal{I}_p|}{p}\right),$$

uniformly for $2 \leqslant z \leqslant y$ *and all* x.

Exercise 2.3. *Let k be a positive, even integer. Assuming Conjecture 1.6 for the pair of linear forms $(n, n+k)$, prove that*

$$\#\{n \leqslant x : n \text{ and } n+k \text{ are } \textbf{consecutive} \text{ primes}\} \sim \frac{a(k)x}{\log^2 x} \qquad (x \to \infty),$$

where

$$a(k) = C \prod_{p|k, p>2} \frac{p-1}{p-2}, \qquad C = 2\prod_{p \geqslant 3} \left(1 - \frac{1}{(p-1)^2}\right),$$

C being the "twin prime constant".

Exercise 2.4. *Show that for any odd N,*

$$\#\{(p_1, p_2, p_3) : N = p_1 + p_2 + p_3\} \ll \frac{N^2}{\log^3 N},$$

the constant being absolute.

Exercise 2.5. *We say that an even number k is a* Goldbach number *if there exists primes p_1, p_2 such that $k = p_1 + p_2$. Show that a positive proportion of all positive integers are Goldbach numbers.*

Exercise 2.6. *Let $\Xi(x, y, z)$ denote the number of integers $n \leqslant x$ which have no prime factor in $(y, z]$. Prove that uniformly in $1.5 \leqslant y \leqslant z \leqslant x$, we have*

$$\Xi(x, y, z) \ll \frac{x \log y}{\log z}.$$

Exercise 2.7. *Prove that*

$$\sum_{p \leqslant x} \Omega(p-1) \ll \frac{x \log_2 x}{\log x},$$

where $\Omega(n)$ is the number of prime power divisors of n.

Exercise 2.8. *Let $\mathcal{S}$ denote the set of integers that are the sum of two squares.*

(a) *Show that $\mathcal{S}$ has counting function $O(x(\log x)^{-1/2})$.*

(b) *Let $h \in \mathbb{N}$. Show that, uniformly in h,*

$$\#\{n \leqslant x : n \in \mathcal{S}, n+h \in \mathcal{S}\} \ll \prod_{\substack{p|h \\ p \equiv 3 \pmod 4}} \left(1 + \frac{1}{p}\right) \frac{x}{\log x}.$$

Exercise 2.9. *(Landau-Lehmer problem). Let $q \in \mathbb{N}$, $q \geqslant 2$, and let $\mathcal{R}$ be a collection of reduced residue classes modulo q. Prove that the number of integers $n \leqslant x$ lacking any prime factor that lies in one of the congruence classes of $\mathcal{R}$ is bounded above by*

$$\ll_q \frac{x}{(\log x)^{|\mathcal{R}|/\phi(q)}}.$$

Exercise 2.10. *Let $\mathscr{A}$ denote the set of positive, squarefree integers.*

(a) Show that the Schnirelmann density $\sigma(\mathscr{A}) > 1/2$.

(b) Show that every integer $n \geqslant 2$ is the sum of exactly two squarefree integers.(This does not *follow immediately from Theorem 2.9 (b)).*

Exercise 2.11. *Let $\mathcal{B}$ be the set of primes that are of the form a^2+b^4 for some integers a, b. Prove that*

$$\#\{p \leqslant x : p \in \mathcal{B}\} \ll \frac{x^{3/4}}{\log x}.$$

Exercise 2.12. *Let $1 \leqslant a \leqslant q$, $(a,q)=1$. Prove that uniformly in a, q and $x \geqslant 10$ we have*

$$\sum_{\substack{q\leqslant p\leqslant x\\ p\equiv a \pmod q}} \frac{1}{p} \ll \frac{\log_2 x}{\phi(q)}.$$

Note that the condition $p \geqslant q$ is necessary, as the least prime that is $\equiv a \pmod q$ might be a, which could be very small.

3. The Brun-Hooley sieve: lower bounds

Our lower bound for $S(\mathcal{A}, z)$ is derived from the upper bound in Lemma 2.2 by subtracting off appropriate quantities. We use the following simple inequality:

Lemma 3.1. *Suppose that $0 \leqslant x_j \leqslant y_j$ for $1 \leqslant j \leqslant t$. Then*

$$x_1 \cdots x_t \geqslant y_1 \cdots y_t - \sum_{\ell=1}^{t} (y_\ell - x_\ell) \prod_{\substack{j=1\\ j\neq \ell}}^{t} y_j.$$

Proof. The inequality is an equality when $t=1$, and follows by induction on t using

$$\begin{aligned} y_1 \cdots y_t - x_1 \cdots x_t &= (y_1 \cdots y_{t-1} - x_1 \cdots x_{t-1}) y_t + x_1 \cdots x_{t-1}(y_t - x_t) \\ &\leqslant (y_1 \cdots y_{t-1} - x_1 \cdots x_{t-1}) y_t + y_1 \cdots y_{t-1}(y_t - x_t). \end{aligned}$$

□

As before, partition the primes in $[2, z]$ into sets $\mathcal{P}_1, \dots, \mathcal{P}_t$ and let $k_1, \dots, k_t$ be nonnegative even integers. We apply Lemma 3.1 with

$$x_j = \sum_{d \mid (n, P_j)} \mu(d), \qquad y_j = \sum_{\substack{d\mid (n,P_j)\\ \omega(d)\leqslant k_j}} \mu(d) \quad (1 \leqslant j \leqslant t).$$

From Lemma 1.6 (with $g(p)=1$ for each p) we have that

$$0 \leqslant y_\ell - x_\ell \leqslant \sum_{\substack{d\mid (n,P_\ell)\\ \omega(d)=k_\ell+1}} 1.$$

Hence by Lemma 2.2 and Lemma 3.1, for any n with $P^+(n) \leqslant z$ we obtain

$$\mathbb{1}_{n=1} = \sum_{d\mid n} \mu(d) \geqslant \prod_{j=1}^{t} \sum_{\substack{d_j\mid (n,P_j)\\ \omega(d_j)\leqslant k_j}} \mu(d_j) - \sum_{\ell=1}^{t} \Bigg(\sum_{\substack{d_\ell\mid (n,P_\ell)\\ \omega(d_\ell)=k_\ell+1}} 1 \Bigg) \prod_{\substack{j=1\\ j\neq \ell}}^{t} \sum_{\substack{d_j\mid (n,P_j)\\ \omega(d_j)\leqslant k_j}} \mu(d_j). \tag{3.1}$$

Thus, we have a lower bound sieve with coefficients (here $d = d_1 \cdots d_t$ with $d_i = (d, P_i)$ for each i)

$$\lambda_d^- = \begin{cases} \mu(d) & \text{if } d_j \mid P_j, \omega(d_j) \leqslant k_j \ \ (1 \leqslant j \leqslant t) \\ \mu(d) & \text{if } d_j \mid P_j \ (1 \leqslant j \leqslant t),\ \exists \ell : \omega(d_j) \leqslant k_j \ (j \neq \ell), \omega(d_\ell) = k_\ell + 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

From (1.7),

$$S(\mathcal{A},z) \geqslant \sum_{\substack{d_1,\dots,d_t\\ \forall j: d_j\mid P_j, \omega(d_j)\leqslant k_j}} \mu(d_1)\cdots\mu(d_j)A_{d_1\cdots d_t} - \sum_{\ell=1}^{t} \sum_{\substack{d_1,\dots,d_t\\ d_j\mid P_j,\omega(d_j)\leqslant k_j\ (j\neq\ell)\\ d_\ell\mid P_\ell,\omega(d_\ell)=k_\ell+1}} \mu\left(\frac{d_1\cdots d_t}{d_\ell}\right) A_{d_1\cdots d_t}.$$

As before, introduce the conditions (g) and (r), and the shorthand notation (V). Adopting our previous notation (2.3) for quantities U_j, we can rewrite this as

$$S(\mathcal{A},z) \geqslant XU_1\cdots U_t\left(1-\sum_{\ell=1}^{t}\frac{1}{U_\ell}\sum_{\substack{d_\ell\mid P_\ell\\ \omega(d_\ell)=k_\ell+1}} g(d_\ell)\right) - \widetilde{R},$$

where

$$\widetilde{R} = \sum_{d\in\mathcal{D}} |r_d|,$$

and $\mathcal{D}$ is the set of squarefree numbers d that satsify $P^+(d)\leqslant z$, $\omega((d,P_j))\leqslant k_j+1$ for all j and $\omega((d,P_j))=k_j+1$ for at most one j. Define V_j and L_j by (2.4). By Lemma 1.6, $U_i\geqslant V_i$ for every i, thus in particular

$$U_1\cdots U_t \geqslant V(z), \qquad 1/U_\ell \leqslant e^{L_\ell \mathbb{1}_{k_\ell>0}}.$$

Another invocation of the Erdős trick (cf., the proof of Lemma 1.6) gives

$$\sum_{\substack{d_\ell\mid P_\ell\\ \omega(d_\ell)=k_\ell+1}} g(d_\ell) \leqslant \frac{1}{(k_\ell+1)!}\Bigg(\sum_{p\in P_\ell} g(p)\Bigg)^{k_\ell+1} \leqslant \frac{L_\ell^{k_\ell+1}}{(k_\ell+1)!}.$$

We partition the primes in $[2,z]$ as in (2.9). Then

$$\widetilde{R} \leqslant \sum_{\substack{d\in\mathscr{P}(z)\\ d\leqslant D}} |r_d|, \qquad D = z_1^{k_1+1}z_2^{k_2}\cdots z_t^{k_t}.$$

We arrive at a first general lower bound sieve.

Theorem 3.2. *Let $\mathcal{A}$ be a sieve problem and assume (g), (r) and (V). Partition the set of primes in $[2,z]$ as in (2.9) and let $k_1,\dots,k_t$ be nonnegative even integers. Then*

$$S(\mathcal{A},z) \geqslant XV(z)(1-E) - R'',$$

where E is given by (2.6) and

$$R'' = \sum_{\substack{d\in\mathscr{P}(z)\\ d\leqslant z_1^{k_1+1}z_2^{k_2}\cdots z_t^{k_t}}} |r_d|, \tag{3.3}$$

In terms of the lower bound sieve coefficients (3.2), we have shown that

$$(1-E)V(z) \leqslant \sum_d \lambda_d^- g(d) \leqslant V(z), \tag{3.4}$$

where E is given by (2.6) and the upper bound comes from (1.8).

Incorporating (Ω), we can derive a lower bound of general utility.

Theorem 3.3. *Adopt the notation (g), (r) and (V). Assume also (Ω) with $\kappa > 0$. Then, if z is large enough as a function of κ, B, we have*

$$S(\mathcal{A}, z) \geqslant 0.00001XV(z) - \sum_{\substack{d \in \mathscr{P}(z) \\ d \leqslant z^{f(\kappa)}}} |r_d|,$$

where $f(\kappa)$ is a function satisfying

(i) $f(1) = 3.35$;
(ii) $f(2) = 7.8$;
(iii) $f(\kappa) = 1 + O(2^{-1/\kappa})$ *for* $0 < \kappa \leqslant 1$;
(iv) $f(\kappa) = \xi\kappa + O(\kappa^{1/2}\log(\kappa+3))$ *for* $\kappa \geqslant 1$, *where* $\xi = 3.5911214766\ldots$ *is the unique solution of* $1/\xi - \log\xi + 1 = 0$.

Proof. Fix κ, suppose $1 = \alpha_1 > \alpha_2 > \alpha_3 > \cdots$ with

$$\alpha_j > 0 \ \ (j \geqslant 1), \quad \alpha_{j+1} \geqslant \alpha_j^2 \ \ (j \text{ large}).$$

Suppose also $\ell_1, \ell_2, \ldots$ are non-negative, even integers with

$$\ell_j \geqslant 2 \quad (j \geqslant 1).$$

The numbers α_j and ℓ_j are chosen to depend on κ alone. Define

$$f = 1 + \sum_{j=1}^{\infty} \alpha_j \ell_j, \qquad E' = \sum_{j=1}^{\infty} \left(\frac{\alpha_j}{\alpha_{j+1}}\right)^{\kappa \mathbb{1}_{\ell_j > 0}} \frac{(\kappa \log(\alpha_j/\alpha_{j+1}))^{\ell_j+1}}{(\ell_j+1)!},$$

and assume that the numbers α_j, ℓ_j are chosen so that f is finite and $E' \leqslant 1$.

Suppose we have a sieve problem, number X and function g satisfying (g), and adopt the notation (r) and (V). Suppose that (Ω) holds for some constant B. Define numbers $z_1 = z > z_2 > \cdots > z_t > z_{t+1} = 2$ and sets $\mathcal{P}_j$ as in (2.9) with $z_j = z^{\alpha_j}$ for $1 \leqslant j \leqslant t$, and such that t is the smallest index with $1/(\log\log z)^8 \leqslant \alpha_t \leqslant 1/(\log\log z)^4$. Since $a_{j+1} \geqslant \alpha_j^2$ for large j, such t exists whenever z is large enough. Furthermore, since $\ell_j \geqslant 2$ for $j \geqslant 2$ and f is finite, we have that $a_j \ll 1/j$ and consequently

$$t \ll (\log\log z)^4.$$

Set

$$k_j = \ell_j \ \ (1 \leqslant j \leqslant t-1), \quad k_t = 2\lfloor (\log\log z)^2 \rfloor.$$

We then have $z_t^{k_t} = z^{o(1)}$ as $z \to \infty$. Adopting the notation R'' from Theorem 3.2, we see that

$$R'' \leqslant \sum_{\substack{d \in \mathscr{P}(z) \\ d \leqslant D}} |r_d|, \quad D = z^{f+o(1)} \ \ (z \to \infty).$$

By (2.4) and (Ω),

$$L_j \leqslant \kappa \log(\alpha_j/\alpha_{j+1}) + \frac{B}{\log z_{j+1}} = \kappa \log(\alpha_j/\alpha_{j+1}) + \frac{B}{\alpha_{j+1}\log z}. \tag{3.5}$$

In particular,

$$L_j \leqslant \kappa \log(\alpha_j/\alpha_{j+1}) + O\left(\frac{(\log\log z)^8}{\log z}\right) \qquad (1 \leqslant j \leqslant t-1)$$

and

$$L_t \leqslant \log\left(\frac{\log z}{\log 2}\right) + O(1) \ll \log\log z.$$

For brevity, let $K_j = \kappa\log(\alpha_j/\alpha_{j+1})$ and $\delta = (\log\log z)^8/\log z$. Recalling the definition (2.6) of E, we have

$$\begin{aligned} E &\leqslant e^{O(\delta)} \sum_{j=1}^{t-1} \mathrm{e}^{K_j \mathbb{1}_{k_j>0}} \frac{(K_j+O(\delta))^{k_j+1}}{(k_j+1)!} + e^{O(\log\log z)} \frac{(O(\log\log z))^{k_t+1}}{(k_t+1)!} \\ &= (1+o(1)) \sum_{j=1}^{t-1} \mathrm{e}^{K_j \mathbb{1}_{\ell_j>0}} \frac{(K_j+O(\delta))^{\ell_j+1}}{(\ell_j+1)!} + o(1) \qquad (z\to\infty). \end{aligned}$$

Since $K_j \ll \log_3 z$ for $1 \leqslant j \leqslant t-1$, the binomial theorem implies

$$(K_j+O(\delta))^{\ell_j+1} = K_j^{\ell_j+1} + O(\delta)2^{\ell_j+1}(O(\log_3 z))^{\ell_j}.$$

Therefore,

$$\begin{aligned} \sum_{j=1}^{t-1} \mathrm{e}^{K_j \mathbb{1}_{\ell_j>0}} \frac{(K_j+O(\delta))^{\ell_j+1}}{(\ell_j+1)!} &\leqslant E' + O(\delta t) \max_{1\leqslant j\leqslant t-1} \frac{(O(\log_3 z))^{\ell_j+1}}{(\ell_j+1)!} \\ &\leqslant E' + O(\delta t)\mathrm{e}^{O(\log_3 z)} \\ &\leqslant E' + o(1) \quad (z\to\infty). \end{aligned}$$

Invoking Theorem 3.2, we find that for sufficiently large z,

$$S(\mathcal{A},z) \geqslant (E'-0.00001)XV(z) - \sum_{\substack{d\in\mathscr{P}(z)\\ d\leqslant z^{f+0.00001}}} |r_d|.$$

The desired concluding in the theorem then follows provided that

$$f(\kappa) \geqslant f + 0.00001, \qquad E' \leqslant 1 - 0.00002. \tag{3.6}$$

To prove the various parts of the theorem, we need only exhibit sequences α_j and ℓ_j satisfying (3.6), $\ell_j \geqslant 2$ for $j \geqslant 2$ and $\alpha_{j+1} \geqslant \alpha_j^2$ for large j.

(i) When $\kappa = 1$, take $\ell_1 = 0$, $\ell_j = 2$ for $j \geqslant 1$, and

$$\alpha_j = \frac{96}{(j+2.5)^{3.5}}.$$

A computer calculation then gives $f \leqslant 3.348$ and $E' \leqslant 0.998$.

(ii) Let $\ell_1 = 0$, $\ell_j = 2$ for $j \geqslant 1$ and

$$\alpha_j = \frac{8774}{(j+8)^4}.$$

A computer calculation gives $f \leqslant 7.79$ and $E' \leqslant 0.9997$.

(iv) When $0 < \kappa \leqslant 1$, let $\ell_1 = 0$, $\ell_j = 2$ for $j \geqslant 1$ and

$$\alpha_j = \left(\frac{96}{(j+2.5)^{3.5}}\right)^{1/\kappa}.$$

By the calculation in part (i), $E' \leqslant 0.998$. Also,

$$f = 1 + O(\alpha_2) = 1 + O(2^{-1/\kappa}).$$

(iv) Let $c = 1/\xi = 0.278\cdots$ so that $1 + c + \log c = 0$. For some positive integers m_j, to be determined later, set

$$\ell_j = 2j - 2 + 2m_j \quad (j \geqslant 1).$$

Also let

$$\alpha_j = \exp\Big\{ - \frac{c(j-1)^2}{\kappa} \Big\}.$$

Utilizing the inequality $k! \geqslant (k/\mathrm{e})^k$, we have

$$\begin{aligned} E' &= \sum_{j=1}^{\infty} e^{c(2j-1)} \frac{(c(2j-1))^{2j-1+2m_j}}{(2j-1+2m_j)!} \\ &\leqslant \sum_{j=1}^{\infty} e^{c(2j-1)} \frac{(c(2j-1))^{2j-1}(c(2j-1))^{2m_j}}{(2j-1)!(2j-1)^{2m_j}} \\ &\leqslant \sum_{j=1}^{\infty} c^{2m_j} \exp\big\{c(2j-1) + (2j-1)(1+\log c)\big\} \\ &= \sum_{j=1}^{\infty} c^{2m_j}. \end{aligned}$$

Taking $m_j = \lceil \log j \rceil + 1$ gives

$$E \leqslant \sum_{j=1}^{\infty} c^2/j^2 < 0.2.$$

We now estimate f. Since $e^{-cx^2/\kappa}$ is a decreasing function for $x \geqslant 0$, we have

$$\begin{aligned} f &\leqslant 1 + 2\ell_1 + \sum_{j=2}^{\infty} \ell_j \int_{j-2}^{j-1} e^{-cx^2/\kappa}\, dx \\ &\leqslant O(1) + \int_0^{\infty} (2x + O(\log(x+3))) e^{-cx^2/\kappa}\, dx \\ &= \kappa/c + O(\kappa^{1/2}\log(\kappa+3)). \end{aligned}$$

□

3.0.1. *Application: twin primes.* We take $\mathcal{A} = [1,x] \cap \mathbb{Z}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p | n(n+2)\}$, $X = x$. As we saw previously, $g(p) = \rho(p)/p$ with $\rho(2) = 1$, $\rho(p) = 2$ for $p > 2$, so (Ω_0) holds with $\kappa = 2$, as does Ω. Taking $z = x^{1/7.9}$ we have

$$\sum_{d \leqslant z^{f(2)}} \mu^2(d)|r_d| \leqslant \sum_{d \leqslant z^{f(2)}} \tau(d) \ll z^{f(2)} \log z \ll x^{0.99},$$

which is negligible. Since $V(z) \gg 1/\log^2 z \gg 1/\log^2 x$ we conclude from Theorem 3.3 that

$$S(\mathcal{A}, z) \gg \frac{x}{\log^2 x}.$$

We conclude that there are $\gg x/\log^2 x$ integers $n \leqslant x$, such that each of n and $n+2$ have at most 7 prime factors, as these factors are $> z$.

Now we approach the problem as in Example (c) (this is A. Rènyi's approach), taking

$$\mathcal{A} = \{p+2 : p \text{ prime } \leqslant x\}, \quad X = \mathrm{li}(x),$$

and $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$. Here

$$A_d = \pi(x; d, -2) = \frac{\mathrm{li}(x)}{\phi(d)} + r_d,$$

where r_d is expected to be small by the prime number theorem for arithmetic progressions. Taking $g(d) = 1/\phi(d)$, we easily verify the (Ω) holds with $\kappa = 1$. Indeed, for $y \geqslant 3$, Mertens' bounds give

$$\begin{aligned}\prod_{y\leqslant p\leqslant w} (1-g(p))^{-1} &= \prod_{y\leqslant p\leqslant w} \left(1-\frac{1}{p-1}\right)^{-1} = \prod_{y\leqslant p\leqslant w} \left(1-\frac{1}{p}\right)^{-1}\left(1-\frac{1}{(p-1)^2}\right)^{-1} \\ &\leqslant \frac{\log w}{\log y}\left(1+O\left(\frac{1}{\log y}\right)\right).\end{aligned}$$

For the error terms r_d, we use the famous theorem of Bombieri and Vinogradov:

Theorem 3.4 (Bombieri-A.I.Vinogradov, 1965)**.** *For every $A > 0$ there is a $B > 0$ so that*

$$\sum_{q\leqslant x^{1/2}(\log x)^{-B}} \max_{y\leqslant x} \max_{(a,q)=1} \left|\pi(y;q,a) - \frac{\mathrm{li}(y)}{\phi(q)}\right| \ll \frac{x}{(\log x)^A}.$$

In Theorem 3.3, we have $f(1) = 3.35$. Take $z = x^{\frac{1}{6.8}}$ so that $z^{f(1)} \leqslant x^{0.493}$. Hence, Theorem BV implies

$$\sum_{d\leqslant z^{f(1)}} |r_d| = \sum_{d\leqslant z^{f(1)}} \left|\pi(x;d,-2) - \frac{\mathrm{li}(x)}{\phi(d)}\right| \ll \frac{x}{(\log x)^{10}}.$$

We conclude that for large x,

$$S(\mathcal{A}, z) \gg XV(z) \gg \frac{x}{\log^2 x}.$$

Finally, we observe that $S(\mathcal{A}, z)$ counts primes p such that $p+2$ has no prime factor $\leqslant z$; in particular, $p+2$ has at most 6 prime factors.

3.0.2. *Prime values of polynomials.* Generalizing greatly the study of twin primes, we can use the sieve to study prime values of arbitrary polynomials.

Theorem 3.5. *Let $F_1, \dots, F_k$ be distinct, irreducible polynomials in $\mathbb{Z}[x]$, each with positive leading coefficient. Put $F = F_1 \cdots F_k$, $\ell = \deg(F)$. Suppose $(F_1, \dots, F_k)$ is admissible. Then*

$$\#\{n \leqslant x : F_i(n) \text{ prime for every } i\} \ll_F \frac{x}{\log^k x}. \tag{3.7}$$

Further, there is an integer $m = O(k\ell)$ such that for large x,

$$\#\{n \leqslant x : \Omega(F_i(n)) \leqslant m \ \ (1 \leqslant i \leqslant k)\} \gg_F \frac{x}{\log^k x}. \tag{3.8}$$

Proof. We will show the upper bound (3.7), leaving (3.8) as an exercise (cf., Exercise 3.1 below). Let x be large, $\mathcal{A} = [1, x] \cap \mathbb{N}$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|F(n)\}$ for primes p, $X = x$, $z = x^{1/2}$. Generically write

$$\rho_G(d) = \#\{0 \leqslant n < d : G(n) \equiv 0 \pmod d\},$$

where G is any polynomial. As with twin primes, we have

$$A_d = \frac{\rho_F(d)}{d}X + r_d, \qquad |r_d| \leqslant \rho_F(d),$$

and so we set $g(d) = \rho_F(d)/d$. Now ρ_F is multiplicative by the Chinese remainder theorem. Also,

$$\rho_F(p) < p \quad \text{by hypothesis,}$$
$$\rho_F(p) \leqslant \ell \quad \text{by Lagrange's theorem.}$$

In particular, (Ω_0) holds with $\kappa = \ell$ and $\delta = 1/\ell$. Thus,

$$\sum_{d\leqslant z}|r_d| \leqslant \sum_{d\leqslant x^{1/2}} dg(d) \leqslant x^{1/2}\sum_{d\in\mathscr{P}(z)} g(d) = x^{1/2}\prod_{p\leqslant z}(1+g(p))$$
$$\leqslant x^{1/2}\prod_{p\leqslant z}\left(1+\frac{\ell}{p}\right) \ll_\ell x^{1/2}(\log x)^\ell \ll_\ell x^{2/3}.$$

By Theorem 2.4 and the trivial lower bound $V(z) \gg_F (\log x)^{-\ell}$, we have

$$S(\mathcal{A}, z) \ll XV(z).$$

As before, the left side is at least as large as the count of $k \leqslant x$ such that each $F_i(k)$ is prime and $> z$ (and there are $O_F(z)$ values of k with each of $F_i(k)$ prime, and one of them is $\leqslant z$). It remains to bound $V(z)$ from above. To do this, we need the following two facts:

$$\rho_F(p) = \rho_{F_1}(p) + \cdots + \rho_{F_k}(p) \text{ for all but finitely many } p, \tag{3.9}$$

$$\sum_{p\leqslant y}\frac{\rho_{F_i}(p)}{p} = \log_2 y + C(F_i) + O_{F_i}\left(\frac{1}{\log y}\right) \quad (1\leqslant i\leqslant k) \tag{3.10}$$

where $C(F_i)$ is a constant depending on F_i. From (3.9) and (3.10) we deduce that

$$\log V(z) = \sum_{p\leqslant z}\log\left(1-\frac{\rho_F(p)}{p}\right) = O_F(1) - \sum_{p\leqslant z}\frac{\rho_F(p)}{p}$$
$$= O_F(1) - \sum_{p\leqslant z}\sum_{j=1}^{k}\frac{\rho_{F_j}(p)}{p} = -k\log_2 z + O_F(1),$$

and thus $V(z) \asymp_F (\log z)^{-k}$. In fact, (3.9) and (3.10) imply that (Ω) holds with $\kappa = k$ (this will be needed in the lower bound argument). In conclusion,

$$\#\{k\leqslant x : F_i(k) \text{ prime for every } i\} \leqslant O_F(z) + S(\mathcal{A}, z) \ll_F \frac{x}{(\log x)^k}.$$

Proof of (3.9): Suppose the equation in (3.9) does not hold. Then there are $i \neq j$ and some n with $p|F_i(n)$ and $p|F_j(n)$, i.e., $p|(F_i(n), F_j(n))$. As F_i and F_j are distinct, irreducible and have no fixed prime factor (in particular, neither is a multiple of the other), $(F_i, F_j) = 1$ over $\mathbb{Q}[x]$. Hence, there are $G, H \in \mathbb{Q}[x]$ such that $F_iG + F_jH = 1$. Clearing denominators gives $\widetilde{G}, \widetilde{H} \in \mathbb{Z}[x]$ with $F_i\widetilde{G} + F_j\widetilde{H} = C_{ij}$, where $C_{ij} \in \mathbb{Z}$. It follows that $p|C_{ij}$. As there are only finitely many pairs i, j, there are finitely many possible p.

Comments on (3.10): this is a consequence of the Prime Ideal Theorem of Landau [**101**]; see also [**23**, p. 33–38]. In some special cases, it follows from Mertens theorem for primes in arithmetic progressions, e.g. if $F(x) = x^2 + 1$, then

$$\sum_{p\leqslant x} \frac{\rho_F(p)}{p} = 1 + 2 \sum_{\substack{p\leqslant x \\ p\equiv 1 \pmod 4}} \frac{1}{p} = \log_2 x + O(1).$$

□

3.1. **Sieving limits.** One sees from Theorem 3.3 the effect of the dimension κ on the quality of the estimates. In order to ensure the remainder terms $\sum |r_d|$ are small, one needs $z^{f(\kappa)}$ to be smaller than the sieving limit for the sieve problem $\mathcal{A}$ (this is generally less than x), hence the range of permissible z is limited by the value of $f(\kappa)$.

Definition 2. *In the literature, the* sieve limit *(or sifting limit) $\beta(\kappa)$, for a given κ, is the infimum of exponents u so that for any sieve problem $\mathcal{A}$ satisfying (g), (r), and (Ω), we have for some positive c the bound*

$$S(\mathcal{A}, z) \geqslant cXV(z) - \sum_{d\leqslant z^u} |r_d|.$$

Some authors, e.g. Selberg, refer to $1/\beta(\kappa)$ as the *sieve limit*.

Theorem (3.3) implies that $\beta(\kappa) \leqslant f(\kappa) \leqslant \xi\kappa + O(\kappa^{1/2}\log(\kappa+3))$ for all $\kappa > 0$, where $\xi = 3.591\ldots$ is defined in Theorem 3.3 (iv) and that $\beta(\kappa) \leqslant 1 + o(1)$ as $\kappa \to 0^+$. When κ is large, this matches the asymptotic upper bound for $\beta(\kappa)$ achievable from the very complicated "β-sieve"; see [**61**, Ch. 11].

The exact value of $\beta(\kappa)$ is known only for $\kappa \in [0, 1/2] \cup \{1\}$, in these cases $\beta(\kappa) = 1$ for $\kappa \leqslant 1/2$ and $\beta(1) = 2$. See [**61**, Ch. 11] for details. The lower bound $\beta(1) \geqslant 2$ is clear from Selberg's examples in Section 1.7.4. The very complicated combinatorial sieves given in [**27**] give the best known upper bounds on $\beta(\kappa)$ for small κ (less than 10, say), see Table 1. Selberg [**136**, eq. (14.40)] showed that $\beta(\kappa) \leqslant 2\kappa + 0.4454$ for sufficiently large κ. The exact value of $\beta(\kappa)$ is unknown in all cases $\kappa > 1/2$ except for $\kappa = 1$.

κ	$\beta(\kappa) \leqslant$
0.5	1.0
1.0	2.0
2.0	4.2665
3.0	6.6409

TABLE 1. Known upper bounds on the sieving limit $\beta(\kappa)$

3.2. **The small z case. Asymptotic formula for the sifting function (Fundamental Lemma).** We combine the upper and lower bound sieve theorems to achieve an asymptotic formula for the sifting function when z is small compared with x. The bounds below are of the same strength, as s becomes large, as [**76**, Theorem 2.5], which is derived from a later, complicated sieve

developed by Brun. In particular, this gives the promised asymptotic formula in (1.4) as long as κ is fixed, $z = x^{o(1)}$ and the remainders $|r_d|$ are small enough on average.

Theorem 3.6 (Fundamental Lemma)**.** *(a) For any pair (z, D) of positive integers with $2 \leqslant z \leqslant D^{1/2}$, there are sieves $\boldsymbol{\lambda}^+$ and $\boldsymbol{\lambda}^-$ satisfying (λ^+) and (λ^-), respectively, satisfying $|\lambda_d^{\pm}| \leqslant 1$ for all d, with support in $\mathscr{D}(z, D) := \{d \in \mathbb{N} : \mu^2(d) = 1, P^+(d) \leqslant z, d \leqslant D\}$, and such that for any multiplicative function g satisfying (Ω), we have*

$$\sum_d \lambda_d^{\pm} g(d) = \left(1 + O\left(e^{-s\log s + s\log_3 s + O_{\kappa,B}(s)}\right)\right) \prod_{p \leqslant z}(1 - g(p)),$$

where $s = \max(100, \frac{\log D}{\log z})$.

(b) Let $\kappa_0 > 0$ and $B_0 > 0$. Assume $\mathcal{A}$ is a sieve problem, such that (g), (r) and that (Ω) holds for some constants $0 \leqslant \kappa \leqslant \kappa_0$ and $0 \leqslant B \leqslant B_0$. For any $2 \leqslant z \leqslant D^{1/2}$, we have

$$S(\mathcal{A}, z) = XV(z)\left(1 + O\left(e^{-s\log s + s\log_3 s + O_{\kappa_0,B_0}(s)}\right)\right) + \Delta \sum_{\substack{d \leqslant D \\ d \in \mathscr{P}(z)}} |r_d|,$$

where $s = \max(100, \frac{\log D}{\log z})$ and $|\Delta| \leqslant 1$.

Proof. Firstly, we observe that part (b) is immediate from part (a), the bound $|\lambda_d^{\pm}| \leqslant 1$ and (1.9). To prove (a), we take Brun-Hooley sieves, that is, sieves given in Theorems 2.3 (upper bound sieve) and 3.2 (lower bound sieve) and additionally supported on $\mathscr{D}(z, D)$. In light of (2.8), (2.11) and (3.4), it suffices to prove that there is a choice of parameters z_j and k_j such that

$$\begin{aligned} z_1^{k_1+1} z_2^{k_2} \cdots z_t^{k_t} &\leqslant D, \\ E &\ll e^{-s\log s + s\log_3 s + O_{\kappa_0,B_0}(s)}. \end{aligned} \tag{3.11}$$

As before, we will partition the primes in $[2, z]$ as in (2.9) and let $k_1, \dots, k_t$ be nonnegative even integers. Let $s_0(\kappa_0, B_0) \geqslant 100$ be sufficiently large. When $s \leqslant s_0(\kappa_0, B_0)$, we need only establish $E \ll_{\kappa_0,B_0} 1$. We'll take, similar to the proof of Theorem 2.4, the parameters $k_1 = 0$, $k_j = 2^{j-1}$ $(j \geqslant 2)$, and $z_j = z^{1/4^{j-1}}$ for each $j \geqslant 1$. Then $z_1^{k_1+1} z_2^{k_2} \cdots \leqslant z^2 \leqslant D$. As before, $L_j \leqslant L$ for all j, where L is a constant that depends only on κ_0, B_0. It then follows from (2.6) that $E \ll_{\kappa_0,B_0} 1$, and this implies (3.11) in this case.

When $s > s_0(\kappa_0, B_0)$ we take the parameters

$$\theta = \log s, \quad z_j = z^{1/\theta^{j-1}}, \qquad (1 \leqslant j \leqslant t),$$

where t is maximally chosen to satisfy (2.9), and

$$k_j = 2\left\lfloor \frac{s}{2}\left(1 - \frac{1}{\theta}\right)\right\rfloor - 6 + 2j \qquad (1 \leqslant j \leqslant t).$$

Note that $\theta \geqslant \log 100 > 4$ and $k_j \leqslant s(1-1/\theta) - 6 + 2j$ for all j. Thus, $z_1^{k_1+1} z_2^{k_2} \cdots \leqslant z^C$, where

$$\begin{aligned} C &= 1 + \sum_{j=1}^{\infty} \frac{2(1-1/\theta) - 6 + 2j}{\theta^{j-1}} \\ &= 1 + \frac{s(1-1/\theta) - 6}{1-1/\theta)} + \frac{2}{(1-1/\theta)^2} \\ &\leqslant s. \end{aligned}$$

Since $z^s = D$, this proves the first part of (3.11). In the notation (2.4), we have by (Ω) the bound

$$L_j \leqslant \kappa \log\theta + \frac{B}{\log z_{j+1}} \leqslant \kappa_0 \log\theta + 2B_0 =: L.$$

For large enough $s_0(\kappa_0, B_0)$, if $s > s_0(\kappa_0, B_0)$ then

$$k_1 \geqslant s(1-1/\theta) - 6 \geqslant 3s/4 - 6 \geqslant 2L.$$

Thus, using Stirling's formula, we find that

$$\begin{aligned} E &\leqslant e^L \sum_{j=1}^{t} \frac{L^{k_j+1}}{(k_j+1)!} \\ &\ll e^L \frac{L^{k_1+1}}{(k_1+1)!} \leqslant e^L \left(\frac{eL}{k_1+1}\right)^{k_1+1} \\ &\ll e^{O(B_0)} (\log s)^{\kappa_0} \exp\Big\{ -k_1 \log k_1 + O(k_1) + k_1 \log(\kappa_0 \log\theta + 2B_0) \Big\} \\ &\ll e^{-s\log s + s\log_3 s + O_{\kappa_0, B_0}(s)}. \end{aligned}$$

This gives (3.11) and completes the proof. □

3.2.1. *Application: Buchstab's function.* Buchstab's function

$$\Phi(x,z) = \#\{n \leqslant x : P^-(n) > z\}$$

is perhaps the most basic sieve function. Here we take $g(p) = 1/p$ for all p so that (Ω) holds with $\kappa = 1$.

Theorem 3.7 (No small prime factors)**.** *(i) Uniformly for $x \geqslant 0$ and $2 \leqslant z \leqslant y$, we have*

$$\Phi(x+y, z) - \Phi(x,z) \ll \frac{y}{\log z};$$

(ii) Uniformly for $x \geqslant 2z \geqslant 4$ we have

$$\Phi(x,z) \gg \frac{x}{\log z}.$$

(iii) Uniformly for $\exp\{(\log_2 x)^2\} \leqslant z \leqslant x$, we have

$$\Phi(x,z) = \big(1 + O(e^{-u\log u + O(u\log_3 u)})\big) x \prod_{p \leqslant z} (1 - 1/p),$$

where $u = \frac{\log x}{\log z}$.

Proof. For (i), we may suppose that $z \leqslant \sqrt{y}$, for the estimate with $z > \sqrt{y}$ follows from that for $z = \sqrt{y}$. We let $\mathcal{A} = (x, x+y] \cap \mathbb{N}$, $X = y$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$, $g(d) = 1/d$ and then (Ω) holds with $\kappa = 1$. Part (i) is then immediate from Theorem 2.4. For parts (ii) and (iii), let $\mathcal{A} = [1, x] \cap \mathbb{N}$, $X = x$, $\mathcal{A}_p = \{n \in \mathcal{A} : p|n\}$, $g(d) = 1/d$ and again (Ω) holds with $\kappa = 1$. Part (ii) follows from Theorem 3.3 when $z \leqslant x^{1/5}$ and x is large, since $f(1) < 5$ and thus $\sum_{d \leqslant z^{f(1)}} |r_d| \ll x^{4/5}$. When $x^{1/5} < z \leqslant x/2$ and x is large, say $x \geqslant x_0$, the Prime number theorem implies

$$\Phi(x, z) \geqslant \pi(x) - \pi(z) \gg \frac{x}{\log x} \gg \frac{x}{\log z}.$$

Finally, when $4 \leqslant x < x_0$, we have $\Phi(x, z) \geqslant 1 \gg x/\log z$. This proves (ii).

For part (iii), when $u \leqslant 10$ this follows from (i) and (ii). Now assume $u > 10$. take $D = \frac{x}{\log x} e^{-u \log u}$ in Theorem 3.6, and define s by $D = z^s$. Then the error terms satisfy

$$\sum_{d \leqslant D} |r_d| \leqslant D \ll e^{-u \log u} XV(z)$$

since $V(z) \asymp 1/\log z \gg 1/\log x$. By hypothesis, $u \leqslant \frac{\log x}{(\log_2 x)^2} \leqslant \frac{\log x}{\log^2 u}$ and hence

$$s = \frac{u}{\log x} (\log x - \log_2 x - u \log u) = u \left(1 - O\left(\frac{1}{\log u}\right)\right).$$

Inserting this into the conclusion of Theorem 3.6 completes the proof of (iii). □

3.3. **Small κ. The petite sieve.** In the previous sections, we showed an asymptotic formula for $S(\mathcal{A}, z)$ whenever κ, B are bounded and $\frac{\log D}{\log z} \to \infty$. In this section, we show an asymptotic formula for $S(\mathcal{A}, z)$ when $D = z^{O(1)}$ and $\kappa \to 0$. First, we mention a very general principle.

Definition 3. *Consider a sieve problem, and let p be a prime and $p > w \geqslant 2$. Then*

$$S(\mathcal{A}_p, w) := M\mathbb{P}\Big(\mathcal{A}_p \setminus \bigcup_{q \leqslant w} \mathcal{A}_q\Big).$$

This corresponds to $S(\widetilde{\mathcal{A}}, w)$, where $\widetilde{\mathcal{A}}$ is the sieve problem with space $(\widetilde{\mathcal{A}}, \mathcal{F}, \mathbb{P})$ inherited from $\mathcal{A}$, that is, $\widetilde{\mathcal{A}} = \mathcal{A}_p$, $\widetilde{\mathcal{A}}_d = \mathcal{A}_p \cap \mathcal{A}_d$, $\widetilde{X} = g(p)X$, $\tilde{g}(d) = g(d)$ and $\tilde{r}_d = r_{pd}$.

In the case where $\mathcal{A}$ is a finite set with uniform probability, $S(\mathcal{A}_p, w)$ counts those elements of $\mathcal{A}$ which have property $\mathcal{A}_p$ but not any of the properties $\mathcal{A}_q$ for $q \leqslant w$.

Lemma 3.8 (Buchstab's identity)**.** *Consider a sieve problem, and $z > w \geqslant 2$. Then*

$$S(\mathcal{A}, z) = S(\mathcal{A}, w) - \sum_{w < p \leqslant z} S(\mathcal{A}_p, p - 1). \tag{3.12}$$

Proof. Suppose that $\mathcal{A}_q$ does not occur for all primes $q \leqslant w$, but $\mathcal{A}_p$ does holds for some $p \in (w, z]$. Letting p be the largest such prime and summing over p yields the lemma. □

Theorem 3.9 (The petite sieve)**.** *Let $\mathcal{A}$ be a sieve problem and assume (g), (r) and define $V(z)$ by (V). Assume g satisfies (Ω) for some $\kappa \in (0, 1]$ and $B > 0$. Then, uniformly for $2 \leqslant s \leqslant \log z$ with $\kappa \log s \leqslant 1$ we have*

$$S(\mathcal{A}, z) = \left(1 + O_B\left(\kappa \log s + \frac{s}{\log z} + e^{-\frac{1}{2} s \log s}\right)\right) XV(z) + O\Big(\sum_{d \leqslant z^{1+1/s}} |r_d|\Big).$$

Proof. Let

$$w = z^{1/s}$$

so that $w \geqslant e > 2$. By Buchstab's identity (3.12) and the fact that $S(\mathcal{A}_p, z)$ is decreasing in z, we have

$$S(\mathcal{A}, w) \geqslant S(\mathcal{A}, z) \geqslant S(\mathcal{A}, w) - \sum_{w<p\leqslant z} S(\mathcal{A}_p, w).$$

By Theorem 3.6, with $\kappa_0 = 1$ and $B_0 = B$, we have

$$S(\mathcal{A}, w) = (1 + O_B(e^{-\frac{1}{2}s\log s}))XV(w) + \Delta \sum_{d\leqslant z} |r_d|$$

where $|\Delta| \leqslant 1$. By Theorem 2.4, again with $\kappa_0 = 1$ and $B_0 = B$ we have

$$S(\mathcal{A}_p, w) \leqslant O_B(Xg(p)V(w)) + \sum_{d\leqslant w} |r_{pd}|.$$

Summing over all $p \in (w, z]$ we use that

$$\begin{aligned}\sum_{w<p\leqslant z} g(p) &\leqslant -\sum_{w<p\leqslant z} \log(1-g(p)) \\ &\leqslant \kappa \log\left(\frac{\log z}{\log w}\right) + \frac{B}{\log w} \\ &= \kappa \log s + \frac{Bs}{\log z}.\end{aligned}$$

We conclude that

$$S(\mathcal{A}, z) = \left(1 + O_B\left(e^{-\frac{1}{2}s\log s} + \kappa\log s + \frac{s}{\log z}\right)\right) XV(w) + O\Big(\sum_{d\leqslant zw} |r_d|\Big). \tag{3.13}$$

Finally,

$$\begin{aligned}V(z) \leqslant V(w) = V(z)\frac{V(w)}{V(z)} &\leqslant V(z)\left(\frac{\log z}{\log w}\right)^{\kappa}(1 + O_B(1/\log w)) \\ &= V(z)s^{\kappa}(1 + O_B(1/\log w)) \\ &= V(z)(1 + O_B(\kappa\log s + s/\log z)).\end{aligned}$$

Inserting this into (3.13) completes the proof. □

Corollary 3.10. *Assume that $z \to \infty$, $\kappa \to 0$, and that the remainders satisfy*

$$\sum_{d\leqslant z^{1+\delta}} |r_d| = o(XV(z))$$

for some fixed $\delta > 0$. Then

$$S(\mathcal{A}, z) \sim XV(z)$$

Proof. Take $s = 2 + \min(\log(1/\kappa), \log_2 z)$ in Theorem 3.9. For large z, $s \geqslant 1/\delta$. By hypothesis, $s \to \infty$ and $\frac{s}{\log z} \to 0$, and $\kappa \log s \to 0$. □

3.4. **Exercises.**

Exercise 3.1. *Prove the lower bound* (3.8) *in Theorem 3.5.*

Exercise 3.2. *Show that for all sufficiently large even N, there is a prime p and a number q with $\Omega(q) \leqslant 6$ such that $N = p + q$.*

Exercise 3.3. Almost primes in short intervals. *Prove that for every $\delta > 0$, there is a natural number k so that whenever $x \geqslant x_0(\delta)$, then the interval $(x, x + x^\delta]$ contains an integer q with $\Omega(q) \leqslant k$.*

Exercise 3.4 $(*)$**.** *Let $\mathcal{S}$ denote the set of integers which are the sum of two squares*

(a) *Show that a positive proportion of integers have the form $a^2 + b^2 + 2^{c^2}$ for non-negative integers a, b, c. You may use as a fact that the counting function of $\mathcal{S}$ is $\gg x/\sqrt{\log x}$.*

(a) *Show that a positive proportion of integers* ***do not*** *have the form $a^2 + b^2 + 2^{c^2}$ for non-negative integers a, b, c.*

4. Selberg's sieve

For *any* real numbers $\lambda(d)$ for squarefree d, satisfying $\lambda(1)=1$, and for any natural number m, we have

$$\mathbb{1}_{m=1}=\sum_{d|m}\mu(d)\leqslant\Big(\sum_{d|m}\lambda(d)\Big)^2=\sum_{e|m}\sum_{[d_1,d_2]=e}\lambda(d_1)\lambda(d_2).$$

Thus, any choice of $\lambda(d)$ produces a valid upper bound sieve $\boldsymbol{\lambda}^+$ with

$$\lambda_e^+=\sum_{[d_1,d_2]=e}\lambda(d_1)\lambda(d_2).$$

Selberg calls this the Λ^2-sieve, terminology used also by Friedlander and Iwaniec [**61**]. Our goal is to optimize the choice of $(\lambda(d))_{d\geqslant 1}$. Let $\mathcal{A}$ be a sieve problem, and assume (g) and (r). Then

$$\begin{aligned}S(\mathcal{A},z)&\leqslant\sum_{\substack{d_1,d_2\\P^+(d_1d_2)\leqslant z}}\lambda(d_1)\lambda(d_2)A_{[d_1,d_2]}\\&\leqslant X\sum_{\substack{d_1,d_2\\P^+(d_1d_2)\leqslant z}}\lambda(d_1)\lambda(d_2)g([d_1,d_2])+\sum_{\substack{d_1,d_2\\P^+(d_1d_2)\leqslant z}}|\lambda(d_1)\lambda(d_2)r_{[d_1,d_2]}|\\&=:XG+R,\end{aligned}$$

say. Minimizing $XG+R$ is quite difficult. However, if we restrict the support of $\lambda(d)$ to $d\leqslant\sqrt{D}$, it is relatively easy to minimize G. Since primes with $g(p)=0$ do not contribute anything to G, we let

$$P=\prod_{\substack{p\leqslant z\\g(p)>0}}p$$

and restrict the support of $\lambda(d)$ to $d|P$. Define

$$h(d)=\prod_{p|d}\frac{g(p)}{1-g(p)}\qquad\text{for}\quad d|P \tag{4.1}$$

and $h(d)=0$ otherwise. Since $[d_1,d_2]=\frac{d_1d_2}{(d_1,d_2)}$ is squarefree, we have

$$g([d_1,d_2])=\frac{g(d_1)g(d_2)}{g((d_1,d_2))}.$$

Inverting (4.1) gives

$$\frac{1}{g(m)}=\prod_{p|m}\Big(1+\frac{1}{h(p)}\Big)=\sum_{d|m}\frac{1}{h(d)}\qquad(m|P),$$

so that

$$\begin{aligned}G&=\sum_{d_1|P,d_2|P}\lambda(d_1)\lambda(d_2)g(d_1)g(d_2)\sum_{d|(d_1,d_2)}\frac{1}{h(d)}\\&=\sum_{d|P}\frac{1}{h(d)}\sum_{\substack{d_1|P,d_2|P\\d|d_1,d|d_2}}\lambda(d_1)\lambda(d_2)g(d_1)g(d_2)=\sum_{d|P}\frac{1}{h(d)}\Big(\sum_{d|m}\lambda(m)g(m)\Big)^2.\end{aligned}$$

If we define

$$\xi(d) = \frac{\mu(d)}{h(d)} \sum_{d|m} \lambda(m) g(m) \qquad (d|P, d \leqslant \sqrt{D}), \tag{4.2}$$

then we have

$$G = \sum_d h(d) \xi(d)^2.$$

This quadratic form in the $\lambda(d)$ is minimized using Cauchy's inequality. First, we invert (4.2). Since

$$\begin{aligned} \sum_k h(kl)\xi(kl) &= \mu(l) \sum_k \mu(k) \sum_{kl|m} \lambda(m) g(m) \\ &= \mu(l) \sum_{l|m} \lambda(m) g(m) \sum_{k|m/l} \mu(k) = \mu(l)\lambda(l) g(l), \end{aligned}$$

and thus

$$\lambda(l) = \frac{\mu(l)}{g(l)} \sum_{l|d} h(d) \xi(d) \qquad (l|P). \tag{4.3}$$

In particular, by Cauchy's inequality,

$$1 = \lambda(1)^2 = \left(\sum_{d \leqslant \sqrt{D}} h(d)^{1/2} \xi(d) h(d)^{1/2} \right)^2 \leqslant \left(\sum_{d \leqslant \sqrt{D}} h(d) \xi(d)^2 \right) J = GJ, \quad J = \sum_{d \leqslant \sqrt{D}} h(d),$$

with equality if and only if $\xi(d)$ is constant; i.e., $\xi(d) = 1/J$ for every $d \leqslant \sqrt{D}, d|P$. By (4.3), this is attained by taking

$$\lambda(l) = \frac{\mu(l)}{Jg(l)} \sum_{\substack{l|d \\ d \leqslant \sqrt{D}}} h(d) = \frac{\mu(l) h(l)}{Jg(l)} \sum_{\substack{m \leqslant \sqrt{D}/l \\ (m,l)=1}} h(m). \tag{4.4}$$

We conclude that

$$S(\mathcal{A}, z) \leqslant \frac{X}{J} + R.$$

To handle R, we need an upper bound on $\lambda(l)$. Observe that for any $l \geqslant 1$,

$$\begin{aligned} J &= \sum_{k|l} \sum_{\substack{d \leqslant \sqrt{D} \\ (d,l)=k}} h(d) = \sum_{k|l} h(k) \sum_{\substack{m \leqslant \sqrt{D}/k \\ (m, l/k)=1}} h(m) \\ &\geqslant \sum_{k|l} h(k) \sum_{\substack{m \leqslant \sqrt{D}/l \\ (m,l)=1}} h(m) = \frac{h(l)}{g(l)} \sum_{\substack{m \leqslant \sqrt{D}/l \\ (m,l)=1}} h(m). \end{aligned}$$

Comparing with (4.4), we see that $|\lambda(l)| \leqslant 1$ for all l, and therefore that

$$R \leqslant \sum_{\substack{d_1, d_2 \leqslant \sqrt{D} \\ P^+(d_1 d_2) \leqslant z}} |r_{[d_1,d_2]}| \leqslant \sum_{\substack{d \leqslant D \\ P^+(d) \leqslant z}} 3^{\omega(d)} |r_d|. \tag{4.5}$$

Since $h(d) = 0$ for $d \nmid P$, we arrive at the following general theorem.

Theorem 4.1 (Selberg's sieve)**.** *Let $\mathcal{A}$ be a sieve problem, and assume (g), (r). Let $z \geqslant 2$ and $D \geqslant 1$. Then*

$$S(\mathcal{A}, z) \leqslant \frac{X}{J} + \sum_{\substack{d_1, d_2 \leqslant \sqrt{D} \\ P^+(d_1 d_2) \leqslant z}} |r_{[d_1,d_2]}| \leqslant \frac{X}{J} + \sum_{\substack{d \leqslant D \\ P^+(d) \leqslant z}} 3^{\omega(d)} |r_d|,$$

where $J = \sum_{n \leqslant \sqrt{D}} \mu^2(n) h(n)$, and h is the multiplicative function defined on primes $p \leqslant z$ by $h(p) = \frac{g(p)}{1-g(p)}$.

4.1. **Application. The Brun-Titchmarch inequality revisited.**

Theorem 4.2 (Brun-Titichmarch inequality, v.2)**.** *For $x \geqslant y \geqslant q \geqslant 1$ and $(a,q) = 1$, we have*

$$\pi(x; q, a) - \pi(x - y; q, a) \leqslant \frac{2y}{\phi(q) \log(5y/q)} \left(1 + O\left(\frac{\log_2(5y/q)}{\log(5y/q)}\right)\right).$$

We need the following auxilliary lemma.

Lemma 4.3. *For $k \in \mathbb{N}$ and $x \geqslant 1$, define*

$$H_k(x) := \sum_{\substack{d \leqslant x \\ (d,k)=1}} \frac{\mu^2(d)}{\phi(d)}.$$

Then

(i) *$H_k(x) \geqslant \frac{\phi(k)}{k} H_1(x)$ for all k, x;*
(ii) *$H_1(x) \geqslant \log x$ for $x \geqslant 1$.*

Proof. We first write

$$H_1(x) = \sum_{\ell \mid k} \sum_{\substack{d \leqslant x \\ (d,k)=\ell}} \frac{\mu^2(d)}{\phi(d)}.$$

Write $d = \ell h$ and observe that any nonzero summand corresponds to squarefree d, so $(h, \ell) = 1$ and $(h, k/\ell) = 1$. Thus, $(h, k) = 1$ and we have

$$H_1(x) = \sum_{\ell \mid k} \frac{\mu^2(\ell)}{\phi(\ell)} \sum_{\substack{h \leqslant x/\ell \\ (h,k)=1}} \frac{\mu^2(h)}{\phi(h)} = \sum_{\ell \mid k} \frac{\mu^2(\ell)}{\phi(\ell)} H_k(x/\ell) \leqslant H_k(x) \sum_{\ell \mid k} \frac{\mu^2(\ell)}{\phi(\ell)}.$$

Inequality (i) follows from the identity

$$\sum_{\ell \mid k} \frac{\mu^2(\ell)}{\phi(\ell)} = \prod_{p \mid k} \left(1 + \frac{1}{p-1}\right) = \frac{k}{\phi(k)}.$$

Next, define $s(d) = \prod_{p \mid d} p$ to be the squarefree kernel of d. Then

$$\sum_{s(d)=m} \frac{1}{d} = \prod_{p \mid m} \left(\frac{1}{p} + \frac{1}{p^2} + \cdots\right) = \frac{\mu^2(m)}{\phi(m)}.$$

Hence, for $x \geqslant 2$ we have

$$H_1(x) = \sum_{m \leqslant x} \sum_{s(h)=m} \frac{1}{h} = \sum_{s(h) \leqslant x} \frac{1}{h} \geqslant \sum_{h \leqslant x} \frac{1}{h} \geqslant \int_1^{\lfloor x+1 \rfloor} \frac{dt}{t} = \log \lfloor x+1 \rfloor \geqslant \log x. \qquad \square$$

Proof of Theorem 4.2. As with the proof of Theorem 2.7, let

$$\mathcal{A} = \{x - y < n \leqslant x : n \equiv a \pmod q\}, \;\; X = y/q,$$

and let $g(p) = 1/p$ for $p \nmid q$ and $g(p) = 0$ for $p|q$. Then

$$A_d = g(d)X + r_d, \quad |r_d| \leqslant 1$$

Also

$$S(\mathcal{A}, z) \geqslant \pi(x; q, a) - \pi(x - y; q, a) - z.$$

In the notation of Theorem 4.1, we have

$$h(d) = \frac{\mu^2(d)}{\phi(d)} \quad \text{for } (d, q) = 1.$$

Let $z = \sqrt{D}$. By Lemma 4.3,

$$J = H_q(\sqrt{D}) \geqslant \frac{\phi(q)}{2q} \log D.$$

By Theorem 4.1,

$$S(\mathcal{A}, z) \leqslant \frac{X}{J} + \sum_{d_1, d_2 \leqslant \sqrt{D}} 1 \leqslant \frac{X}{J} + D \leqslant \frac{2y}{\phi(q) \log D} + D,$$

hence

$$\pi(x; q, a) - \pi(x - y; q, a) \leqslant \frac{2y}{\phi(q) \log D} + D + \sqrt{D}.$$

A near-optimal choice for D is

$$D = \max\left(2, \frac{5y/q}{\log^2(5y/q)}\right),$$

and this gives the theorem upon using that

$$\frac{1}{\log D} = \frac{1}{\log(5y/q)} \left(1 + O\left(\frac{\log_2(5y/q)}{\log(5y/q)}\right)\right).$$

$\square$

4.2. **Selberg's sieve is best possible in dimension 1.** Consider Selberg's example

$$\mathcal{A} = \{n \leqslant x : \lambda(n) = -1\},$$

where $\lambda(n) = (-1)^{\Omega(n)}$ is Liouville's function, and set $X = x/2$. As we saw previously, by the Prime Number Theorem, for some constant $c > 0$,

$$A_d = \sum_{d \leqslant x} \frac{1 - \lambda(d)}{2} = \frac{x/2}{d} + O\left(\frac{x}{d} e^{-c\sqrt{\log(x/d)}}\right).$$

Let $g(d) = 1/d$ for $d \in \mathscr{P}(z)$, and thus $h(p) = 1/(p-1)$ for $p \leqslant z$. Take $z = \sqrt{D} = x^{1/2-\varepsilon(x)}$ for $\varepsilon(x) = 1/\log\log x$. Then

$$\sum_{d\leqslant D} \mu^2(d)3^{\omega(d)}|r_d| \ll xe^{-c\sqrt{2\varepsilon(x)\log x}} \sum_{d\leqslant D} \frac{\mu^2(d)3^{\omega(d)}}{d}$$
$$\ll \frac{x}{\log^{10} x} \prod_{p\leqslant x}(1+3/p) \ll \frac{x}{\log^7 x}.$$

By Lemma 4.3,

$$J = H_1(\sqrt{D}) \geqslant \tfrac{1}{2}\log D,$$

and thus, by Theorem 4.1,

$$S(\mathcal{A}, z) \leqslant \frac{X}{J} + O\left(\frac{x}{\log^7 x}\right) = \frac{x}{(1-2\varepsilon(x))\log x} + O\left(\frac{x}{\log^2 x}\right) \sim \frac{x}{\log x}$$

as $x \to \infty$. But $S(\mathcal{A}, z) = \pi(x) - \pi(\sqrt{x}) + 1 \sim x/\log x$ by the prime number theorem. Thus, Selberg's upper bound sieve cannot be improved, for general sieve problems.

4.3. **Asymptotic formulas.** It is possible, under two-sided bounds on $g(p)$, to obtain an asymptotic formula for J. The theorem below illustrates some useful techniques for analyzing sums of multiplicative functions.

Theorem 4.4. *Let $A_1, A_2, L, \kappa > 0$, and let g be a multiplicative function supported on squarefree integers satisfying*

$$0 \leqslant g(p) \leqslant 1 - A_1 \qquad (p \text{ prime}) \tag{4.6}$$

and

$$-L \leqslant \sum_{w\leqslant p\leqslant y} g(p)\log p - \kappa \log\frac{y}{w} \leqslant A_2 \qquad (2 \leqslant w \leqslant y). \tag{4.7}$$

Let h be the multiplicative function supported on squarefree integers and defined by

$$h(p) = \frac{g(p)}{1-g(p)} \qquad (p \text{ prime}).$$

Then, uniformly for $x \geqslant 2$,

$$\sum_{n\leqslant x} \mu^2(n)h(n) = \mathfrak{S}(g)\frac{(\log x)^\kappa}{\Gamma(\kappa+1)} + O_{\kappa,A_1,A_2}\left(\mathfrak{S}(g)\left[(L+1)(\log x)^{\kappa-1} + (L+1)^\kappa\right]\right),$$

where

$$\mathfrak{S}(g) = \prod_p (1-g(p))^{-1}(1-1/p)^\kappa.$$

Remarks. It is important that the dependence on L be made explicit in the error term, as in many applications L is not uniformly bounded, whereas κ, A_1, A_2 are usually uniformly bounded. That is, we may wish to apply the result with a sequence of functions $g_1, g_2, \ldots$, and typically the same κ, A_1, A_2 will work for every g_i but the corresponding numbers L (call them L_i, say) may not be bounded; see Theorem TP2 below for an example.

The sum on the left side depends only on $g(p)$ for primes $p \leqslant x$, whereas the singular series is defined in terms of $g(p)$ for all primes p. It might seem more natural to replace the infinite product

defining $\mathfrak{S}(g)$ with a finite product over primes $p \leqslant x$, but the given form of the theorem is easier to apply in practice. The inclusion of "extraneous factors" in $\mathfrak{S}(g)$ does force us to include the additional error term $\mathfrak{S}(g)(L+1)^\kappa$, which is in fact best possible as the following example shows: fix $\kappa > 0$, consider a large L (say going to ∞ as $x \to \infty$), $g(p) = 0$ for $p < e^{L/2\kappa}$ and $g(p) = \kappa/p$ for $p \geqslant e^{L/2\kappa}$. One easily verifies (4.6) and (4.7), and computes that $\mathfrak{S}(g) \ll L^{-\kappa}$. If $1 < x < e^{o(L)}$, then the sum in the lemma is just $h(1) = 1$, while the "main term" and first error term on the right side are both $o(1)$.

The second error term $\mathfrak{S}(g)(L+1)^\kappa$ is frequently missing in versions of Theorem gh appearing in the literature, e.g. some of the recent work on prime gaps [**66**, Lemma 4], [**108**, Lemma 6.1]. These two works ultimately depend on [**76**, Lemma 5.3 (2.5)], which claims, in our notation, that

$$\prod_{p \leqslant z} (1 - g(p)) = \mathfrak{S}(g)^{-1} \frac{e^{-\gamma\kappa}}{(\log z)^\kappa} \left(1 + O\left(\frac{L+1}{\log z}\right)\right)$$

for *all* $z \geqslant 2$. The above example clearly shows this claim to be false in some cases when $2 \leqslant z < e^{o(L)}$. The error in the proof can be traced to [**76**, p. 146], two lines after (2.7), where it is written that $(1 + O(L/\log z))^{-1} = (1 + O(L/\log z))$, a relation which is only true when $z > e^{CL}$ for some positive constant C.

Proof. We will use a technique due to E. Wirsing. Assume throughout that g and h are supported on squarefree integers, so that all sums are over squarefree integers, and let

$$H(x) = \sum_{n \leqslant x} h(n).$$

By partial summation,

$$H(x) \log x = \sum_{n \leqslant x} h(n) \left(\log n + \log \frac{x}{n}\right) = \sum_{n \leqslant x} h(n) \log n + \int_1^x \frac{H(u)}{u}\, du.$$

We rewrite the right hand sum as

$$\sum_{n \leqslant x} h(n) \log n = \sum_{n \leqslant x} h(n) \sum_{p \leqslant x} \log p = \sum_{p \leqslant x} \log p \sum_{\substack{m \leqslant x/p \\ p \nmid m}} h(p)h(m).$$

Using the relation $h(p) = g(p) + g(p)h(p)$, the right side transforms into

$$= \sum_{p \leqslant x} g(p) \log p \left\{ \sum_{m \leqslant x/p} h(m) - \sum_{\substack{m \leqslant x/p \\ p \mid m}} h(m) + h(p) \sum_{\substack{m \leqslant x/p \\ p \nmid m}} h(m) \right\}$$

$$= \sum_{p \leqslant x} g(p) \log p \left\{ \sum_{m \leqslant x/p} h(m) - h(p) \sum_{\substack{l \leqslant x/p^2 \\ p \nmid l}} h(l) + h(p) \sum_{\substack{m \leqslant x/p \\ p \nmid m}} h(m) \right\}$$

$$= \sum_{p \leqslant x} g(p) \log p \left\{ \sum_{m \leqslant x/p} h(m) + h(p) \sum_{\substack{x/p^2 < m \leqslant x/p \\ p \nmid m}} h(m) \right\}.$$

Interchanging the order of summation yields

$$\sum_{n\leqslant x} h(n)\log n = \sum_{m\leqslant x} h(m)\Big\{ \sum_{p\leqslant x/m} g(p)\log p + \sum_{\substack{\sqrt{x/m}<p\leqslant x/m\\ p\nmid m}} g(p)h(p)\log p\Big\}. \tag{4.8}$$

Applying (4.7) (with $y=w=p$) followed by (4.6), we see that for every p,

$$g(p)\leqslant \frac{A_2}{\log p} \quad \text{and} \quad h(p)\leqslant \frac{A_2/\log p}{A_1}.$$

Hence, by another application of (4.7), we obtain

$$\sum_{\sqrt{x/m}\leqslant p\leqslant x/m} g(p)h(p)\log p \leqslant \frac{A_2/A_1}{\log\sqrt{x/m}} \sum_{\sqrt{x/m}\leqslant p\leqslant x/m} g(p)\log p \ll 1 \qquad (m\leqslant x/4),$$

the sum being trivially $O(1)$ when $m>x/4$ (here and throughout he proof, constants implied by O-symbols may depend on κ, A_1, A_2 but not on any other parameter). By (4.7),

$$\sum_{p\leqslant x/m} g(p)\log p = \kappa\log\frac{x}{m} + O(L+1),$$

and thus relation (4.8) becomes

$$\sum_{n\leqslant x} h(n)\log n = \sum_{m\leqslant x} h(m)\Big\{\kappa\log\frac{x}{m}+O(L+1)\Big\} = \kappa\int_1^x \frac{H(u)}{u}\,du + O((L+1)H(x)).$$

Hence

$$H(x)\log x = (\kappa+1)\int_1^x \frac{H(u)}{u}\,du + \Delta^*(x), \quad \Delta^*(x)\ll (L+1)H(x).$$

It is convenient to slightly alter this relation, using the fact that $H(u)=1$ for $1\leqslant u<2$. Thus

$$H(x)\log x = (\kappa+1)\int_2^x \frac{H(u)}{u}\,du + \Delta(x) \quad (x\geqslant 2), \quad \Delta(x)\ll (L+1)H(x). \tag{4.9}$$

We treat $\Delta(x)$ like an error term, and consider the integral equation for H given by (4.9). The "unperturbed" equation $f(x)\log x = (\kappa+1)\int_2^x f(u)/u\,du$ has general solution $f(t)=C(\log t)^\kappa$ for some constant C, and our aim is to prove an asymptotic of this shape for $H(x)$. First, we bound $\Delta(x)$ using the crude estimate

$$H(x)\leqslant \prod_{p\leqslant x}(1+h(p)) = S(x)\prod_{p\leqslant x}\left(1-\frac{1}{p}\right)^{-\kappa} \ll (\log x)^\kappa \mathfrak{S}(x), \tag{4.10}$$

where

$$S(x) = \prod_{p\leqslant x}(1-g(p))^{-1}(1-1/p)^\kappa.$$

We must relate $S(x)$ to $\mathfrak{S}(g)$ by estimating the contribution of the large primes $p>x$ to $\mathfrak{S}(g)$; that is, the "extraneous" primes that are not involved in the sum $\sum_{n\leqslant x} h(n)$ (cf. the Remarks). Let $x\geqslant 2$. Now

$$\log(S(x)/\mathfrak{S}(g)) = \log\prod_{p>x}(1-g(p))(1-1/p)^{-\kappa} \leqslant \sum_{p>x}\left(\frac{\kappa}{p}-g(p)+\frac{\kappa}{p^2}\right).$$

By (4.7) and partial summation, when $x \geqslant e^{L+1}$ we get

$$\sum_{p>x} \frac{\kappa}{p} - g(p) = \sum_{p>x} \frac{\frac{\kappa \log p}{p} - g(p)\log p}{\log p} = \int_x^\infty \sum_{x<p\leqslant t} \left(\frac{\kappa \log p}{p} - g(p)\log p\right) \frac{dt}{t\log^2 t} \ll \frac{L+1}{\log x} \ll 1.$$

When $x < e^{L+1}$, however, we do better using the trivial observation $g(p) \geqslant 0$:

$$\sum_{p>x} \frac{\kappa}{p} - g(p) \leqslant \sum_{x<p\leqslant e^{L+1}} \frac{\kappa}{p} + \sum_{p>e^{L+1}} \frac{\kappa}{p} - g(p) = \kappa \log\left(\frac{L+1}{\log x}\right) + O(1).$$

Therefore we have

$$S(x) = \mathfrak{S}(g)\frac{S(x)}{\mathfrak{S}(g)} \ll \left(1 + \left(\frac{L+1}{\log x}\right)^\kappa\right)\mathfrak{S}(g). \tag{4.11}$$

The example described in the Remarks shows that (4.11) is in fact best possible. Combining (4.11) with (4.10) yields the estimates

$$H(x) \ll ((\log x)^\kappa + (L+1)^\kappa)\,\mathfrak{S}(g) \qquad (x \geqslant 2), \tag{4.12}$$

$$\Delta(x) \ll \big((L+1)(\log x)^\kappa + (L+1)^{\kappa+1}\big)\,\mathfrak{S}(g) \qquad (x \geqslant 2). \tag{4.13}$$

Returning to (4.9), we replace x with t, divide through by $t(\log t)^{\kappa+2}$ and integrate from 2 to x. This gives

$$\int_2^x \frac{H(t)}{t(\log t)^{\kappa+1}}\,dt - \int_2^x \frac{\kappa+1}{t(\log t)^{\kappa+2}} \int_2^t \frac{H(u)}{u}\,du = \int_2^x \frac{\Delta(t)}{t(\log t)^{\kappa+2}}\,dt.$$

After interchanging the order of integration, we see that the double integral equals

$$\int_2^x \frac{H(u)}{u}\left(\frac{1}{(\log u)^{\kappa+1}} - \frac{1}{(\log x)^{\kappa+1}}\right)du,$$

which implies that

$$\int_2^x \frac{H(u)}{u}\,du = (\log x)^{\kappa+1}\int_2^x \frac{\Delta(t)}{t(\log t)^{\kappa+2}}\,dt.$$

Comparing this with (4.9) yields

$$H(x) = (\kappa+1)(\log x)^\kappa \int_2^x \frac{\Delta(t)}{t(\log t)^{\kappa+2}}\,dt + \frac{\Delta(x)}{\log x} \qquad (x \geqslant 2). \tag{4.14}$$

By (4.13), the integral above converges as $x \to \infty$ (the tail being $\ll (L+1)\mathfrak{S}(g)/\log x$ for $x > e^{L+1}$), thus

$$\int_2^\infty \frac{\Delta(t)}{t(\log t)^{\kappa+2}}\,dt = C$$

for some constant C. Therefore, combining (4.14) and (4.13) (when $x \geqslant e^{L+1}$) and (4.12) (when $x < e^{L+1}$), we conclude that

$$H(x) = C(\kappa+1)(\log x)^\kappa + O\left(\mathfrak{S}(g)\left[(L+1)(\log x)^{\kappa-1} + (L+1)^\kappa\right]\right) \qquad (x \geqslant 2). \tag{4.15}$$

It remains to determine C. For real $s > 0$ let

$$\zeta_h(s) := \sum_{m=1}^\infty h(m)m^{-s} = s\int_1^\infty H(x)x^{-s-1}\,dx = s\int_0^\infty H(e^t)e^{-st}\,dt.$$

By (4.15), as $s \to 0^+$ we have

$$\zeta_h(s) \sim sC(\kappa+1)\int_0^\infty t^\kappa e^{-st}\,dt = s^{-\kappa}C(\kappa+1)\int_0^\infty u^\kappa e^{-u}\,du = s^{-\kappa}C(\kappa+1)\Gamma(\kappa+1).$$

Since $\zeta(s+1) \sim \frac{1}{s}$ as $s \to 0^+$, we obtain

$$\lim_{s\to 0^+} \zeta(s+1)^{-\kappa}\zeta_h(s) = C(\kappa+1)\Gamma(\kappa+1).$$

On the other hand,

$$\zeta(s+1)^{-\kappa}\zeta_h(s) = \prod_p \left(1-\frac{1}{p^{s+1}}\right)^\kappa \left(1+\frac{h(p)}{p^s}\right).$$

Using (4.7), we see that the Euler product is uniformly convergent for $s \geqslant 0$ (Exercise). Therefore,

$$C(\kappa+1)\Gamma(\kappa+1) = \lim_{s\to 0^+}\prod_p \left(1-\frac{1}{p^{s+1}}\right)^\kappa \left(1+\frac{h(p)}{p^s}\right) = \prod_p \left(1-\frac{1}{p}\right)^\kappa (1+h(p)) = \mathfrak{S}(g). \quad \square.$$

4.4. **Application: twin primes.**

Theorem 4.5. *Let k be an even, positive integer. Then, uniformly in $k \leqslant x$,*

$$\#\{p \leqslant x : p+k \text{ prime}\} \leqslant C\prod_{\substack{p\mid k\\ p>2}} \left(\frac{p-1}{p-2}\right) \frac{x}{\log^2 x}\left(4+O\left(\frac{\log_2 x}{\log x}\right)\right),$$

where

$$C = 2\prod_p \left(1-\frac{1}{(p-1)^2}\right)$$

is the "twin prime constant".

Proof. Let $\mathcal{A} = \{p+k : p \leqslant x\}$. Then

$$A_d = Xg(d) + r_d, \quad X = \mathrm{li}(x),\ g(d) = \begin{cases} 1/\phi(d) & \text{if } (d,k)=1 \\ 0 & \text{if } (d,k)>1. \end{cases}$$

Let $D = z = x^{1/2}(\log x)^{-B}$, where B is sufficiently large. We have $|r_d| \ll x/d$ trivially, and therefore by the Bombieri-Vinogradov theorem and Cauchy's inequality,

$$\begin{aligned}\sum_{d\leqslant D} 3^{\omega(d)}\mu^2(d)|r_d| &\ll \sum_{d\leqslant D}\mu^2(d)3^{\omega(d)}\left(\frac{x}{d}\right)^{1/2}|r_d|^{1/2} \\ &\leqslant x^{1/2}\left(\sum_{d\leqslant D}\frac{\mu^2(d)9^{\omega(d)}}{d}\right)^{1/2}\left(\sum_{d\leqslant D}|r_d|\right)^{1/2} \\ &\ll x^{1/2}\left(\prod_{p\leqslant D}\left(1+\frac{9}{p}\right)\right)^{1/2}\left(\frac{x}{(\log x)^{20}}\right)^{1/2} \ll \frac{x}{(\log x)^5}.\end{aligned}$$

Since $g(p) = \frac{1}{p-1}$ for primes $p \nmid k$, the hypotheses of Theorem 4.4 hold with $\kappa = 1$, A_1 and A_2 absolute constants, and with

$$L = O(1) + \sum_{p\mid k}\frac{\log p}{p-1} \ll \sum_{p=O(\log k)}\frac{\log p}{p-1} \ll \log_2 k.$$

We also have $h(p) = \frac{1}{p-2}$ for $p \nmid k$. Theorem 4.4 implies that

$$J = \mathfrak{S}(g)\left(\log\sqrt{D} + O(\log_2 k)\right) = \mathfrak{S}(g)\left(\tfrac{1}{4}\log x + O(\log_2 x)\right),$$

where

$$\mathfrak{S}(g) = \frac{1}{2C}\prod_{\substack{p\mid k\\ p>2}}\frac{p-2}{p-1}.$$

Finally, Theorem 4.1 gives

$$\begin{aligned}\#\{p \leqslant x : p+k \text{ prime}\} &\leqslant D + S(\mathcal{A}, D) \leqslant D + \frac{X}{J} + \sum_{d\leqslant D}\mu^2(d)3^{\omega(d)}|r_d| \\ &\leqslant C\prod_{\substack{p\mid k\\ p>2}}\left(\frac{p-1}{p-2}\right)\frac{x}{\log^2 x}\left(4 + O\left(\frac{\log_2 x}{\log x}\right)\right). \qquad \square\end{aligned}$$

Remarks. Assuming the Elliott-Halberstam conjecture, we may take $D = x^{1-\varepsilon}$ for any positive ε in the above argument, and replace the 4 with $2 + o(1)$ in the conclusion. This is only a factor two worse than the Hardy-Littlewood conjectured asymptotic formula for the left hand side.

For *fixed* k, Bombieri, Friedlander and Iwaniec showed that one may replace 4 with 3.5 (the error term is not uniform in k).

Using the very same analysis, one can also prove, for even $N \geqslant 4$,

$$\#\{p_1, p_2 : N = p_1 + p_2\} \leqslant C\prod_{\substack{p\mid N\\ p>2}}\frac{p-1}{p-2}\frac{N}{\log^2 N}\left(4 + O\left(\frac{\log_2 N}{\log N}\right)\right).$$

We leave this as an exercise.

4.5. **Exercises.**

Exercise 4.1. *As in the proof of Lemma 4.3, let*

$$H_1(x) = \sum_{n\leqslant x}\frac{\mu^2(n)}{\phi(n)}.$$

Show that $H_1(x) = \log x + c + o(1)$, *where*

$$c = \gamma + \sum_p \frac{\log p}{p(p-1)} = 1.332\ldots$$

and $\gamma = 0.5772\ldots$ *is Euler's constant.*

Exercise 4.2. *Prove that for even* $N \geqslant 4$,

$$\#\{p_1, p_2 : N = p_1 + p_2\} \leqslant C\prod_{\substack{p\mid N\\ p>2}}\frac{p-1}{p-2}\frac{N}{\log^2 N}\left(4 + O\left(\frac{\log_2 N}{\log N}\right)\right).$$

Exercise 4.3. *(Asymptotic for J in Selberg's sieve). Let $A_1, A_2, L, \kappa > 0$, and let g be a multiplicative function satisfying* (4.6) *and* (4.7). *Let h be the multiplicative function defined by $h(p) = \frac{g(p)}{1-g(p)}$ for prime p.*

(i) Prove that uniformly for $2 \leqslant w \leqslant y$ and $s \geqslant 0$,

$$\prod_{w\leqslant p\leqslant y} \left(1 - \frac{1}{p^{s+1}}\right)^{\kappa} \left(1 + \frac{h(p)}{p^s}\right) = 1 + O\left(\frac{L+1}{\log w}\right).$$

The implied constant in the $O-$term may depend only on A_1, A_2, κ. Hint: be careful. You may want to consider separately the cases $w < e^{L+1}$ and $w \geqslant e^{L+1}$.

(ii) Use (i) to show that

$$\lim_{s\to 0^+} \prod_p \left(1 - \frac{1}{p^{s+1}}\right)^{\kappa} \left(1 + \frac{h(p)}{p^s}\right) = \prod_p \left(1 - \frac{1}{p}\right)^{\kappa} (1 + h(p)).$$

Exercise 4.4. *We say that a tuple $(h_1, \dots, h_k)$ of integers is* admissible *if the set of linear forms $(n+h_1, \dots, n+h_k)$ is admissible. For each $x \geqslant 2$, let $\rho^*(x)$ be the maximum size k of an admissible k-tuple lying in $[0, x]$. For example, $\rho^*(12) = 5$ by taking $0, 2, 6, 8, 12$ Show that*

$$(1 + o(1))x/\log x \leqslant \rho^*(x) \leqslant (2 + o(1))x/\log x.$$

5. Smooth numbers

In what follows we will need the Erdős notation for iterates of the logarithm:

$$\log_k x = \underbrace{\log \cdots \log}_{k} x.$$

The basic function

$$\Psi(x,y) = \#\{n \leqslant x : P^+(n) \leqslant y\}$$

is widely used in number theory. In contrast with $\Phi(x,z)$, which eliminates numbers with small prime factors, $\Psi(x,y)$ eliminates numbers with large prime factors. One can use the small sieve (Theorem 2.4) to bound $\Psi(x,y)$ but we obtain only

$$\Psi(x,y) \ll x\frac{\log y}{\log x}$$

with no lower bound. The truth, however, is very different, owing to the non-independence of large primes dividing n, e.g. the events $p|n$ and $p'|n$ are exclusive if $p > p' > \sqrt{x}$.

Theorem 5.1. *Uniformly for $x \geqslant 10$ and $\log x \leqslant y \leqslant x$ we have*

$$\Psi(x,y) = xe^{-u\log u + O(u\log_2(10u))}, \qquad u = \frac{\log x}{\log y}.$$

Remarks. There is a change of behavior around $y = \log x$, due to the fact that for smaller y, $\prod_{p\leqslant y} p < x$, and this forces at least some of the exponents of primes dividing n to be large.

We note some special cases which we will find useful for applications:

$$\Psi(x,\log x) = \exp\left\{O\left(\frac{(\log x)\log_3 x}{\log_2 x}\right)\right\} = x^{o(1)} \quad (x\to\infty), \tag{5.1}$$

$$\Psi(x,(\log x)^c) = x^{1-1/c+o(1)} \qquad (x\to\infty) \tag{5.2}$$

for any fixed $c \geqslant 1$, and

$$\Psi(x, x^{c(\log_3 x)/\log_2 x}) = \frac{x}{(\log x)^{c+o(1)}} \quad (x\to\infty). \tag{5.3}$$

We first need a combinatorial lemma, which is similar to devices we used to develop the Brun-Hooley sieve.

Lemma 5.2. *Let I be a finite set of positive integers, and $k \in \mathbb{N}$. Then*

$$\sum_{\substack{n_1,\dots,n_k\in I\\ n_1<n_2<\cdots<n_k}} \frac{1}{n_1\cdots n_k} \geqslant \frac{1}{k!}\left(\sum_{n\in I}\frac{1}{n} - \frac{k-1}{\min I}\right)^k,$$

provided that the expression in parentheses is $\geqslant 0$.

Proof. This is straightforward. We have

$$\sum_{\substack{n_1,\dots,n_k\in I\\ n_1<n_2<\cdots<n_k}} \frac{1}{n_1\cdots n_k} = \frac{1}{k!}\sum_{\substack{n_1,\dots,n_k\in I\\ \text{distinct}}} \frac{1}{n_1\cdots n_k}$$
$$= \frac{1}{k!}\sum_{n_1\in I}\frac{1}{n_1}\sum_{\substack{n_2\in I\\ n_2\neq n_1}}\frac{1}{n_2}\cdots \sum_{\substack{n_k\in I\\ n_k\notin\{n_1,\dots,n_{k-1}\}}}\frac{1}{n_k}.$$

Each sum over n_i is at least $\sum_{n\in I} 1/n - (i-1)/\min I$, independent of the choice of $n_1,\dots,n_{i-1}$, and the proof is complete. □

Proof of Theorem 5.1. We begin with the upper bound, and we will in fact prove a stronger bound

$$\Psi(x,y) \leqslant xe^{-u\log u+O(u)}. \tag{5.4}$$

Define

$$\alpha = 1-\frac{\log u}{\log y}.$$

By our hypothesis that $\log x\leqslant y\leqslant x$,

$$1\leqslant u\leqslant \frac{\log x}{\log_2 x},\qquad \frac{\log_3 x}{\log_2 x}\leqslant \alpha\leqslant 1. \tag{5.5}$$

For all $w>0$, $\log(w^\alpha)\leqslant w^\alpha$ and thus $\log w\leqslant \alpha^{-1}w^\alpha$. Hence,

$$(\log x)\Psi(x,y) = \sum_{\substack{n\leqslant x\\ P^+(n)\leqslant y}} \log(x/n)+\log n \leqslant \alpha^{-1}x^\alpha S + \sum_{\substack{n\leqslant x\\ P^+(n)\leqslant y}}\sum_{p^a\mid n}\log p, \tag{5.6}$$

where

$$S=\sum_{P^+(n)\leqslant y}\frac{1}{n^\alpha}.$$

In the double-sum on the right side of (5.6), let $n=p^am$, and separate into cases depending on $a=1$ or $a>1$. The $a=1$ terms contribute

$$\leqslant \sum_{\substack{m\leqslant x\\ P^+(m)\leqslant y}}\sum_{p\leqslant \min(y,x/m)}\log p \ll \sum_{\substack{m\leqslant x\\ P^+(m)\leqslant y}}\min(y,x/m)\leqslant \sum_{P^+(m)\leqslant y} y^{1-\alpha}\left(\frac{x}{m}\right)^\alpha = ux^\alpha S.$$

The terms with $a>1$ contribute

$$\leqslant \sum_{\substack{p\leqslant y\\ 2\leqslant a\ll\log x}}\log p\sum_{\substack{m\leqslant x/p^a\\ P^+(m)\leqslant y}} 1 \leqslant \sum_{\substack{p\leqslant y\\ 2\leqslant a\ll\log x}}\log p\sum_{P^+(m)\leqslant y}\left(\frac{x}{p^am}\right)^\alpha = x^\alpha ST,$$

where

$$T=\sum_{\substack{p\leqslant y\\ 2\leqslant a\ll\log x}}\frac{\log p}{p^{a\,\alpha}}.$$

If $\alpha\geqslant 2/3$ then clearly $T\ll 1$. If $0<\alpha<2/3$ then $y\leqslant(\log x)^3$ and $u\geqslant \frac{\log x}{3\log_2 x}$, thus crudely

$$T\ll(\log x)\sum_{p\leqslant y}\log p\ll y\log x\ll(\log x)^4\ll u^5.$$

Putting together the $a = 1$ and $a > 1$ terms, we conclude from (5.5) and (5.6) that

$$(\log x)\Psi(x,y) \ll x^\alpha S(\alpha^{-1} + u + u^5) \ll u^5 x^\alpha S = xu^5 e^{-u\log u} S, \tag{5.7}$$

since $\alpha^{-1} = \frac{\log y}{\log(y/u)} \ll \log u$. It remains to bound S. We have

$$S = \prod_{p\leqslant y}\left(1 + \frac{1}{p^\alpha} + \frac{1}{p^{2\alpha}} + \cdots\right) = \prod_{p\leqslant y}\left(1 + \frac{1}{p^\alpha - 1}\right) \leqslant \exp\left\{\sum_{p\leqslant y}\frac{1}{p^\alpha - 1}\right\}.$$

First consider the case $\alpha \geqslant 2/3$. For any $0 \leqslant z \leqslant 1$ and $c \geqslant 0$ we have

$$e^{cz} = \sum_{k=0}^\infty \frac{(cz)^k}{k!} \leqslant 1 + z\sum_{k=1}^\infty \frac{c^k}{k!} \leqslant 1 + e^c z,$$

hence

$$\frac{1}{p^\alpha} = \frac{1}{p} e^{(\log u)\frac{\log p}{\log y}} \leqslant \frac{1}{p}\left(1 + e^{\log u}\frac{\log p}{\log y}\right).$$

Thus, by Mertens bounds,

$$\begin{aligned}\log S &\leqslant O(1) + \sum_{p\leqslant y}\frac{1}{p^\alpha}\\ &\leqslant O(1) + \sum_{p\leqslant y}\frac{1}{p} + \frac{u}{\log y}\sum_{p\leqslant y}\frac{\log p}{p}\\ &\leqslant \log_2 y + O(u+1).\end{aligned}$$

Thus, $S \ll (\log y)e^{O(u)}$. Combining this with (5.7), we get (5.4) in this case.

Now assume $0 < \alpha < 2/3$. Then $y \leqslant u^3 \leqslant \log^3 x$ and $u \asymp \frac{\log x}{\log_2 x}$. We then have, using (5.5) again,

$$\begin{aligned}\log S &\ll \sum_{p\leqslant 2^{1/\alpha}}\frac{1}{\alpha\log p} + \int_1^y \frac{dt}{t^\alpha}\\ &\ll \alpha^{-1}2^{1/\alpha} + \frac{y^{1-\alpha}}{1-\alpha}\\ &\ll (\log x)^{o(1)} + u \ll u.\end{aligned}$$

Again, plugging this into (5.7) yields the claimed upper bound in (5.4).

Now we prove the lower bound in Theorem 5.1. This breaks into two cases,

(i) $\log x \leqslant y \leqslant \exp\{(\log_2 x)^{10}\}$;
(ii) $\exp\{(\log_2 x)^{10} \leqslant y \leqslant x$.

In each case we may assume that $x \geqslant x_0$, a sufficiently large constant, since for $x \leqslant x_0$, $\Psi(x,y) \geqslant 1$ and the result follows.

We begin with the easier case (i), following ideas from Tenenbaum [**140**, Ch. III.5.1]. Here

$$\frac{\log x}{(\log_2 x)^{10}} \leqslant u \leqslant \frac{\log x}{\log_2 x}. \tag{5.8}$$

Let $v = \lfloor u \rfloor$ and $k = \pi(y)$. Let $p_1, \dots, p_k$ be the primes $\leqslant y$. Then $\Psi(x,y)$ counts at least every number of the form $p_1^{\nu_1} \cdots p_k^{\nu_k}$ where $\nu_1 + \cdots + \nu_k \leqslant v$. Thus,

$$\Psi(x,y) \geqslant \binom{k+v}{v}.$$

By (5.8), $u \ll y/\log y \asymp k$. Thus, by Stirling's formula and $v = u + O(1)$, we deduce

$$\begin{aligned}\log \Psi(x,y) &\geqslant (k+v)\log(k+v) - k\log k - v\log v + O(\log(k+v)) \\ &= (k+u)\log(k+u) - k\log k - u\log u + O(\log y) \\ &= -u\log u + u\log k + (k+u)\log(1+u/k) + O(\log y).\end{aligned}$$

Again, (5.8) implies that $\log_2 y \ll \log_3 x \ll \log_2 u$, hence

$$u\log k = u\log y + O(u\log_2 y) = \log x + O(u\log_2 u).$$

Also,

$$(k+u)\log(1+u/k) \leqslant \frac{(k+u)u}{k} \leqslant u + \frac{u^2}{k} \ll u.$$

Hence,

$$\log \Psi(x,y) \geqslant \log x - u\log u + O(u\log_2 u),$$

as desired.

In case (ii), we will first show a weaker bound

$$\Psi(x,y) \gg x e^{-3u\log u} \qquad (1 \leqslant y \leqslant x). \tag{5.9}$$

If $y < \log x$ then $3u\log u > \log x$ and (5.9) is trivial, and if $\log x \leqslant y \leqslant \exp\{(\log_2 x)^{10}\}$ then (5.9) follows from the bound proved in case (i). Now suppose $y > \exp\{(\log_2 x)^{10}\}$. Let $k = \lfloor u \rfloor + 1$ and let I be the set of primes in the interval $(x^{\frac{1}{k+1}}, x^{\frac{1}{k}}]$. Then $\Psi(x,y)$ counts at least all numbers of the form $n = p_1 \cdots p_k m \leqslant x$ with $p_i \in I$ for each i and $p_1 < \cdots < p_k$ (these numbers are distinct since $m \leqslant x^{\frac{1}{k+1}}$). We observe that

$$x^{\frac{1}{k+1}} \geqslant x^{\frac{1}{u+2}} \geqslant x^{\frac{1}{3u}} = y^{1/3} > \exp\{(\log_2 x)^9\}.$$

By the Prime Number Theorem, we thus have for some constant $c > 0$

$$\sum_{p\in I} \frac{1}{p} = \log\left(\frac{k+1}{k}\right) + O\Big(e^{-c\sqrt{\log x^{1/(k+1)}}}\Big) = \log\left(\frac{k+1}{k}\right) + O(1/\log^{10} x).$$

Hence, by Lemma 5.2,

$$\begin{aligned}\Psi(x,y) &\geqslant \sum_{\substack{p_1,\dots,p_k\in I\\ p_1<p_2<\cdots<p_k}} \left\lfloor \frac{x}{p_1\cdots p_k}\right\rfloor \geqslant \sum_{\substack{p_1,\dots,p_k\in I\\ p_1<p_2<\cdots<p_k}} \frac{x}{2p_1\cdots p_k}\\ &\geqslant \frac{x/2}{k!}\Big(\sum_{p\in I}\frac{1}{p}-\frac{k}{x^{\frac{1}{k+1}}}\Big)^k\\ &\geqslant \frac{x/2}{k!}\Big(\log(1+1/k)-O\Big(\frac{1}{k^2\log^2 x}\Big)\Big)^k\\ &\geqslant \frac{x/2}{k!}\Big(\frac{1}{k}-\frac{1}{2k^2}--O\Big(\frac{1}{k^2\log^2 x}\Big)\Big)^k\\ &\gg \frac{x}{k^k\,k!}\gg xe^{-2k\log k+O(k)}\end{aligned}$$

and (5.9) follows since $k\leqslant u+1$.

Now we use (5.9) as a bootstrap for a better estimate. Suppose that $\exp\{(\log_2 x)^{10}\}\leqslant y\leqslant x$. Assume $u\geqslant e^4$, since the desired lower bound in Theorem 5.1 follows from (5.9) for $1\leqslant u\leqslant e^4$. Let $k=\lfloor u\rfloor+1$, let $\eta=\frac{1}{\log u}\leqslant\frac14$ and I be the set of primes in the interval $(y^{1-2\eta},y^{1-\eta}]$. Consider integers $n=p_1\cdots p_km$, where $p_i\in I$ for all i, $p_1<\cdots<p_k$ and $P^+(m)\leqslant y^{1-2\eta}$. These integers are distinct and counted by $\Psi(x,y)$. Then

$$\Psi(x,y)\geqslant \sum_{\substack{p_1,\dots,p_k\in I\\ p_1<\cdots<p_k}}\Psi\Big(\frac{x}{p_1\cdots p_k},y^{1-2\eta}\Big).$$

Now $p_1\cdots p_k\geqslant y^{k(1-2\eta)}\geqslant y^{u(1-2\eta)}=x^{1-2\eta}$ and thus

$$\frac{x}{p_1\cdots p_k}\leqslant x^{2\eta}=y^{2\eta u}\leqslant (y^{1-2\eta})^{4\eta u},$$

and $p_1\cdots p_k\leqslant (xy)^{1-\eta}=x^{(1+1/u)(1-\eta)}\leqslant x$, using that $\eta\leqslant\frac14$. Hence, by (5.9),

$$\begin{aligned}\Psi(x,y)&\gg x\sum_{\substack{p_1,\dots,p_k\in I\\ p_1<\cdots<p_k}}\frac{1}{p_1\cdots p_k}e^{-12\eta u\log(4\eta u)}\\ &\gg xe^{-O(u)}\sum_{\substack{p_1,\dots,p_k\in I\\ p_1<\cdots<p_k}}\frac{1}{p_1\cdots p_k}.\end{aligned}$$

Invoking Lemma 5.2 and using the Prime Number Theorem as in case (i), we conclude that

$$\begin{aligned}\Psi(x,y)&\gg \frac{xe^{-O(u)}}{k!}\Big(\sum_{p\in I}\frac{1}{p}+O\Big(\frac{1}{k^{100}\log x}\Big)\Big)^k\\ &\gg \frac{xe^{-O(u)}}{k!}\Big(\frac{1}{2\log u}\Big)^k\\ &= xe^{-u\log u+O(u\log_2(10u))}.\end{aligned}$$

This completes the proof of the lower bound in Theorem 5.1. □

Exercise 5.1. *(a) Show, using elementary estimates, that*

$$\Psi(x, x^{1/u}) = x(1 - \log u) + O(x/\log x)$$

uniformly for $1 \leqslant u \leqslant 2$.

(b) Show that for $y < z < x$ *that*

$$\Psi(x, y) = \Psi(x, z) - \sum_{y < p \leqslant z} \Psi(x/p, p).$$

(c) Let $\rho(u)$ *be the Dickman-de Bruijn function, defined recursively by*

$$\rho(u) = 1 \quad (0 \leqslant u \leqslant 1), \qquad \rho(u) = 1 - \int_1^u \frac{\rho(v-1)}{v}\, dv \quad (u > 1).$$

Using (b) and induction on $\lfloor u \rfloor$, *show that for any* $k > 1$, *uniformly for* $1 \leqslant u \leqslant k$ *we have*

$$\Psi(x, x^{1/u}) = \rho(u)x + O_k\left(\frac{x}{\log x}\right).$$

Exercise 5.2. *Prove that uniformly for* $10 \leqslant \log z \leqslant y \leqslant z \leqslant x/2$,

$$\Theta(x, y, z) := \#\left\{ n \leqslant x : \prod_{\substack{p^v \| n \\ p \leqslant y}} p^v > z \right\} = xe^{-u \log u + O(u \log_2(3u))}, \qquad u = \frac{\log z}{\log y}.$$

That is, counting numbers which have a large y*-smooth part.*

Exercise 5.3. *Let* $\mathcal{S}$ *denote the set of integers with* $(P^+(n))^2 | n$. *Show that*

$$\#\{n \leqslant x : n \in \mathcal{S}\} = x \exp\Big\{ -(\sqrt{2} + o(1))\sqrt{\log x \log_2 x} \Big\}.$$

6. Structure of shifted primes

Throughout, we will work with sets $\mathscr{P}_a = \{p+a : p \text{ prime }\}$, for a fixed nonzero a. The structure of these sets have numerous applications, particularly when $a = -1$ or $a = 1$. These include

(1) The study of Euler's totient function $\phi(n)$;
(2) The study of the sum-of-divisors function $\sigma(n)$;
(3) Problems about orders and primitive roots modulo primes;
(4) Carmichael numbers; analyzing primality tests

6.1. **The number of prime factors of shifted primes.** In 1935, Erdős [**32**] showed that for any fixed $a \neq 0$, the function $\omega(p+a)$ has normal order $\log_2 p$. To accomplish this, Erdős proved an upper bound of Hardy-Ramanujan type for the number of primes $p \leqslant x$ with $\omega(p+a) = k$. Here we prove a best-possible version of this estimate. Theorem 6.2 below improves the uniformity of Theorem 1 of Timofeev [**141**], who showed this bound uniformly for $k \leqslant b \log_2 x$ for any fixed b.

Lemma 6.1. *Fix $b \geqslant 0$. Uniformly for $x \geqslant 2$ and $\ell \geqslant 0$ we have*

$$\sum_{\substack{\omega(r)=\ell \\ rP^+(r)\leqslant x}} \frac{1}{\phi(r)\log^b(x/r)} \ll_b \frac{(\log_2 x + O(1))^\ell}{\ell! \log^b x}.$$

Proof. If $\ell = 0$ then $r = 1$ and the result is trivial. Now suppose $\ell \geqslant 1$. Then $2 \leqslant r \leqslant x/2$. We separately consider r in log-dyadic ranges. Let $Q_j = x^{1/2^j}$ for $j \geqslant 0$ and define

$$\mathscr{T}_j = \left\{ 2 \leqslant r \leqslant x/2 : \omega(r) = \ell, rP^+(r) \leqslant x, \frac{x}{Q_{j-1}} < r \leqslant \frac{x}{Q_j} \right\}.$$

For $r \in \mathscr{T}_j$, $P^+(r) \leqslant x/r \leqslant Q_{j-1}$. Also, if $\mathscr{T}_j$ is nonempty then $Q_{j-1} \geqslant 2$. We have

$$\sum_{r\in\mathscr{T}_j} \frac{1}{\phi(r)\log^b(x/r)} \leqslant \frac{1}{\log^b Q_j} \sum_{r\in\mathscr{T}_j} \frac{1}{\phi(r)}.$$

For the sum on the right side, we'll use a trick that was useful in the study of $\Psi(x,y)$. Let $\alpha = \frac{1}{10\log Q_j}$. Since $Q_{j-1} \geqslant 2$, $Q_j \geqslant \sqrt{2}$ and thus $0 < \alpha \leqslant \frac{1}{3}$. We'll encode the condition $r \geqslant x/Q_{j-1}$ with

$$\frac{1}{\phi(r)} = \frac{1}{\phi(r)^\alpha \phi(r)^{1-\alpha}} \ll \frac{1}{r^{\alpha/2}\phi(r)^{1-\alpha}} \ll \frac{1}{x^{\alpha/2}\phi(r)^{1-\alpha}},$$

since $(x/r)^{\alpha/2} \leqslant Q_{j-1}^{\alpha/2} = Q_j^\alpha = \mathrm{e}^{1/10}$. Now

$$\sum_{\substack{P^+(r)\leqslant Q_{j-1} \\ \omega(r)=\ell}} \frac{1}{\phi(r)^{1-\alpha}} \leqslant \frac{1}{\ell!}\left\{ \sum_{p\leqslant Q_{j-1}} \frac{1}{(p-1)^{1-\alpha}} + \frac{1}{(p(p-1))^{1-\alpha}} + \cdots \right\}^\ell$$

$$= \frac{1}{\ell!}\left\{ O(1) + \sum_{p\leqslant Q_{j-1}} \frac{1}{p^{1-\alpha}} \right\}^\ell.$$

Since $\log p \leqslant \log Q_{j-1} = 2\log Q_j$, we have $p^\alpha = 1 + O(\alpha \log p)$. It follows that

$$\sum_{p\leqslant Q_{j-1}} \frac{1}{p^{1-\alpha}} \leqslant \sum_{p\leqslant Q_{j-1}} \frac{1}{p} + O\left(\frac{\alpha\log p}{p}\right) \leqslant \log_2 Q_{j-1} + O(1) \leqslant \log_2 x + O(1).$$

Putting everything together, we see that

$$\sum_{r\in\mathscr{T}_j}\frac{1}{\phi(r)\log^b(x/r)}\ll\frac{1}{x^{\alpha/2}\log^b Q_j}\sum_{\substack{P^+(r)\leqslant Q_{j-1}\\ \omega(r)=\ell}}\frac{1}{\phi(r)^{1-\alpha}}$$
$$\leqslant\frac{1}{x^{\alpha/2}\log^b Q_j}\,\frac{(\log_2 x+O(1))^\ell}{\ell!}$$
$$=\frac{2^{bj}\exp\{-\frac{1}{20}\cdot 2^j\}}{\log^b x}\,\frac{(\log_2 x+O(1))^\ell}{\ell!}.$$

Summing over j completes the proof. □

Theorem 6.2 (K. Ford [**51**])**.** *Fix $a\neq 0$. Uniformly for $x\geqslant 4|a|$ and $k\in\mathbb{N}$ we have*

$$\#\{2+|a|<p\leqslant x:\omega(p+a)=k\}\ll_a\frac{x}{\log^2 x}\,\frac{(\log_2 x+O(1))^{k-1}}{(k-1)!}\qquad(2|a)$$

and

$$\#\{2+|a|<p\leqslant x:\omega(\tfrac{p+a}{2})=k\}\ll_a\frac{x}{\log^2 x}\,\frac{(\log_2 x+O(1))^{k-1}}{(k-1)!}\qquad(2\nmid a).$$

Proof. Let

$$s=\begin{cases}1 & 2|a\\ 2 & 2\nmid a.\end{cases}$$

Suppose $p>2+|a|\geqslant 3$ is prime and $\omega(\frac{p+a}{s})=k$. Define $q=P^+(\frac{p+a}{s})$ and write $\frac{p+a}{s}=qr$. Either $\omega(r)=k-1$ or $\omega(r)=k$, and r is odd if $2|a$. Also, $rP^+(r)\leqslant rq\leqslant(x+a)/s$. Define, for $\ell\geqslant 0$,

$$\mathscr{R}_\ell=\{r\in\mathbb{N}:rP^+(r)\leqslant\tfrac{x+a}{s},\omega(r)=\ell;r\text{ odd if }2|a\}.$$

We separately consider two cases. First, suppose that $1\leqslant k\leqslant\log_2 x$. Let $L=\exp\{\sqrt{\log x}\}$. Using Theorem 5.1, we have

$$\#\{2+|a|<p\leqslant x:q^2|\tfrac{p+a}{s}\}\leqslant\underbrace{\Psi(x+a,L)}_{q\leqslant L}+\sum_{q>L}\frac{x+a}{q^2}\ll_a\frac{x}{L}.\tag{6.1}$$

We use Theorem 2.5 to bound the contribution of those p with $q\|(p+a)$ thus:

$$\#\{2+|a|<p\leqslant x:\omega(\tfrac{p+a}{s})=k,q^2\nmid\tfrac{p+a}{s}\}\leqslant\sum_{r\in\mathscr{R}_{k-1}}\#\{q\leqslant\tfrac{x+a}{rs}:\underbrace{q,qsr-a}_{E=|rsa|}\text{ both prime}\}$$
$$\ll\sum_{r\in\mathscr{R}_{k-1}}\frac{|ars|}{\phi(|ars|)}\,\frac{x+a}{rs\log^2\frac{x+a}{rs}}$$
$$\ll_a x\sum_{r\in\mathscr{R}_{k-1}}\frac{1}{\phi(r)\log^2\frac{x+a}{r}}.$$

Invoking Lemma 6.1, we conclude that

$$\#\{2+|a|<p\leqslant x:\omega(\tfrac{p+a}{s})=k\}\ll_a\frac{x}{\log^2 x}\,\frac{(\log_2 x+O(1))^{k-1}}{(k-1)!}+\frac{x}{L}.$$

The final term x/L is negligible since $k \leqslant \log_2 x$ implies that the first term on the right is $\gg x/\log^2 x$. This concludes the proof when $k \leqslant \log_2 x$.

Now assume $k > \log_2 x$. Here we do not separately consider the case $q^2|\frac{p+a}{s}$) and do not use (6.1). We must then retain both cases $\omega(r) = k-1$ and $\omega(r) = k$. Arguing as before, we obtain from Theorem 2.5 and Lemma 6.1 the estimate

$$\begin{aligned}\#\{2+|a| < p \leqslant x : \omega(\tfrac{p+a}{s}) = k\} &\ll_a \sum_{r\in\mathscr{R}_{k-1}\cup\mathscr{R}_k} \frac{x}{\phi(r)\log^2((x+a)/r)} \\ &\ll_a \frac{x}{\log^2 x}\left(\frac{(\log_2 x + O(1))^{k-1}}{(k-1)!} + \frac{(\log_2 x + O(1))^k}{k!}\right) \\ &= \frac{x}{\log^2 x}\frac{(\log_2 x + O(1))^{k-1}}{(k-1)!}\left(1 + \frac{(\log_2 x + O(1))}{k}\right)\end{aligned}$$

Recalling that $k \geqslant \log_2 x$, this concludes the proof. □

6.2. **Large prime factors of shifted primes.** It is conjectured that $P^+(p+a)$ has a distribution roughly the same as the distribution of $P^+(n)$ over all integers $n \leqslant x$. We are far from proving this, in fact the following is not known:

Conjecture 6.3. *For any $0 < u < 1$, there are infinitely many primes p with $P^+(p+a) < p^u$ and infinitely many primes p with $P^+(p+a) > p^u$.*

The second assertion measures progress toward a special case of the Prime $k-$tuples conjecture 1.1. Letting $s = 1$ for even a and $s = 2$ for odd a, it's conjectured that infinitely often, $\frac{p+a}{s}$ is prime.

At present, it is know that $P^+(p+a) < p^{0.2844}$ infinitely often and $P^+(p+a) \geqslant p^{0.677}$ infinitely often. The first result in this direction is due to Erdős.

Theorem 6.4 (Erdős [**32**], 1935)**.** *For some $u < 1$ there are infinitely many primes p with $P^+(p+a) < p^u$.*

Definition 4. *Let μ_1 denote the infimum of all real numbers μ such that for all $a \neq 0$,*

$$\#\{p \leqslant x : P^+(p+a) \leqslant x^\mu\} > x^{1-o(1)}.$$

Upper bounds on μ_1 have been given in a series of papers, see Table 2.

Upper bound on μ_1	Author	year
$2\sqrt{2}-2 = 0.82842\ldots$	Wooldridge [**148**]	1979
$\frac{625}{512\mathrm{e}} = 0.44907\ldots$	Pomerance [**121**]	1980
$\frac{1}{2}e^{-1/2.73} = 0.34664\ldots$	Balog [**4**]	1984
$\frac{1}{2}e^{-1/2} = 0.30326\ldots$	Friedlander [**58**]	1989
0.2961	Baker and Harman [**7**]	1998
$\frac{15}{32}e^{-1/2} = 0.2844\ldots;$	Lichtman [**103**]	2022+

TABLE 2. Progression of results on smooth shifted primes

Here, we present an argument due to Balog which gives quantitatively better constants. The main input is a hypothesis known as the *Brun-Titmarsch on average* hypothesis. Here $N = N(x)$ is some function of x.

Hypothesis BT$(x, N, -a; \lambda)$**.** For all but $o(N)$ moduli $n \in (N, 2N]$ (as $N \to \infty$), we have

$$\pi(x; n, -a) \leqslant \lambda \frac{x}{\phi(n) \log x}. \tag{6.2}$$

By Theorem 4.2, for any $\varepsilon > 0$, any $1 < N < x$ with x sufficiently large (depending on ε), and any a, we have (6.2) with $\lambda = (2+\varepsilon)\frac{\log x}{\log(x/N)}$. The first result showing Hypothesis BT$(x, N, -a; \lambda)$ with a value of λ smaller than that achievable pointwise (that is, for *all* n, a) is due to Hooley [**91**], with further improvements due to Hooley [**93**], Fouvry [**57**], Bombieri, Friedlander and Iwaniec [**15**] and Baker and Harman [**6, 7**].

Theorem 6.5 (Balog [**4**], 1984)**.** *Let x be sufficiently large in terms of a, let $M = x^{\xi} \geqslant x^{1/2} \log^A x$, where A is a large enough constant. Assume that $\lambda \geqslant 1$ and Hypothesis BT$(x, N, -a; \lambda)$ holds for all $M < N \leqslant M \log x$. Then for any u satisfying*

$$\xi > u > \big(e^{-1}\xi^{\lambda}(1-\xi)\big)^{\frac{1}{1+\lambda}}, \tag{6.3}$$

there are $\gg_u x/\log^5 x$ primes $p \leqslant x$ so that $P^+(p+a) < x^u$.

Take $\xi > 1/2$, then Theorem 4.2 implies that BT$(x, x^{\xi}; -a, 4+\varepsilon)$ holds (with no exceptional set), where $\varepsilon \to 0$ as $\xi \to 1/2$ and $x \to \infty$. We conclude that

Corollary 6.6. *For any $u > \frac{1}{2}e^{-1/5} = 0.409365\ldots$, there are $\gg_u x/\log^5 x$ primes $p \leqslant x$ with $P^+(p+a) < x^u$. In particular, $\mu_1 \leqslant \frac{1}{2}e^{-1/5}$.*

Bombieri, Friedlander and Iwaniec [**15**] showed that Hypothesis BT$(x, x^{1/2}\log^A x, a; 1+\varepsilon)$ holds for all $A > 0$ and all $\varepsilon > 0$, if x is sufficiently large. It follows that (6.3) holds for all $u > \frac{1}{2}e^{-1/2} = 0.303\ldots$, giving Friedlander's result [**58**] that $\mu_1 \leqslant \frac{1}{2}e^{-1/2}$.

Proof of Theorem 6.5. Let $y = x^u$, and for $x/\log x < p \leqslant x$ let

$$g(p) = \#\{(m,n) : p + a = mn, P^+(mn) \leqslant y, \tfrac{x+a}{M\log x} < m \leqslant \tfrac{x+a}{M}, M < n \leqslant M\log x\}.$$

By Cauchy-Schwarz (see §2.2.4) we have

$$\#\{p \leqslant x : P^+(p+a) \leqslant y\} \geqslant \frac{\Big(\sum_{p\leqslant x} g(p)\Big)^2}{\sum_{p\leqslant x} g(p)^2}. \tag{6.4}$$

We bound the denominator using a crude argument:

$$\sum_{p\leqslant x} g(p)^2 \leqslant \sum_{p\leqslant x} \tau(p+a)^2 \leqslant \sum_{n\leqslant x+|a|} \tau(n)^2 \ll x\log^3 x. \tag{6.5}$$

Evidently,

$$\begin{aligned} g(p) \geqslant{}& \#\{(m,n) : p+a = mn, P^+(m) \leqslant y, \tfrac{x+a}{M\log x} < m \leqslant \tfrac{x+a}{M}, M < n \leqslant M\log x\} \\ &- \#\{(m,n) : p+a = mn, \tfrac{x+a}{M\log x} < m \leqslant \tfrac{x+a}{M}, M < n \leqslant M\log x, P^+(n) > y\}, \end{aligned}$$

where we have ignored the condition $P^+(m) \leqslant y$ in the subtracted term. Thus

$$\sum_{p \leqslant x} g(p) \geqslant S_1 - S_2, \tag{6.6}$$

where, writing $\mathcal{M}$ for the set of integers in $(\frac{x+a}{M \log x}, \frac{x+a}{M}]$ that have largest prime factor $\leqslant y$,

$$S_1 \geqslant \sum_{m \in \mathcal{M}} \pi(x; m, -a) - \pi(Mm - a; m, -a)$$

and, writing $n = q\ell$ with $q = P^+(n)$,

$$S_2 \leqslant \sum_{y < q \leqslant M \log x} \sum_{\substack{M/q < \ell \leqslant (M \log x)/q \\ P^+(\ell) \leqslant q}} \pi(x; \ell q, -a).$$

Since $\frac{x+a}{M} \ll x^{1/2} (\log x)^{-A}$, we use the Bombieri-Vinogradov theorem for S_1, obtaining

$$\begin{aligned} S_1 &\geqslant \sum_{m \in \mathcal{M}} \frac{1}{\phi(m)} \int_{Mn-a}^{x} \frac{dt}{\log t} + O\left(\frac{x}{\log x}\right) \\ &\geqslant \sum_{m \in \mathcal{M}} \frac{x - Mm}{\phi(m) \log x} + O\left(\frac{x}{\log x}\right) \\ &\geqslant \frac{x}{\log x} \sum_{m \in \mathcal{M}} \frac{1}{\phi(m)} + O\left(\frac{x}{\log x}\right) \end{aligned}$$

using the elementary bound

$$\sum_{m \leqslant z} \frac{m}{\phi(m)} = O(z).$$

We will also need the asymptotic

$$\sum_{m \leqslant z} \frac{1}{\phi(m)} = C_1 \log z + O(1) \tag{6.7}$$

for some constant $C_1 > 0$. If $P^+(m) > y$ write $m = qm'$ where $q = P^+(m)$. Then

$$\sum_{m \in \mathcal{M}} \frac{1}{\phi(m)} \geqslant \sum_{\frac{x+a}{M \log x} < m \leqslant \frac{x+a}{M}} \frac{1}{\phi(m)} - \sum_{y < q \leqslant \frac{x+a}{M}} \frac{1}{q-1} \sum_{\frac{x+a}{Mq \log x} < m' \leqslant \frac{x+a}{Mq}} \frac{1}{\phi(m')}.$$

The inner sum on m' is at most $C_1 \log_2 x + O(1)$ by (6.7). Therefore, by Mertens' theorems,

$$\begin{aligned} S_1 &\geqslant \frac{x}{\log x} (C_1 \log_2 x) \left(1 - \log \frac{\log(x/M)}{\log y} + o(1)\right) \\ &= \frac{x}{\log x} (C_1 \log_2 x) \left(1 - \log \frac{1-\xi}{u} + o(1)\right). \end{aligned} \tag{6.8}$$

Now we bound S_2 from above. The terms corresponding to $M < q \leqslant M \log x$ contribute, by the Brun-Titchmarsh inequality,

$$\ll \frac{x}{\log x} \sum_{M < q \leqslant M \log x} \frac{1}{q} \sum_{M/q < \ell \leqslant (M \log x)/q} \frac{1}{\phi(\ell)} \ll \frac{x (\log_2 x)^2}{\log^2 x}.$$

Let E denote the set of $n \in (M, M\log x]$ such that $\pi(x;n,-a) > \lambda\frac{x}{\phi(n)\log x}$. By our hypothesis (6.2) and the Brun-Titchmarsh inequality,

$$\sum_{n\in E}\pi(x;n,-a) \ll \frac{x}{\log x}\sum_{n\in E}\frac{1}{\phi(n)} = o\left(\frac{x\log_2 x}{\log x}\right).$$

Using (6.7) gain, the terms with $q \leqslant M$ and $q\ell \notin E$ contribute

$$\begin{aligned}&\leqslant \frac{\lambda x}{\log x}\sum_{y<q\leqslant M}\ \sum_{\substack{M/q<\ell\leqslant (M\log x)/q\\ q\ell\notin E}}\frac{1}{\phi(q\ell)}\\ &\leqslant (\lambda\log(\xi/u)+o(1))\frac{x}{\log x}\left(C_1\log_2 x\right).\end{aligned}$$

Therefore,

$$S_2 \leqslant (\lambda\log(\xi/u)+o(1))\frac{x}{\log x}\left(C_1\log_2 x\right). \tag{6.9}$$

Combining (6.8) and (6.9), we conclude that

$$\sum_{p\leqslant x} g(p) \geqslant \frac{x}{\log x}\left(C_1\log_2 x\right)\left(1-\log\frac{1-\xi}{u}-\lambda\log\frac{\xi}{u}+o(1)\right)$$

as $x\to\infty$. The hypothesis (6.3) implies that

$$1-\log\frac{1-\xi}{u}-\lambda\log\frac{\xi}{u} > 0.$$

Inserting the lower bound for $\sum_{p\leqslant x} g(p)$ into (6.4) and recalling (6.5), the proof is complete. □

Now we turn to the problem of showing that $P^+(p+a)$ is often large.

Definition 5. *Let μ_2 denote the supremum of all real numbers μ such that for any $a\neq 0$*

$$\#\{p\leqslant x : P^+(p+a) > x^\mu\} \gg_\mu \frac{x}{\log x}$$

The progression of records for the lower bound on μ_2 is given in Table 3.

Lower bound on μ_2	Author	year
$1-\frac{1}{2}e^{-1/4} = 0.6105\ldots$	Goldfeld [**63**]	1969
$1-\frac{1}{2}e^{-1/4} = 0.6105\ldots$	Motohashi [**112**]	1970
$\frac{1}{2}+\frac{3}{2}(1-e^{-1/12}) = 0.61993\ldots$	Hooley [**91**]	1972
$5/8 = 0.625$	Hooley [**92**]	1973
0.6683	Fouvry [**57**] [3]	1985
0.676	Baker and Harman [**6**]	1996
0.677	Baker and Harman [**7**]	1998

TABLE 3. Progression of results on large prime factors of shifted primes

Here we show the theorem proved independently by Goldfeld and Motohashi. We will then indicate where in the proof the further improvements come from.

Theorem 6.7. *We have* $\mu_2 \geqslant 1 - \frac{1}{2}e^{-1/4}$.

Proof. We use a device due to Chebyshev, in connection with the largest prime factor of polynomials. Let $1/2 < \theta < 1 - \frac{1}{2}e^{-1/4}$ and define

$$M = \log \prod_{|a|<p\leqslant x} (p+a).$$

Let B be the exponent in the Bombieri-Vinogradov Theorem (Theorem 3.4) corresponding to $A = 2$, and let $Q = x^{1/2}(\log x)^{-B}$. Let $\mathscr{P}$ denote the set of primes $|a| < p \leqslant x$ so that there is a prime power $q^b > Q$ with $b \geqslant 2$ and $q^b|(p+a)$. Define exponents $\alpha(q)$ by the prime factorization

$$\prod_{\substack{|a|<p\leqslant x\\ p\notin\mathscr{P}}} (p+a) = \prod_{q\leqslant x+a} q^{\alpha(q)}.$$

Then

$$M = M_1 + M_2 + M_3 + M_4,$$

where

$$\begin{aligned} M_1 &= \log \prod_{p\in\mathscr{P}} (p+a),\\ M_2 &= \log \prod_{q\leqslant Q} q^{\alpha(q)}.\\ M_3 &= \log \prod_{Q<q\leqslant x^\theta} q^{\alpha(q)},\\ M_4 &= \log \prod_{x^\theta<q\leqslant x+a} q^{\alpha(q)}. \end{aligned}$$

Our goal is to prove that $M_4 \gg x$. We accomplish this by first observing that

$$M \sim x \tag{6.10}$$

by the Prime Number Theorem, and then upper bounding M_1, M_2, M_3. For M_1, if $p \in \mathscr{P}$ then $p + a$ is divisible by d^2 for some $d > Q^{1/3}$. Hence

$$M_1 \ll (\log x) \sum_{d>Q^{1/3}} \frac{x}{d^2} \ll x^{6/7}.$$

We handle M_2 with the Bombieri-Vinogradov Theorem. Letting $E(x; s, c) = \pi(x; s, c) - \mathrm{li}(x)/\phi(s)$ we have

$$\begin{aligned} M_2 &= \sum_{q^b\leqslant Q} (\log q)\pi(x; q^b, -a)\\ &= \sum_{q^b\leqslant Q} (\log q)\left(\frac{\mathrm{li}(x)}{\phi(q^b)} + E(x; q^b, -a)\right)\\ &\sim \mathrm{li}(x)\log Q \sim \frac{x}{2}. \end{aligned}$$

We cannot evaluate M_3 asymptotically, but a good upper bound will suffice. The most crude method is to use the Brun-Titchmarsh inequality, Theorem 4.2 with explicit constant. This gives

$$\begin{aligned} M_3 &= \sum_{Q<q\leqslant x^\theta} (\log q)\pi(x;q,-a) \\ &\leqslant (2+o(1)) \sum_{Q<q\leqslant x^\theta} \frac{x\log q}{(q-1)\log(x/q)} \\ &= (2+o(1))x\int_Q^{x^\theta} \frac{dt}{t\log(x/t)} \\ &\sim \Big(2\log\frac{1}{2-2\theta}\Big)x. \end{aligned}$$

Combining the estimates for M_1, M_2 and M_3, we find that

$$M_1+M_2+M_3 \leqslant (1+o(1))\left(\frac{1}{2}+2\log\frac{1}{2-2\theta}\right)x.$$

Recalling the bound on θ, and comparing this with (6.10), we find that $M_4 \gg_\theta x$. As each number $p+a$ is divisible by at most one prime $> x^\theta$ we see that

$$\#\{p\leqslant x : P^+(p+a) > x^\theta\} \geqslant \frac{M_4}{\log(x+a)} \gg_\theta \frac{x}{\log x},$$

as desired. □

Remarks. Most improvements to Theorem 6.7 come from better estimates of M_3 using Hypothesis $\mathrm{BT}(x,N,a;\lambda)$ (an exception is the work of Hooley [**92**], which makes use of a Selberg-type sieve). These in turn rely on a variety of tools, from complicated sieves to bounds on Kloosterman sums.

The bound $\mu_2 > 2/3$ of Fouvry enabled Adleman and Heath-Brown [**1**] to deduce that the first case of Fermat's Last Theorem is true for infinitely many primes p; this is now a historical point in light of Wiles's proof 10 years later.

The bound on μ_2 also played an important role in the original AKS polynomial-time algorithm [**2**] for determining the primality of numbers.

6.3. **Application to Carmichael numbers.** A composite number n is a *Carmichael number* if for all $(a,n)=1$, $a^{n-1}\equiv 1 \pmod n$. Such a number mimics a prime w.r.t. the Fermat congruence. An old criteria of Korselt says that n is a Carmichael number if and only if n is squarefree and $(p-1)|(n-1)$ for every prime $p|n$. The smallest Carmichael numbers are 561, 1105 and 1729. One way to construct Carmichael numbers is to observe that if $k\in\mathbb{N}$, and $6k+1$, $12k+1$ and $18k+1$ are all prime, then $n=(6k+1)(12k+1)(18k+1)$ is a Carmichael number. By the Prime k-tuples conjecture (Conjecture 1.1), there are $\sim C\frac{x}{\log^3 x}$ such numbers $k\leqslant x$.

Alford, Granville and Pomerance showed unconditionally in 1994 [**3**] that there are infinitely many Carmichael numbers. Moreover, they showed a lower bound $C(x)\gg x^{2/7}$ for the counting function $C(x)$ of Carmichael numbers. A key ingredient is finding many smooth shifted primes $p-1$. In fact, Alford, Granville and Pomerance proved that

$$C(x) \gg x^{\frac{5}{12}(1-\mu_1)-o(1)}.$$

With Lichtman's [**103**] value $\mu_1 \leqslant 0.2844$, we get an exponent of $0.2981\ldots$. The $\frac{5}{12}$ fraction comes from lower bounds for prime in progressions to "virtually all" moduli. Harman [**81, 83**] improved the exponent, showing

$$C(x) \gg x^{0.4736(1-\mu_1)-o(1)},$$

where now the exponent is a bit larger than $1/3$.

6.4. **Applications to Euler's function and the sum of divisors function.**

6.4.1. *Popular values of Euler's function.* Let $A(m) = \#\{x : \phi(x) = m\}$. Erdős [**32**] showed that there is a constant $c > 0$ such that $A(m) > m^c$ infinitely often. A key ingredient in the proof is Theorem 6.4. The argument was sharpened by Pomerance [**121**], to connect c with the exponent in Theorem 6.4.

Theorem 6.8 (Erdős; Pomerance)**.** *For every $\varepsilon > 0$ there are infinitely many n such that $A(n) > n^{1-\mu_1-\varepsilon}$.*

Currently, the best bound we know is $\mu_1 \leqslant 0.2961$ due to Baker and Harman [**7**], while it is conjectured that $\mu_1 = 0$. The conclusion then would be best possible since trivially

$$A(m) \leqslant \#\{n : \phi(n) \leqslant m\} \sim Cm,$$

where $C = \zeta(2)\zeta(3)/\zeta(6)$; this last asymptotic is due to Bateman [**10**].

Proof. Fix $\mu_1 < \nu < 1$ and let z be large. Let

$$\mathscr{P} = \{p \leqslant (\log z)^{1/\nu} : P^+(p-1) \leqslant \log z\}.$$

By hypothesis, $\#\mathscr{P} \geqslant (\log z)^{1/\nu - o(1)}$ as $z \to \infty$. Let

$$M = \left\lfloor \frac{\nu \log z}{\log\log z} \right\rfloor,$$

and let $\mathscr{N}$ denote the set of all integers which are the product of M distinct primes from $\mathscr{P}$. Note that for all $n \in \mathscr{N}$, $\phi(n) \leqslant n \leqslant z$ and $P^+(\phi(n)) \leqslant \log z$. Hence, there is an $m \in \{\phi(n) : n \in \mathscr{N}\}$ such that

$$A(m) \geqslant \frac{\#\mathscr{N}}{\Psi(z, \log z)}$$

By Theorem 5.1 (inequality (5.1)), $\Psi(z, \log z) = z^{o(1)}$. Also, as $z \to \infty$, we have

$$\begin{aligned}
\#\mathscr{N} = \binom{\#\mathscr{P}}{M} &\geqslant \left(\frac{\#\mathscr{P}}{M}\right)^M \geqslant \left(\frac{(\log z)^{(1-o(1))/\nu}}{M}\right)^M \\
&\geqslant (\log z)^{\left[\frac{1}{\nu}-1-o(1)\right]\cdot\left[\frac{\nu\log z}{\log_2 z}-1\right]} \\
&= z^{1-\nu-o(1)}.
\end{aligned}$$

Hence there is an $m \leqslant z$ with $A(m) \geqslant z^{1-\nu-o(1)}$. Taking ν arbitrarily close to μ_1, the theorem follows. □

6.4.2. *The range of Euler's function.* Let

$$\mathscr{V} = \{\phi(n) : n \in \mathbb{N}\} = \{1, 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 28, \dots\}$$

denote the range of Euler's totient function $\phi(n)$, and let $V(x) = \#\{m \in \mathscr{V} : m \leqslant x\}$ be its counting function. Evidently $\mathscr{V}$ contains only one odd number, namely 1, but what can we say about the growth of $V(x)$? Since $\phi(p) = p - 1$ for all primes p, trivially

$$V(x) \geqslant \pi(x) \sim \frac{x}{\log x}.$$

Pillai [**116**] gave the first non-trivial upper bound on $V(x)$, namely

$$V(x) \ll \frac{x}{(\log x)^{(\log 2)/e}}.$$

Using sieve methods, Erdős [**32**] improved this to

$$V(x) = \frac{x}{(\log x)^{1-o(1)}} \quad (x \to \infty).$$

Upper and lower bounds for $V(x)$ were sharpened in a series of papers by Erdős [**35**], Erdős and Hall [**41, 42**], Pomerance [**123**], Maier and Pomerance [**106**], and finally by Ford [**47**]. The exact order of $V(x)$ is now known [**47**, Theorem 1], namely

$$V(x) = \frac{x}{\log x} \exp\{C(\log_3 x - \log_4 x)^2 + D \log_3 x - (D + 1/2 - 2C) \log_4 x + O(1)\},$$

where $C = 0.817814646400836322 31\dots$ and $D = 2.17696874355941032173\dots$ are specific constants defined in [**47**]. Sieve methods played a crucial role in these investigations. Below we use the results about the normal behavior of $\omega(n)$ and $\omega(p + a)$ to provide nontrivial upper and lower boudns for $V(x)$.

Theorem 6.9. *We have*

$$V(x) \gg \frac{x \log_2 x}{\log x}.$$

Proof. Let $y = \exp\{\sqrt{\log x}\}$ and let

$$\mathcal{B} = \{pq : q \leqslant y < p \leqslant x/q, p \text{ and } q \text{ prime}\}.$$

Evidently $\phi(m) \leqslant x$ for $m \in \mathcal{B}$. Our goal is to show that there is little overlap among the numbers $\phi(m)$. By simple inclusion-exclusion,

$$V(x) \geqslant |\mathcal{B}| - Z, \quad Z := \#\{m_1, m_2 \in \mathcal{B} : m_1 \neq m_2, \phi(m_1) = \phi(m_2)\}. \tag{6.11}$$

By Mertens,

$$|\mathcal{B}| = \sum_{q \leqslant y} \pi(x/q) - \pi(y) \gg \sum_{q \leqslant y} \frac{x}{q \log x} \gg \frac{x \log_2 x}{\log x}.$$

Also, Z is the number of solutions of

$$(p - 1)(q - 1) = (p' - 1)(q' - 1) \qquad (p \neq p', pq \in \mathcal{B}, p'q' \in \mathcal{B}).$$

Fix q and q', and write

$$g = (q - 1, q' - 1), \quad r = \frac{q - 1}{g}, s = \frac{q' - 1}{g}.$$

Then the above equation reduces to

$$(p-1)r = (p'-1)s,$$

where $(r,s)=1$. Thus, for some $n \leqslant x/(qs)$ we have $p' = 1+rn$ and $p = 1+sn$. Apply Theorem 2.5 with the pair of linear forms $rn+1, sn+1$. Here $E = rs|r-s| \ll x^3$. Hence

$$\begin{aligned}\#\{n \leqslant x/(qs) : nr+1, ns+1 \text{ both prime}\} &\ll \frac{x/(qs)}{\log^2(x/(qs))}(\log_2 x)^2 \\ &\ll \frac{x(\log_2 x)^2}{\log^2 x}\,\frac{g}{qq'}\end{aligned}$$

We now sum over q, q' using Exercise 2.12. This gives

$$\begin{aligned}\sum_{\substack{q,q' \leqslant y \\ q \neq q'}} \frac{(q-1,q'-1)}{qq'} &\leqslant \sum_{g \leqslant y/2} g \Bigg(\sum_{\substack{q \leqslant y \\ q \equiv 1 \pmod g}} \frac{1}{q} \Bigg)^2 \\ &\ll (\log_2 x)^2 \sum_{g \leqslant y/2} \frac{g}{\phi(g)^2} \\ &\leqslant (\log_2 x)^2 \prod_{p \leqslant y/2} \left(1 + \frac{p}{(p-1)^2} + O\Big(\frac{1}{p^2}\Big)\right) \\ &\ll (\log_2 x)^2 \sqrt{\log x}.\end{aligned}$$

It follows that

$$Z \ll \frac{x(\log_2 x)^4}{(\log x)^{3/2}}.$$

Combined with (6.11), this completes the proof. □

Lemma 6.10 (Poisson tails)**.** *Let X be a Poisson random variable with parameter λ. Then*

$$\mathbb{P}(X \leqslant \alpha\lambda) \leqslant \mathrm{e}^{-Q(\alpha)\lambda} \quad (0 \leqslant \alpha \leqslant 1), \qquad \mathbb{P}(X \geqslant \alpha\lambda) \leqslant \mathrm{e}^{-Q(\alpha)\lambda} \quad (\alpha \geqslant 1),$$

where

$$Q(x) = x\log x - x + 1. \tag{6.12}$$

Proof. We use the idea behind Chernoff's inequality, Markov's inequality, and the easy identity $\mathbb{E}c^X = \mathrm{e}^{(c-1)\lambda}$ for $c > 0$. Let $b = \log\alpha$. Then

$$\mathbb{P}(X \leqslant \alpha\lambda) = \mathbb{P}\left(\mathrm{e}^{bX} \geqslant \mathrm{e}^{b\alpha\lambda}\right) \leqslant \frac{\mathbb{E}\mathrm{e}^{bX}}{\mathrm{e}^{b\alpha\lambda}} = \frac{\mathrm{e}^{(\mathrm{e}^b-1)\lambda}}{\mathrm{e}^{b\alpha\lambda}} = \mathrm{e}^{-Q(\alpha)\lambda}.$$

Similarly,

$$\mathbb{P}(X \geqslant \alpha\lambda) = \mathbb{P}\left(\mathrm{e}^{bX} \geqslant \mathrm{e}^{b\alpha\lambda}\right) \leqslant \frac{\mathbb{E}\mathrm{e}^{bX}}{\mathrm{e}^{b\alpha\lambda}} = \frac{\mathrm{e}^{(\mathrm{e}^b-1)\lambda}}{\mathrm{e}^{b\alpha\lambda}} = \mathrm{e}^{-Q(\alpha)\lambda}.$$ □

Lemma 6.11. *Uniformly for $k \geqslant 1$ we have*

$$\#\{n \leqslant x : \Omega(n) \geqslant k\} \ll \frac{xk\log x}{2^k}.$$

Proof. Take $y = 2-1/k$ and write $y^{\Omega(n)} = (1\star g)(n)$, where g is multiplicative, and $g(p^k) = y^k - y^{k-1}$ for prime p. Since $g(n) \geqslant 0$ for all n,

$$\begin{aligned}\sum_{n\leqslant x} y^{\Omega(n)} &\leqslant x\sum_{d\leqslant x}\frac{g(d)}{d} \leqslant x\prod_{p\leqslant x}\left(1+\sum_{k=1}^{\infty}\frac{y^k-y^{k-1}}{p^k}\right)\\ &= x\prod_{p\leqslant x}\left(1+\frac{y-1}{p-y}\right)\\ &\leqslant \frac{x}{2-y}\exp\left\{\sum_{3\leqslant p\leqslant x}\frac{y-1}{p-y}\right\}\\ &\ll (xk)(\log x)^{y-1}.\end{aligned}$$

Then,

$$\begin{aligned}\#\{n\leqslant x : \Omega(n)\geqslant k\} &\leqslant \sum_{n\leqslant x} y^{\Omega(n)-k}\\ &\ll y^{-k}(xk)(\log x)^{y-1}\\ &\leqslant \frac{xk}{2^k(1-\frac{1}{2k})^k}(\log x)^{1-1/k} \ll \frac{xk\log x}{2^k}.\end{aligned}$$

□

Theorem 6.12. *We have*

$$V(x) \ll \frac{x}{\log x}\exp\left\{3.5(\log_3 x)^2\right\}.$$

Proof. Let

$$\varepsilon = 0.0001, \quad \delta = Q(1-\varepsilon) > 0.$$

Let

$$\mathcal{S} = \{p : \omega(p-1) \leqslant (1-\varepsilon)\log_2 p\}$$

and let $S(x)$ be the counting function of $\mathcal{S}$. By Theorem 6.2, and Lemma 6.10,

$$\begin{aligned}S(x) &\leqslant \#\{p\leqslant x : \omega(p-1)\leqslant (1-\varepsilon)\log_2 x\}\\ &\ll \frac{x}{\log^2 x}\sum_{k\leqslant(1-\varepsilon)\log_2 x}\frac{(\log_2 x + O(1))^{k-1}}{(k-1)!}\\ &\ll \frac{x}{\log^2 x}\sum_{k\leqslant(1-\varepsilon)\log_2 x}\frac{(\log_2 x)^{k-1}}{(k-1)!}\\ &\ll \frac{x}{(\log x)^{1+\delta}}.\end{aligned} \tag{6.13}$$

Also define

$$W(x) = \max_{y\leqslant x}\frac{\log y}{y}V(y).$$

Although $\frac{\log y}{y}V(y)$ may not be monotone, $W(x)$ is and this will be convenient in the proof. We note that $V(2) = 2$ and thus $W(x) \geqslant \log 2$ for $x \geqslant 2$.

Let x_0 be sufficiently large and assume that $x \geqslant x_0$. Then there is a $y \leqslant x$ with

$$W(x) = \frac{\log y}{y}V(y). \tag{6.14}$$

By Theorem 6.9, $W(x) \to \infty$ as $x \to \infty$ and hence $y \to \infty$ as $x \to \infty$. We now bound $V(y)$. Let

$$\theta = 0.485, \qquad y_1 = \exp\left\{\frac{(\log y)^\theta}{3\log_2 y}\right\}.$$

Associate to each $v \in \mathscr{V}$ with $v \leqslant y$ a preimage n, that is, $\phi(n) = v$. We have $n \ll y\log_2 y$ by a classical estimate.

We divide these v into four classes:

(a) $V_1 = \{v : \Omega(v) \geqslant 2.9\log_2 y \text{ or } \Omega(n) \geqslant 2.9\log_2 y\}$;
(b) $V_2 = \{v : p^2|n \text{ for some } p > y_1\}$;
(c) $V_3 = \{v \notin V_1 \cup V_2 : p|n \text{ for some } p > y_1, p \in \mathcal{S}\}$;
(d) $V_4 = (\mathscr{V} \cap [1, y]) \setminus (V_1 \cup V_2 \cup V_3)$ (the remaining v).

Lemma 6.11 implies immediately that

$$|V_1| \ll \frac{y(\log_2 y)^2}{(\log y)^{2.9\log 2-1}} \ll \frac{y}{(\log y)^{1.01}}. \tag{6.15}$$

Trivially,

$$|V_2| \ll \sum_{p>y_1} \frac{y\log_2 y}{p^2} \ll \frac{y}{(\log y)^{100}}. \tag{6.16}$$

Now suppose that $v \in V_3$. Then $p|n$ for some $p \in \mathcal{S}, p > y_1$. Since $v \notin V_2$, $v = (p-1)\phi(n/p) = (p-1)w$ for some $w \in \mathscr{V}$, $w \leqslant \frac{y}{p-1}$. Thus,

$$|V_3| \leqslant \sum_{\substack{y_1<p\leqslant y\\ p\in\mathcal{S}}} V\left(\frac{y}{p-1}\right) \ll \underbrace{S(y)}_{p>y/2} + W(y)y \sum_{\substack{y_1<p\leqslant y/2\\ p\in\mathcal{S}}} \frac{1}{p\log(y/p)}.$$

A standard partial summation argument, together with (6.13), yields

$$\begin{aligned}\sum_{\substack{y_1<p\leqslant y/2\\ p\in\mathcal{S}}} \frac{1}{p\log(y/p)} &\ll \frac{S(y)}{y} + \int_{y_1}^{y/2} \frac{S(t)}{t^2\log(y/t)}\,dt \\ &\ll \frac{1}{(\log y)^{1+\delta}} + \frac{1}{(\log y_1)^\delta}\int_{y_1}^{y/2} \frac{dt}{t(\log t)\log(y/t)} \\ &\ll \frac{(\log_2 y)^2}{(\log y)^{1+\delta\theta}}.\end{aligned}$$

Therefore, by the monotonicity of $W()$,

$$|V_3| \ll W(y)\frac{y(\log_2 y)^2}{(\log y)^{1+\delta\theta}} \ll \frac{yW(x)}{(\log y)^{1+\delta\theta/2}}. \tag{6.17}$$

Finally, suppose that $v \in V_4$. Let

$$y_2 = y_1^{3\log_2 y} = \exp\{(\log y)^\theta\}.$$

We claim that n has at most 5 prime factors $> y_1$. Since $v \notin V_2$, the prime factors of n which are $> y_1$ are distinct. Since $v \notin V_3$, any prime factor p of n which is $> y_1$ is not in $\mathcal{S}$, and thus satisfies $\omega(p-1) > (1-\varepsilon)\log_2 p \geqslant (1-\varepsilon)\log_2 y_1$. If n has 6 or more such factors, then

$$\Omega(v) \geqslant 6(1-\varepsilon)\log_2 y_1 > 6(1-\varepsilon)(\theta - o(1))\log_2 y > 2.9\log_2 y,$$

for y large enough, contradicting that $v \notin V_1$. This proves the claim. Thus, $n = qm$, where $P^+(m) \leqslant y_1$, $P^-(q) > y_1$, q is squarefree and $\omega(q) \leqslant 5$. But then $m \leqslant y_1^{\Omega(m)} \leqslant y_1^{\Omega(n)} \leqslant y_2$ since $v \notin V_1$ and $\phi(n) = \phi(q)\phi(m)$. Let $\mathcal{Q}$ denote the set of such numbers q which satisfy $\phi(q) \leqslant y$. Thus,

$$|V_4| \leqslant V(y_2) + \sum_{q\in\mathcal{Q}} V\left(\min\left(y_2, \frac{y}{\phi(q)}\right)\right).$$

For $t \leqslant y_2$, $V(t) \leqslant \frac{t}{\log t} W(t) \leqslant \frac{t}{\log t} W(y_2)$. We have $\phi(q) \gg q$ for such q, so $q \ll y$. When $q > y/2$, $V(y/\phi(q)) \ll 1$ and there are $O(y(\log_2 y)^4/\log y)$ such numbers q by the Hardy-Ramanujan inequality. Therefore, since $W(y_2) \gg 1$,

$$\begin{aligned} |V_4| &\ll W(y_2)\left[\frac{y(\log_2 y)^4}{\log y} + \sum_{\substack{q\leqslant y/2\\ q\in\mathcal{Q}}} \frac{y}{q\log(y/q)}\right] \\ &\ll W(y_2)\frac{y(\log_2 y)^5}{\log y}. \end{aligned}$$

Combining this with (6.14), (6.15), (6.16) and (6.17) gives

$$W(x) \ll \frac{1}{(\log y)^{0.01}} + \frac{W(x)}{(\log y)^{\delta\theta/2}} + W(y_2)(\log_2 y)^5.$$

Again, using the monotonicity of $W()$, and the fact that x and y are sufficiently large,

$$W(y_2) \leqslant W\big(\exp\{(\log x)^\theta\}\big).$$

Also recall that $W(x) \to \infty$ as $x \to \infty$, and we obtain

$$W(x) \leqslant K(\log_2 x)^5 W\left(\exp\{(\log x)^\theta\}\right)$$

for some absolute constant $K > 0$. Iterating this expression yields

$$W(x) \leqslant K^k \theta^{5k(k-1)/2} (\log_2 x)^{5k} W\left(\exp\{(\log x)^{\theta^k}\}\right)$$

provided that $(\log x)^{\theta^k} \geqslant x_0$. Taking

$$k = \left\lfloor \frac{\log_3 x}{-\log\theta} - c \right\rfloor$$

for a large enough constant c produces

$$W(x) \leqslant \exp\left\{\frac{2.5}{-\log(\theta)}(\log_3 x)^2 + O(\log_3 x)\right\} \ll \exp\{3.5(\log_3 x)^2\}.$$

This completes the proof. □

6.5. **Exercises.**

Exercise 6.1. *Show that for every $\alpha < \frac{1}{2}$, a positive proportion of primes p satisfy* **both** $P^+(p-1) > p^\alpha$ *and* $P^+(p+1) > p^\alpha$.

Exercise 6.2. *Improve the upper estimate* (6.5) *for* $\sum_{p\leqslant x} g(p)^2$.

Exercise 6.3 (Ford, 1995 - see [**47**])**.** *(a) Suppose than $m \geqslant 2$ and $A(m) = k$. Also suppose that $p > 2m+1$ is a prime such that*

(i) $2p+1$ *and* $2mp+1$ *are both prime;*
(ii) $dp+1$ *is composite for all* $d|2m$ *except for* $d=2, d=2m$.

Show that $A(2mp)=k+2$.

(b) Assume the Prime k-tuples conjecture, Conjecture 1.1. Show that a prime p satisfying the above conditions always exists as long as $m \geqslant 2$ *and* $m \not\equiv 2 \pmod 3$. *Use this to conclude that for any* $k \geqslant 2$ *that there is a number m with* $A(m)=k$. *[This is a conjecture of Sierpiński (see* [**128**] *and* [**38**]*), which was proved for even k by Ford and Konyagin* [**56**] *and for odd k by Ford* [**48**]*.]*

Exercise 6.4. *Carmichael* [**21, 22**] *conjectured that* $A(m)=1$ *is impossible. Assume that* $A(m)=1$ *and* $\phi(n)=m$.

(a) *(Carmichael-Klee) Show that if* $d\prod_{p|d} p|n$ *and* $q=1+d$ *is prime, then* $q^2|n$.
(b) *Show that* $2^2 3^2 7^2 43^2|n$.
(c) *Show that either* $13^2|n$ *or* $19^2|n$. *Hint: consider the two cases* $3^2\|n$ *and* $3^3|n$.

7. Small gaps between primes

7.1. **Introduction.** Prior to 2005, there were only weak results know about small gaps between primes. Let p_n denote the n-th prime, and set

$$\Delta = \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log n}.$$

The Prime Number Theorem implies $\Delta \leqslant 1$. Prior to 2005, it was not known that $\Delta = 0$, which is itself a very weak bound in the direction of the Twin Prime Conjecture. Here we summarize the major events:

- Hardy-Littlewood, 1926 [**79**] (unpublished): $\Delta \leqslant 2/3$ assuming the Extended Riemann Hypothesis for Dirichlet L-functions.
- Erdős, 1940 [**34**]: $\Delta < 1$ unconditionally.
- Bombieri and Davenport, 1966 [**14**] : $\Delta \leqslant \frac{2+\sqrt{3}}{8} = 0.466\ldots$
- Maier, 1988 [**105**]: $\Delta \leqslant 0.248$.

We begin by showing Erdős' argument from 1940, which incorporated a basic sieve bound.

Theorem 7.1. *Suppose that, uniformly for all even $h \leqslant 2\log x$ we have*

$$\#\{x/2 < n \leqslant x : n \text{ and } n+h \text{ are both prime}\} \leqslant (B+o(1))C_h \frac{x/2}{\log^2 x} \quad (x\to\infty) \tag{7.1}$$

where

$$C_h = C\prod_{\substack{p\mid h\\ p>2}} \frac{p-1}{p-2}, \quad C = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right).$$

Then $\Delta \leqslant 1 - \frac{1}{2B}$.

Basic upper bound sieves, such as Theorem 2.5, show that (7.1) holds for some finite B, while Theorem 4.5 implies that $B = 4$ is admissible.

Proof. The basic idea is that the Prime Number Theorem implies that the average gap between primes up to x is $\sim \log x$, but (7.1) implies that there are not many gaps that are extremely close to $\log x$.

Let x be large and let $p_{n+1},\dots,p_{n+m+1}$ be the primes in $(x/2,x]$. Put $d_j = p_{j+1}-p_j$ for all j. Fix $\frac{1}{3B} < \varepsilon < \frac{1}{2B}$ and let

$$I = \big[(1-\varepsilon)\log x, (1+\varepsilon)\log x\big].$$

Assume that $d_j > (1-\varepsilon)\log x$ for all $n+1 \leqslant j \leqslant n+m$. For each even $k \in I$, let

$$N_k = \#\{n+1 \leqslant j \leqslant n+m : d_j = k\},$$

so that by (7.1), we have

$$N_k \leqslant (B+o(1))C_k \frac{x/2}{\log^2 x}. \tag{7.2}$$

Now

$$\begin{aligned}\frac{x}{2} &\geqslant \sum_{j=n+1}^{n+m} d_j = \sum_{k>(1-\varepsilon)\log x} kN_k \\ &\geqslant \sum_{k\in I} kN_k + (1+\varepsilon)\log x\Big(m - \sum_{k\in I} N_k\Big) \\ &= (1+\varepsilon)m\log x - \sum_{k\in I} N_k\left((1+\varepsilon)\log x - k\right).\end{aligned}$$

The Prime Number Theorem gives $m \sim \frac{x/2}{\log x}$. Hence, by (7.2),

$$\frac{x}{2} \geqslant \left(\frac{1+\varepsilon}{2} + o(1)\right)x - \frac{(B/2+o(1))x}{\log^2 x}\sum_{k\in I} C_k\left((1+\varepsilon)\log x - k\right). \tag{7.3}$$

We claim that

$$\sum_{k\leqslant y} C_k = y + O(\log^2 y). \tag{7.4}$$

Assuming the claim, partial summation gives

$$\sum_{k\in I} kC_k \sim 2\varepsilon\log^2 x,$$

and we conclude from (7.3) that

$$0 \geqslant \frac{\varepsilon}{2} - (B/2)\big(2\varepsilon(1+\varepsilon) - 2\varepsilon\big) + o(1) = \frac{\varepsilon}{2} - B\varepsilon^2 + o(1).$$

This is a contradiction if $\varepsilon < \frac{1}{2B}$, and proves the theorem.

Now we prove the claim (7.4). Write $C_k = C\sum_{d|k} g(d)$, where g is multiplicative, supported on squarefree integers, $g(2) = 0$ and $g(p) = \frac{1}{p-2}$ for primes $p > 2$. Thus,

$$g(d) \ll \frac{1}{d}\left(\frac{d}{\phi(d)}\right)^2 \ll \frac{(\log_2 d)^2}{d} \ll \frac{\log d}{d}.$$

We then compute

$$\begin{aligned}\sum_{k\leqslant y} C_k &= C\sum_{\substack{d\leqslant y\\ d \text{ odd}}} g(d)\sum_{\substack{k\leqslant y\\ (2d)|k}} 1 = C\sum_{\substack{d\leqslant y\\ d \text{ odd}}} g(d)\Big\lfloor\frac{y}{2d}\Big\rfloor \\ &= \frac{Cy}{2}\sum_{d}\frac{g(d)}{d} + O\Big(y\sum_{d>y}\frac{g(d)}{d} + \sum_{d\leqslant y} g(d)\Big) \\ &= \frac{Cy}{2}\prod_{p>2}\left(1+\frac{g(p)}{p}\right) + O(\log^2 y) \\ &= y + O(\log^2 y),\end{aligned}$$

which proves (7.4). □

7.2. **Improved gap detectors: GPY-Zhang-Maynard-Tao.** In 2005 (the paper appeared in 2009), Goldston, Pintz and Yıldırım[**64**] showed that $\Delta = 0$ using a new type of prime detector. Let $(h_1, \dots, h_k)$ be an *admissible* k-tuple of non-negative integers; that is, the set of linear forms $(n + h_1, \dots, n + h_k)$ is admissible in the sense of Definition 1. For some non-negative "weight function" $w(n)$, consider

$$S = \sum_{N<n\leqslant 2N} \Big(\sum_{j=1}^{k} \mathbb{1}_{n+h_j \text{ prime}} - 1 \Big) w(n). \tag{7.5}$$

If $S > 0$ then there is an $n \in (N, 2N]$ such that at least two of the forms $n + h_j$ are simultaneously prime. The object is to choose $w(n)$ so that

(a) $w(n)$ is small when there are few primes among the $n + h_j$
(b) $w(n)$ is large when $n + h_j$ is prime (or nearly prime) for many j;
(c) We can find a good upper bound for $\sum_{N<n\leqslant 2N} w(n)$ and, for every j, a good lower bound for the sum $\sum_{N<n\leqslant 2N} w(n)\mathbb{1}_{n+h_j \text{ prime}}$.

A naive choice is

$$w(n) = \mathbb{1}_{n+h_j \text{ prime } \forall j}.$$

This satisfies (a) and (b) but not (c), because we need something like the Hardy-Littlewood conjectures (Conjecture 1.1) to get the required lower bounds. Motivated by Selberg's sieve, Goldston, Pintz and Yıldırım make the choice

$$w(n) = \Big(\sum_{d|(n+h_1)\cdots(n+h_k)} \lambda_d \Big)^2, \qquad \lambda_d = \mu(d) \left(\frac{\log(z/d)}{\log z} \right)^{k+\ell}, \quad \ell \geqslant 0.$$

Elaborating on the ideas in [**64**], they proved in [**65**] that

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\sqrt{\log n}(\log_2 n)^2} < \infty.$$

More interestingly, they showed that any exponent improvement in the Bombieri-Vinogradov Theorem would imply that bounded gaps exist infinitely often. More precisely,

Theorem 7.2 (Goldston-Pintz-Yıldırım, 2007)**.** *Assume that for some $\theta > \frac{1}{2}$ we have*

$$EH(\theta) : \forall A > 0, \sum_{q\leqslant x^\theta} \max_{(a,q)=1} \max_{y\leqslant x} \left| \pi(y; q, a) - \frac{\mathrm{li}(y)}{\phi(q)} \right| \ll_A \frac{x}{\log^A x}. \tag{7.6}$$

Then there is a constant $C(\theta)$ such that $p_{n+1} - p_n \leqslant C(\theta)$ infinitely often. In particular, $C(0.971) = 16$.

Unconditionally, the Bombieri-Vinogradov Theorem (Theorem 3.4) implies EH(θ) for all $\theta < 1/2$, while Elliott and Halberstam [**31**] conjectured EH(θ) for all $\theta < 1$.

Conjecture 7.3 (Elliott-Halberstam Conjecture)**.** *We have $EH(\theta)$ for every $\theta < 1$.*

Motohashi and Pintz [**113**] showed in fact that the bounded gaps conclusion would follow from a version of (7.6) restricted to smooth moduli up to x^θ, where $\theta > 1/2$, that is moduli with $P^+(q) \leqslant x^\delta$. Yitang Zhang [**152**] then surprised the world in May, 2013 by supplying such a proof of the modified version of (7.6).

Theorem 7.4 (Yitang Zhang, 2013 [**152**])**.** *For infinitely many n, $p_{n+1} - p_n \leqslant 70000000$.*

A large on-line collaborative Polymath project [**119**] refined Zhang's ideas in the summer of 2013 and subsequently reduced the bound to about 4600. Then, in October, 2013, James Maynard [**108**] introduced a better weight function $w(n)$, which reduced the bound further to 600 and furthermore showed that bounded gaps could contain arbitrarily many primes. Around the same time, Terence Tao (unpublished) had a similar idea for improving the weight function.

Theorem 7.5 (Maynard-Tao)**.** *For any $m \geqslant 2$ there is a number $C_m \ll m^2 e^{4m}$ such that $p_{n+m-1} - p_n \leqslant C_m$ infinitely often. Assuming the Elliott-Halberstam Conjecture, we have $C_2 \leqslant 12$ and $C_m \ll m^2 e^{2m}$.*

A second Polymath project [**120**] resulted in further numerical improvements, culminating in the bound $p_{n+1} - p_n \leqslant 246$ infinitely often. This is the current world record.

7.3. **Bounded gaps between primes.** We now prepare the ground for the proof of Theorem 7.5. The method is robust enough to handle general linear forms (not only those of type $n + h_i$), and moreover the bounds can be made uniform in the coefficients of the forms, as in [**109**], with little additional effort.

Theorem 7.6. *For every $m \in \mathbb{N}$ there is a number K_m such that the following holds. Fix $k \in \mathbb{N}$ with $k \geqslant K_m$. Uniformly over all $N \geqslant 100$ and every admissible set $(a_1 n + b_1, \dots, a_k n + b_k)$ of linear forms, with $1 \leqslant a_i \leqslant (\log N)^{100}$ and $-a_i N/2 \leqslant b_i \leqslant N(\log N)^{100}$ for all i, we have*

$$\#\left\{N < n \leqslant 2N : \sum_{j=1}^{k} \mathbb{1}_{a_i n + b_i \text{ prime}} \geqslant m\right\} \geqslant \frac{N}{(\log N)^{c_k}}$$

for some constant c_k depending only on k. Moreover, we have

(a) *$K_m \ll m e^{4m}$ for all $m \geqslant 2$.*
(b) *Assuming the Elliott-Halberstam Conjecture, we have $K_m \ll m e^{2m}$.*

The PolyMath8b project [**120**] implies[4] that $K_2 = 50$ is admissible. Also, Maynard [**108**] showed $K_2 = 5$ is admissible assuming the Elliott-Halberstam conjecture.

Proof of Theorem 7.5 from Theorem 7.6. Let $k = K_m$. Fix $h_1 < \ldots < h_k$ so that $n + h_1, \dots, n + h_k$ is an admissible set of linear forms (we say that the tuple $(h_1, \dots, h_k)$ is admissible). Apply Theorem 7.6, we see that

$$\liminf_{n \to \infty} p_{n+m-1} - p_n \leqslant h_k - h_1.$$

When $m = 2$, there is an admissble tuple with $h_{50} - h_1 = 246$, proved in [**120**]. For general m, let $h_1, \dots, h_k$ be the k smallest primes greater than k. Then $n + h_1, \dots, n + h_k$ is clearly admissible, and by the Prime Number Theorem,

$$h_k - h_1 \sim k \log k = K_m \log K_m \ll m^2 e^{4m}.$$

From Exercise 4.4, for any admissible tuple $(h_1, \dots, h_k)$, $h_k - h_1 \geqslant (1/2 + o(1))k \log k$, and the asymptotic true minimum of $h_k - h_1$ is unknown. Assuming the Elliott-Halberstam Conjecture,

[4]The theorems in [**120**] are stated only in the case $a_i = 1$ and b_i fixed, however it is clear from the proof that the same quantitative bounds hold in the range of uniformity of Theorem 7.6. The relevant parts of [**120**] to be adjusted are Theorems 19 (i), 20 (i) and 26.

the above calculation yields $C_m \ll m^2 e^{2m}$. When $m = 2$, we may take $k = 5$, and we apply Theorem 7.6 with the 5-tuple $(0, 2, 6, 8, 12)$.

□

We now describe the prime-detector which we will use to prove Theorem 7.6. Let $N \geqslant 100$, assume EH(θ) for some $\theta \in [1/3, 1)$, and suppose that the parameters s, R, D, z satisfy

$$\frac{\log_2 N}{2} \leqslant s \leqslant 2\log_2 N, \quad N^{\frac{\theta}{2}-\frac{2}{s}} \leqslant R \leqslant N^{\frac{\theta}{2}-\frac{1}{s}}, \quad D = R^{1/s}, \quad z = D^{1/s} = R^{1/s^2}. \tag{7.7}$$

For each $k \in \mathbb{N}$ we will construct a special sequence $\lambda(\mathbf{d})$ (which depends only on k, z and R) for $\mathbf{d} = (d_1, \dots, d_k) \in \mathbb{N}^k$, which is supported on the set

$$\mathscr{D} = \mathscr{D}(N) := \{\mathbf{d} \in \mathbb{N}^k : d_1 \cdots d_k \leqslant R, \mu^2(d_1 \cdots d_k) = 1, P^-(d_1 \cdots d_k) > z\}. \tag{7.8}$$

We will frequently use the fact that $\mathscr{D}$ is *divisor-closed*, that is, if $\mathbf{d} \in \mathscr{D}$ and $r_j | d_j$ for all j then $\mathbf{r} \in \mathscr{D}$.

Let μ^+ denote an upper bound sieve which is guaranteed by Theorem 3.6, the Fundamental Lemma (with respect to the parameters z, D). In particular we have

$$|\mu^+(t)| \leqslant 1 \qquad (t \in \mathbb{N}). \tag{7.9}$$

Now let $(a_1 n + b_1, \dots, a_k n + b_k)$ be an arbitrary admissible set of linear forms (in particular, $a_i \neq 0$ and $(a_i, b_i) = 1$ for all i), with respect to Definition 1. Let

$$E = E(\mathbf{a}, \mathbf{b}) = \prod_{i=1}^{k} a_i \prod_{i<j} (a_i b_j - a_j b_i), \quad \mathscr{E} = \mathscr{E}(\mathbf{a}, \mathbf{b}) = \Big\{\mathbf{d} \in \mathbb{N}^k : (d_1 \cdots d_k, E) = 1\Big\}. \tag{7.10}$$

Let

$$\rho(d) = \#\{n \mod d : (a_1 n + b_1) \cdots (a_k n + b_k) \equiv 0 \pmod d\},$$

and note that $\rho(p) < p$ for all p (because the set of forms is admissible) and $\rho(p) = k$ for all $p \nmid E$; see Theorem 2.5. Let

$$V = \prod_{p \leqslant z} \left(1 - \frac{\rho(p)}{p}\right). \tag{7.11}$$

Finally, motivated by Selberg's sieve and using an idea from Koukoulopulos [**100**, Ch. 28], let

$$w(n) = \Bigg(\sum_{t | (a_1 n + b_1) \cdots (a_k n + b_k)} \mu^+(t)\Bigg)\Bigg(\sum_{\substack{\mathbf{d} \in \mathscr{D} \cap \mathscr{E} \\ \forall j : d_j | a_j n + b_j}} \lambda(\mathbf{d})\Bigg)^2. \tag{7.12}$$

By Theorem 3.6, $w(n) \geqslant 0$ for all n. The first factor is a familiar upper bound sieve that we used in Theorem 2.5, for example, and very efficiently sieves out those n for which some $a_i n + b_i$ has a prime factor $\leqslant z$. The second factor is the new ingredient, a kind of k-dimensional version of the Selberg method. We will optimize $\lambda(\mathbf{d})$ later, and at this point only impose the normalizing condition

$$|\lambda(\mathbf{d})| \leqslant 1 \qquad (\mathbf{d} \in \mathscr{D}). \tag{7.13}$$

The purpose of imposing $\mathbf{d} \in \mathscr{E}$ is to ensure that $d_1, \dots, d_k$ are pairwise relatively prime.

Our main object is to show that

$$S := \sum_{N<n\leqslant 2N} \left(\sum_{j=1}^{k} \mathbb{1}_{a_j n+b_j \text{ prime}} - (m-1) \right) > 0,$$

which will show that for some $n \in (N, 2N]$, at least m of the forms $a_j n + b_j$ are simultaneously prime. We will in fact get a lower bound on S that will imply that there are many such n.

Lemma 7.7. *Suppose that $m_j | a_j n + b_j$ for every j and $\mathbf{m} \in \mathscr{E}$. Then $m_1, \dots, m_k$ are pairwise relative prime, and $(m_j, a_j) = 1$ for all j.*

Proof. If $p|m_i$ and $p|m_j$ then p divides $a_i(a_j n + b_j) - a_j(a_i n + b_i) = a_i a_j - a_j a_i$, and so $p|E$. This proves the first claim. Since $(a_1 n + b_1, \dots, a_k n + b_k)$ is an admissible set, $(a_j, b_j) = 1$ for all j. Hence, if p is prime, $p|m_j|(a_j n + b_j)$ and $p|a_j$ then $p|b_j$, a contradiction. Thus, $(a_j, m_j) = 1$. □

Notational Convention. Throughout the remainder of this section, constants implied by O and $\ll$ symbols may depend on k and θ, but not on any other parameter. We also suppose that N is sufficiently large, in terms of k and θ. In particular, our bounds are uniform in the coefficients a_i, b_i in some range.

Proposition 7.8. *Assume* (7.7), (7.8), *and let μ^+ be an upper bound sieve function from Theorem 3.6 with parameters z, D. Let $\lambda(\mathbf{d})$ satisfy* (7.13) *and be supported on $\mathscr{D}$. For $\mathbf{r} \in \mathscr{D}$ define*

$$\xi(\mathbf{r}) = \sum_{\mathbf{d}} \frac{\lambda(r_1 d_1, \dots, r_k d_k)}{d_1 \cdots d_k}. \tag{7.14}$$

Let $(a_1 n + b_1, \dots, a_k n + b_k)$ be an admissible set of linear forms, with $k \geqslant 2$ and such that

$$1 \leqslant |a_i| \leqslant N^2, \qquad |b_i| \leqslant N^2 \qquad (1 \leqslant i \leqslant k). \tag{7.15}$$

Define $E, \mathscr{E}$ by (7.10), *V by* (7.11) *and $w(n)$ by* (7.12). *Then*

$$\sum_{N<n\leqslant 2N} w(n) = VN \sum_{\mathbf{r}\in\mathscr{D}} \frac{\xi(\mathbf{r})^2}{r_1 \cdots r_k} + O\left(\frac{N}{(\log N)^{99k}} \right).$$

Proof. From (7.12), we have

$$\sum_{N<n\leqslant 2N} w(n) = \sum_{t\leqslant D} \mu^+(t) \sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\leqslant 2N \\ t|(a_1 n+b_1)\cdots(a_k n+b_k) \\ \forall j:[d_j,e_j]|(a_j n+b_j)}} 1.$$

Since $\mathbf{d}, \mathbf{e} \in \mathscr{E}$, by Lemma 7.7 we have $(d_i e_i, d_j e_j) = 1$ for all $i \neq j$ and $(d_i e_i, a_i) = 1$ for all i. Thus, the simultaneous conditions $[d_j, e_j]|(a_j n + b_j)$, $1 \leqslant j \leqslant k$, define a single residue class n modulo $\prod_{j=1}^{k} [d_j, e_j]$. Also, the condition $t|(a_1 n+b_1)\cdots(a_k n+b_k)$ defines $\rho(t)$ residue classes modulo t. But $P^+(t) \leqslant z < P^-(d_j e_j)$ for each j, hence n runs over $\rho(t)$ residue classes modulo $t[d_1, e_1]\cdots[d_k, e_k]$. Hence,

$$\begin{aligned} \sum_{N<n\leqslant 2N} w(n) &= \sum_{t\leqslant D} \mu^+(t) \sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}} \lambda(\mathbf{d})\lambda(\mathbf{e}) \left(\frac{\rho(t)N}{t[d_1, e_1]\cdots[d_k, e_k]} + O(\rho(t)) \right) \\ &= NV^+B + T, \end{aligned} \tag{7.16}$$

where

$$V^+ = \sum_{t\leqslant D} \frac{\mu^+(t)\rho(t)}{t}, \qquad B = \sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{[d_1,e_1]\cdots[d_k,e_k]},$$

and, by (7.13) and (7.9),

$$|T| \ll |\mathscr{D}|^2 \sum_{t\leqslant D} \rho(t).$$

We begin with the error terms T. We have

$$\begin{aligned} |\mathscr{D}| &\leqslant \sum_{\substack{r\leqslant R\\ P^-(r)>z}} k^{\omega(r)}\mu^2(r) \leqslant R \sum_{\substack{r\leqslant R\\ P^-(r)>z}} \frac{k^{\omega(r)}\mu^2(r)}{r} \\ &\leqslant R\prod_{z<p\leqslant R}\left(1+\frac{k}{p}\right) \ll R(s^2)^k = R(\log_2 N)^{2k}, \end{aligned}$$

and

$$\sum_{t\leqslant D}\rho(t) \leqslant D\sum_{t\leqslant D}\frac{\rho(t)}{t} \leqslant D\prod_{p\leqslant D}\left(1+\frac{k}{p}\right) \ll D(\log N)^k.$$

Hence,

$$T \ll R^2 D(\log N)^{3k} \ll N^{\theta} \ll \frac{N}{(\log N)^{99k}}. \tag{7.17}$$

The function $g(t) = \rho(t)/t$ satisfies (Ω) with $\kappa = k$ since $\rho(p) \leqslant k$ for all p. Therefore, by the Fundamental Lemma (Theorem 3.6),

$$V^+ = V\big(1+O(e^{-\frac{1}{2}s\log s})\big) = V + O\left(\frac{1}{(\log N)^{100k}}\right). \tag{7.18}$$

This is a genuine asymptotic since $V \gg (\log z)^{-k}$. We now record a very crude bound for B. For any $m_i = [d_i, e_i]$, there are at most $3^{\omega(m_i)}$ choices for d_i, e_i. Hence, from (7.13),

$$\Big|\sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{m_1\cdots m_k}\Big| \leqslant \prod_{i=1}^{k}\sum_{\substack{m_i\leqslant R^2\\ P^-(m_i)>z}} \frac{3^{\omega(m_i)}}{m_i} \leqslant \prod_{z<p\leqslant R^2}\left(1+\frac{3}{p}\right)^k \ll s^{6k} \ll \log N. \tag{7.19}$$

To asymptotically bound B, we first remove the conditions $\mathbf{d},\mathbf{e}\in\mathscr{E}$. Now from (7.10) and (7.15),

$$E \ll N^{2k+4(k^2/2)}.$$

Hence there are $\ll \frac{\log N}{\log z} \ll s^2$ prime factors of E which are $> z$. If $\mathbf{d}\notin\mathscr{E}$ or $\mathbf{e}\notin\mathscr{E}$ then there is a $p|E$ with $p|m_j$ for some j. Write $m_j = pm_j'$, then analogously to (7.19) we have

$$\Big|\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathscr{D}\\ \mathbf{d}\notin\mathscr{E}\text{ or }\mathbf{e}\notin\mathscr{E}}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{m_1\cdots m_k}\Big| \leqslant \sum_{j=1}^{k}\sum_{\substack{p|E\\ p>z}}\sum_{\substack{m_j'\leqslant R^2\\ P^-(m_j')>z}} \frac{3^{\omega(m_j')+1}}{m_j'p}\prod_{i\neq j}\sum_{\substack{m_i\leqslant R^2\\ P^-(m_i)>z}}\frac{3^{\omega(m_i)}}{m_i} \ll \frac{s^2}{z}s^{6k} \ll \frac{\log N}{z}.$$

Therefore,

$$B = O\left(\frac{\log N}{z}\right) + B', \quad B' = \sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{[d_1,e_1]\cdots[d_k,e_k]}. \tag{7.20}$$

Combining (7.16) with (7.17), (7.18) and (7.20), plus the crude bound (7.19), we find that

$$\begin{aligned}\sum_{N<n\leqslant 2N} w(n) &= NVB' + T + O\left(\frac{NV\log N}{z} + \frac{NB'}{(\log N)^{100k}}\right)\\ &= NVB' + O\left(\frac{N}{(\log N)^{99k}}\right).\end{aligned} \tag{7.21}$$

It remains to bound B'. Write

$$\frac{1}{[d,e]} = \frac{(d,e)}{de} = \frac{1}{de}\sum_{r|(d,e)}\phi(r). \tag{7.22}$$

Then

$$\begin{aligned}B' &= \sum_{\mathbf{r}\in\mathscr{D}}\phi(r_1)\cdots\phi(r_k)\Big(\sum_{\forall j: r_j|d_j}\frac{\lambda(\mathbf{d})}{d_1\cdots d_k}\Big)\Big(\sum_{\forall j: r_j|e_j}\frac{\lambda(\mathbf{e})}{e_1\cdots e_k}\Big)\\ &= \sum_{\mathbf{r}\in\mathscr{D}}\frac{\phi(r_1)\cdots\phi(r_k)}{r_1^2\cdots r_k^2}\xi(\mathbf{r})^2.\end{aligned}$$

For any $r\leqslant R$, r has at most s^2 prime factors $>z$. Hence, for all r_i,

$$\frac{\phi(r_i)}{r_i} = \prod_{p|r_i}(1-1/p) = 1 + O\left(\frac{s^2}{z}\right) = 1 + \left(\frac{\log N}{z}\right). \tag{7.23}$$

We have the crude bound

$$\xi(\mathbf{r}) \leqslant \Big(\sum_{\substack{d\leqslant R\\ P^-(d)>z}}\frac{1}{d}\Big)^k \leqslant \prod_{z<p\leqslant R}\Big(1+\frac{1}{p}\Big)^k \ll s^{2k} \ll \log N. \tag{7.24}$$

Therefore,

$$B' = \sum_{\mathbf{r}\in\mathscr{D}}\frac{\xi(\mathbf{r})^2}{r_1\cdots r_k}\left(1+O\left(\frac{\log N}{z}\right)\right) = \sum_{\mathbf{r}\in\mathscr{D}}\frac{\xi(\mathbf{r})^2}{r_1\cdots r_k} + O\left(\frac{\log^4 N}{z}\right).$$

The big-O term is $O(1/\log^{99k} N)$. Combining this with (7.21) completes the proof. □

Proposition 7.9. *Assume $1/3<\theta<1$ and EH(θ) holds. Assume (7.7), (7.8), and let μ^+ be an upper bound sieve function from Theorem 3.6 with parameters z, D. Let $\lambda(\mathbf{d})$ satisfy (7.13) and be supported on $\mathscr{D}$. For $\mathbf{r}\in\mathscr{D}$ and $1\leqslant m\leqslant k$ define*

$$\zeta_m(\mathbf{r}) = \mathbb{1}_{r_m=1}\sum_{\substack{\mathbf{d}\in\mathscr{D}\\ d_m=1}}\frac{\lambda(r_1d_1,\dots,r_kd_k)}{d_1\cdots d_k}. \tag{7.25}$$

Let $(a_1n+b_1,\dots,a_kn+b_k)$ be an admissible set of linear forms, with $k\geqslant 2$, such that

$$\begin{aligned}&1\leqslant a_m\leqslant(\log N)^{100}, \qquad -\frac{N}{2}a_m\leqslant b_m\leqslant N\log^{100}N\\ &1\leqslant|a_i|\leqslant N^2, \qquad |b_i|\leqslant N^2 \qquad (i\neq m).\end{aligned} \tag{7.26}$$

Define $E,\mathscr{E}$ by (7.10), V by (7.11) and $w(n)$ by (7.12). Then, for $1\leqslant m\leqslant k$ we have

$$\sum_{N<n\leqslant 2N} w(n)\mathbb{1}_{a_mn+b_m \text{ prime}} = \frac{VY_m}{\prod_{p\leqslant z}(1-1/p)}\sum_{\mathbf{r}\in\mathscr{D}}\frac{\zeta_m(\mathbf{r})^2}{r_1\cdots r_k} + O\left(\frac{N}{(\log N)^{40k^2}}\right),$$

where

$$Y_m := \frac{\operatorname{li}(2a_mN+b_m)-\operatorname{li}(a_mN+b_m)}{a_m} \sim \frac{N}{\log N}.$$

Proof. We begin with

$$\sum_{N<n\leqslant 2N} w(n)\mathbb{1}_{a_mn+b_m \text{ prime}} = \sum_{t\leqslant D}\mu^+(t)\sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}}\lambda(\mathbf{d})\lambda(\mathbf{e}) \sum_{\substack{N<n\leqslant 2N\\ t|(a_1n+b_1)\cdots(a_kn+b_k)\\ \forall j:[d_j,e_j]|(a_jn+b_j)\\ a_mn+b_m \text{ prime}}} 1. \tag{7.27}$$

Since $[d_m,e_m]\leqslant R^2 = N^{\theta-2/s} < N/2$, $p=a_mn+b_m$ is prime and

$$p = a_mn+b_m \geqslant \frac{a_mN}{2} \geqslant \frac{N}{2}, \tag{7.28}$$

we see that

$$d_m = e_m = 1.$$

The range of n implies that

$$a_mN+b_m < p \leqslant 2a_mN+b_m.$$

Since $\mathbf{d},\mathbf{e}\in\mathscr{E}$, by Lemma 7.7 we have $(d_ie_i,d_je_j)=1$ for all $i\neq j$ and $(d_ie_i,a_i)=1$ for all i. Therefore, for each $i\neq m$, the condition $[d_i,e_i]|a_in+b_i$ is equivalent to

$$p \equiv a_i^{-1}(a_ib_m-a_mb_i) \pmod{[d_i,e_i]},$$

and thus p lies in a single residue class modulo $[d_i,e_i]$, and moreover this residue class is coprime to $[d_i,e_i]$ since $\mathbf{d},\mathbf{e}\in\mathscr{E}$. We have $t|(a_1n+b_1)\cdots(a_kn+b_k)$ and $(p,t)=1$, thus to

$$\prod_{i\neq m}(a_in+b_i)\equiv 0 \pmod t, \qquad (a_mn+b_m,t)=1.$$

This defines $\rho^*(t)$ residue classes for n modulo t, where ρ^* is multiplicative by the Chinese Remainder Theorem, and moreover for primes q we have

$$\rho^*(q) = \begin{cases} \rho(q) & \text{if } q|a_m \\ \rho(q)-1 & \text{if } q\nmid a_m. \end{cases} \tag{7.29}$$

To see this, note that when $q|a_m$, the congruence $a_mn+b_m\equiv 0 \pmod q$ has no solutions, and for any n we have $(a_mn+b_m,q)=1$ since $(a_m,b_m)=1$. When $q\nmid a_m$, there are $\rho(q)$ solutions of

$$(a_1n+b_1)\cdots(a_kn+b_k)\equiv 0 \pmod q,$$

including the single solution of $a_mn+b_m\equiv 0 \pmod q$, which must be removed.

Therefore, the prime p lies in one of $\rho^*(t)$ reduced residue classes modulo a_mt. By assumption, $1\leqslant a_m\leqslant \log^{100}N < z$, and so a_mt is coprime to $[d_1,e_1]\cdots[d_k,e_k]$. Hence, the inner sum in (7.27) defines precisely $\rho^*(t)$ reduced residue classes for the prime p modulo $a_mt[d_1,e_1]\cdots[d_k,e_k]$. Define $E(u)$ by

$$E(u) = \max_{(u,s)=1}\left|\pi(2a_mN+b_m;u,s)-\pi(a_mN+b_m;u,s)-\frac{\operatorname{li}(2a_mN+b_m)-\operatorname{li}(a_mN+b_m)}{\phi(u)}\right|$$

and write $u = a_m t[d_1, e_1] \cdots [d_k, e_k]$. Then, by (7.27),

$$\begin{aligned}\sum_{N<n\leqslant 2N} w(n)\mathbb{1}_{a_m n+b_m \text{ prime}} &= \sum_{t\leqslant D}\mu^+(t)\sum_{\substack{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}\\ d_m=e_m=1}}\lambda(\mathbf{d})\lambda(\mathbf{e})\left[\rho^*(t)\frac{a_m Y_m}{\phi(u)}+O(\rho^*(t)E(u))\right]\\ &= a_m Y_m V^* B^* + T^*,\end{aligned} \tag{7.30}$$

where, since $P^+(a_m t) \leqslant z < P^-([d_1, e_1] \cdots [d_k, e_k])$,

$$V^* = \sum_{t\leqslant D}\frac{\mu^+(t)\rho^*(t)}{\phi(a_m t)},$$

$$B^* = \sum_{\substack{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}\\ d_m=e_m=1}}\frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{\phi([d_1, e_1]\cdots[d_k, e_k])},$$

$$|T^*| \ll \sum_{\substack{t\leqslant D\\ P^+(t)\leqslant z}}\rho^*(t)\sum_{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}} E(u).$$

We use Hypothesis EH(θ) to handle the error term T^*. Note that $x := 2a_m N + b_m \geqslant 3N/2$ by hypothesis, and since $\mathbf{d}, \mathbf{e} \in \mathscr{D}$, the moduli u satisfy

$$u \leqslant a_m t d_1 \cdots d_k e_1 \cdots e_k \leqslant a_m D R^2 \leqslant N^{\theta-\frac{1}{2s}} \leqslant x^\theta$$

if N is large enough. For each squarefree $q = [d_1, e_1] \cdots [d_k, e_k]$, there are $\leqslant (3k)^{\omega(q)}$ ways to choose $d_1, e_1, \ldots, d_k, e_k$. Also, $\rho^*(t) \leqslant \rho(t) \leqslant k^{\omega(t)}$. Thus, by Cauchy-Schwarz and the trivial bound

$$E(u) \ll \frac{x}{u} \ll \frac{N(\log N)^{100}}{u},$$

we have the estimate

$$\begin{aligned}|T^*| &\ll \sum_{\substack{t\leqslant D\\ P^+(t)\leqslant z}}\mu^2(t)k^{\omega(t)}\sum_{\substack{P^-(q)>z\\ q\leqslant R^2}}\mu^2(q)(3k)^{\omega(q)}E(a_m t q)\\ &\leqslant \sum_{r\leqslant DR^2}\mu^2(r)(3k)^{\omega(r)}E(a_m r)^{1/2}\left(\frac{N\log^{100} N}{r}\right)^{1/2}\\ &\ll (N\log^{100} N)^{1/2}\bigg(\sum_{P^+(r)\leqslant N}\frac{\mu^2(r)(3k)^{2\omega(r)}}{r}\bigg)^{1/2}\bigg(\sum_{r\leqslant DR^2}E(a_m r)\bigg)^{1/2}\\ &\ll (N\log^{100} N)^{1/2}(\log N)^{9k^2/2}\left(\frac{x}{(\log N)^{1000k^2}}\right)^{1/2}.\end{aligned}$$

From the bound $x \ll N(\log N)^{100}$, we conclude that

$$T^* \ll \frac{N}{(\log N)^{100k^2}}. \tag{7.31}$$

As in the proof of Proposition 7.8, we have

$$\prod_{i=1}^m \frac{[d_i, e_i]}{\phi([d_i, e_i])} = 1 + O\left(\frac{\log N}{z}\right).$$

Hence, by the argument in (7.19),

$$B^* = O\left(\frac{\log^2 N}{z}\right) + \sum_{\substack{\mathbf{d},\mathbf{e}\in\mathscr{D}\cap\mathscr{E}\\ d_m=e_m=1}} \frac{\lambda(\mathbf{d})\lambda(\mathbf{e})}{[d_1,e_1]\cdots[d_k,e_k]}.$$

As in the proof of Proposition 7.8, the "missing terms", those with either $\mathbf{d}\notin\mathscr{E}$ or $\mathbf{e}\notin\mathscr{E}$, have total $O(\frac{\log N}{z})$. Using (7.22) and (7.23) again, we get

$$B^* = O\left(\frac{\log^2 N}{z}\right) + \sum_{\mathbf{r}\in\mathscr{D}} \frac{\zeta_m(\mathbf{r})^2}{r_1\cdots r_k}. \tag{7.32}$$

Finally, apply the Fundamental Lemma of the sieve (Theorem 3.6) with the function

$$g(t) = \frac{\rho^*(t)\phi(a_m)}{\phi(a_m t)},$$

where, by (7.29), for primes p we have

$$g(p) = \begin{cases} \frac{\rho^*(p)}{p} = \frac{\rho(p)}{p} & \text{if } p|a_m,\\ \frac{\rho^*(p)}{p-1} = \frac{\rho(p)-1}{p-1} & \text{if } p\nmid a_m. \end{cases}$$

Again, $g(p)\leqslant 2k/p$ for all p, thus (Ω) holds with $\kappa = 2k$. Then, by Theorem 3.6,

$$V^* = \frac{1}{\phi(a_m)}\left(1+O(e^{-\frac{1}{2}s\log s})\right)\prod_{p\leqslant z}(1-g(p)) = \left(1+O\left(\frac{1}{(\log N)^{100k^2}}\right)\right)V^{**},$$

where, using (7.29),

$$\begin{aligned} V^{**} &= \frac{1}{\phi(a_m)}\prod_{p|a_m}\left(1-\frac{\rho(p)}{p}\right)\prod_{\substack{p\leqslant z\\ p\nmid a_m}}\left(1-\frac{\rho(p)-1}{p-1}\right)\\ &= \frac{V}{\phi(a_m)}\prod_{\substack{p\leqslant z\\ p\nmid a_m}}\left(1-\frac{\rho(p)}{p}\right)^{-1}\left(1-\frac{\rho(p)-1}{p-1}\right)\\ &= \frac{V}{\phi(a_m)}\prod_{p\leqslant z}\left(1-\frac{1}{p}\right)^{-1}\prod_{p|a_m}\left(1-\frac{1}{p}\right)\\ &= \frac{V}{a_m}\prod_{p\leqslant z}\left(1-\frac{1}{p}\right)^{-1}. \end{aligned}$$

Therefore,

$$V^* = \frac{V}{a_m}\prod_{p\leqslant z}\frac{p}{p-1} + O\left(\frac{1}{(\log N)^{90k^2}}\right). \tag{7.33}$$

The same argument leading to (7.19) yields $B^*\ll\log N$. Combining (7.30) with (7.31), (7.32) and (7.33) yields the claimed asymptotic. □

Comparing the conclusions of Propositions (7.8) and (7.9), we see that our goal is to maximize the ratio

$$\left(\sum_{\mathbf{r}\in\mathscr{D}}\frac{\zeta_m(\mathbf{r})^2}{r_1\cdots r_k}\right)\Big/\left(\sum_{\mathbf{r}\in\mathscr{D}}\frac{\xi(\mathbf{r})^2}{r_1\cdots r_k}\right). \tag{7.34}$$

Next, we prove two relatively easy inversion formulas.

Lemma 7.10. *For all $\mathbf{r}\in\mathscr{D}$ and $1\leqslant m\leqslant k$,*

$$\zeta_m(\mathbf{r}) = \mathbb{1}_{r_m=1}\sum_{b\in\mathbb{N}}\frac{\mu(b)\xi(r_1,\dots,r_{m-1},b,r_{m+1},\dots,r_k)}{b}.$$

Proof. Let $r_m=1$. By (7.14), the right side equals

$$\begin{aligned}&=\sum_b\frac{\mu(b)}{b}\sum_{\mathbf{d}}\frac{\lambda(r_1d_1,\dots,r_{m-1}d_{m-1},d_mb,r_{m+1}d_{m+1},\dots)}{d_1\cdots d_k}\\ &=\sum_{d_i:i\neq m}\frac{1}{\prod_{i\neq m}d_i}\sum_{\ell\in\mathbb{N}}\frac{\lambda(r_1d_1,\dots,r_{m-1}d_{m-1},\ell,r_{m+1}d_{m+1},\dots)}{\ell}\sum_{b|\ell}\mu(b)=\zeta_m(\mathbf{r}).\end{aligned}$$

□

Lemma 7.11. *For all $\mathbf{d}$,*

$$\lambda(\mathbf{d}) = \mathbb{1}_{\mathbf{d}\in\mathscr{D}}\sum_{\mathbf{b}}\frac{\mu(b_1)\cdots\mu(b_k)\xi(b_1d_1,\dots,b_kd_k)}{b_1\cdots b_k}.$$

Proof. Let $\mathbf{d}\in\mathscr{D}$. By (7.14), the right side is

$$\begin{aligned}&=\sum_{\mathbf{b}}\frac{\mu(b_1)\cdots\mu(b_k)}{b_1\cdots b_k}\sum_{\mathbf{e}}\frac{\lambda(b_1d_1e_1,\dots,b_kd_ke_k)}{e_1\cdots e_k}\\ &=\sum_{\mathbf{l}}\frac{\lambda(l_1d_1,\dots,l_kd_k)}{l_1\cdots l_k}\prod_{i=1}^k\sum_{b_i|l_i}\mu(b_i)=\lambda(\mathbf{d}).\end{aligned}$$

□

Now we describe how we will construct the functions $\xi()$ and $\lambda()$. Let $\mathscr{F}_k$ denote the set of functions $F(\mathbf{x})$ that are continuously differentiable, symmetric and supported on the set

$$\mathcal{R}_k = \{\mathbf{x}\in[0,1]^k : x_1+\cdots+x_k\leqslant 1\}, \tag{7.35}$$

and such that

$$|F(\mathbf{x})|\leqslant 1 \qquad (\mathbf{x}\in\mathcal{R}_k), \tag{7.36}$$

and F is not identically zero on $\mathcal{R}_k$. For such an F, we then take

$$\xi(\mathbf{r}) = \mathbb{1}_{\mathbf{r}\in\mathscr{D}}\mu(r_1)\cdots\mu(r_k)\underbrace{\prod_{z<p\leqslant R}(1+k/p)^{-1}}_{\text{constant}}F\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right). \tag{7.37}$$

In this way, by Lemma 7.11, for any $\mathbf{d}\in\mathscr{D}$ we have

$$\left|\lambda(\mathbf{d})\prod_{z<p\leqslant R}(1+k/p)\right| \leqslant \sum_{\mathbf{b}\in\mathscr{D}}\frac{1}{b_1\cdots b_k}\leqslant\sum_{\substack{P^-(\ell)>z\\P^+(\ell)\leqslant R}}\frac{\mu^2(\ell)k^{\omega(\ell)}}{\ell}=\prod_{z<p\leqslant R}(1+k/p),$$

that is, $|\lambda(\mathbf{d})| \leqslant 1$ for all $\mathbf{d} \in \mathscr{D}$, as required by (7.36).

We next insert the definition (7.37) into Propositions 7.8 and 7.9, thus relating the sums in question to integrals over F. We first need some general tools, first relating a sum over a 1-variable function to an integral, and using it to handle multivariate functions.

Lemma 7.12. *Suppose $g \in C^1[a,b]$, $0 \leqslant a < b \leqslant 1$, with $|g(t)| \leqslant 1$ and $|g'(t)| \leqslant \log R$ throughout $[a,b]$. Suppose that $e^{\sqrt{\log R}} \leqslant z \leqslant R$ and write $R = z^u$. Then*

$$\sum_{\substack{R^a < n \leqslant R^b \\ P^-(n) > z}} \frac{1}{n} g\left(\frac{\log n}{\log R}\right) = e^{-\gamma} \frac{\log R}{\log z} \left(\int_a^b g(t)\, dt + O\left(\frac{\log u}{u}\right)\right).$$

Proof. Let $\delta = \frac{\log u}{u}$ and $a' = \max(\delta, a)$. We separate the sum into two parts:

$$S_1 = \sum_{\substack{R^a < n \leqslant R^{a'} \\ P^-(n) > z}} \frac{1}{n} g\left(\frac{\log n}{\log R}\right), \qquad S_2 = \sum_{\substack{R^{a'} < n \leqslant R^b \\ P^-(n) > z}} \frac{1}{n} g\left(\frac{\log n}{\log R}\right),$$

If $\delta < a$ then $S_1 = 0$, otherwise we have the crude bound

$$|S_1| \leqslant \sum_{\substack{P^+(n) \leqslant R^\delta \\ P^-(n) > z}} \frac{1}{n} \ll \frac{\log R^\delta}{\log z} = \delta \frac{\log R}{\log z}.$$

We use partial summation on S_2, first writing

$$S_2 := \sum_{\substack{R^{a'} < n \leqslant R^b \\ P^-(n) > z}} \frac{1}{n} g\left(\frac{\log n}{\log R}\right) = \int_{R^{a'}}^{R^b} \frac{1}{t} g\left(\frac{\log t}{\log R}\right) d\Phi(t,z).$$

By Hypothesis, $\frac{\log t}{\log z} \leqslant \sqrt{\log t}$ for $t \leqslant R$. Hence, by Theorem 3.7 (iii), when $t \geqslant R^\delta = z^{u\delta}$ we have

$$\Phi(t,z) = \frac{e^{-\gamma} t}{\log z} + E(t), \quad E(t) = O\left(\frac{t}{e^{u\delta} \log z}\right) = O\left(\frac{t}{u \log z}\right).$$

Therefore,

$$\begin{aligned} S_2 &= \frac{e^{-\gamma}}{\log z} \int_{R^{a'}}^{R^b} g\left(\frac{\log t}{\log R}\right) \frac{dt}{t} + \frac{E(t) g(\frac{\log t}{\log R})}{t}\Bigg|_{R^{a'}}^{R^b} - \int_{R^{a'}}^{R^b} E(t) \left[\frac{g'(\frac{\log t}{\log R})}{t^2 \log R} - \frac{g(\frac{\log t}{\log R})}{t^2}\right] dt \\ &= \frac{e^{-\gamma} \log R}{\log z} \int_{a'}^b g(y)\, dy + O\left(\frac{\log R}{u \log z}\right). \end{aligned}$$

Recalling that $\delta = \frac{\log u}{u}$ and our bound on S_1, the lemma follows; if $\delta < a$ then $S = S_2$ and otherwise we use that $\int_a^{a'} |g| \leqslant \int_0^\delta 1 = \delta$. □

Lemma 7.13. *Suppose that $f \in C^1(\mathcal{R}_k)$, $|f(\mathbf{x})| \leqslant 1$ on $\mathcal{R}_k$, and*

$$\max_{\mathbf{x} \in \mathcal{R}_k} \max_i \left| \frac{\partial f}{\partial x_i}(\mathbf{x}) \right| \leqslant \log R.$$

Let $R = z^u$, where $3 \leqslant u \leqslant \sqrt{\log R}$. then

$$\sum_{\substack{P^-(n_1\cdots n_k)>z\\ \mu^2(n_1\cdots n_r)=1}} \frac{f(\frac{\log n_1}{\log R},\dots,\frac{\log n_k}{\log R})}{n_1\cdots n_k} = \left(e^{-\gamma}\frac{\log R}{\log z}\right)^k \left[\int_{\mathcal{R}_k} f + O_k\left(\frac{\log u}{u}\right)\right].$$

Proof. We first insert "missing" summands corresponding to $\mu^2(n_1\cdots n_k) = 0$. These have total at most

$$\sum_{\substack{P^-(n_1\cdots n_k)>z\\ \mu^2(n_1\cdots n_k)=0}} \frac{1}{n_1\cdots n_k} \leqslant k^2 \sum_{p>z}\frac{1}{p^2}\Bigg(\sum_{\substack{P^-(m)>z\\ m\leqslant R}}\frac{1}{m}\Bigg)^k \ll_k \frac{1}{z}\left(\frac{\log R}{\log z}\right)^k,$$

which is tiny. Repeated application of Lemma 7.12 (each time with $a = 0$) shows that

$$\sum_{P^-(n_1\cdots n_k)>z} \frac{f(\frac{\log n_1}{\log R},\dots,\frac{\log n_k}{\log R})}{n_1\cdots n_k} = \left(\frac{e^{-\gamma}\log R}{\log z}\right)^k \left[\int_{\mathcal{R}_k} f + O_k\left(\frac{\log u}{u}\right)\right],$$

which implies the desired conclusion. □

Proposition 7.14. *Let $F \in \mathscr{F}_k$ with $I(F) > 0$ and $J(F) > 0$. Define ξ by* (7.37).

(i) Under the hypotheses of Proposition 7.8, we have

$$\sum_{N<n\leqslant 2N} w(n) \sim VN\left(e^{-\gamma}\frac{\log z}{\log R}\right)^k I(F) \qquad (N\to\infty),$$

where

$$I(F) = \int_{\mathcal{R}_k} F^2(\mathbf{x})\,d\mathbf{x}. \tag{7.38}$$

(ii) Under the hypotheses of Proposition 7.9, we have

$$\sum_{N<n\leqslant 2N}\sum_{m=1}^{k} w(n)\mathbb{1}_{a_mn+b_m \text{ prime}} \sim VN\left(e^{-\gamma}\frac{\log z}{\log R}\right)^k \frac{k\theta}{2}J(F) \qquad (N\to\infty),$$

where

$$J(F) = \int_{x_2,\dots,x_k}\cdots\int\left(\int F(\mathbf{x})\,dx_1\right)^2 dx_2\cdots dx_n. \tag{7.39}$$

Proof. From the definitions (7.7), we have $R = z^u$, where

$$u = s^2\left(\frac{\theta}{2} - \frac{1}{s}\right) \asymp s^2.$$

By hypothesis, all first-order partial derivatives of F are bounded, and the same is true of the functions $F(\mathbf{x})^2$ and $(\int F(\mathbf{x})\,dx_1)^2$.

(i) From Proposition 7.8, the definition (7.37) of ξ and Lemma 7.13, we have

$$\sum_{N<n\leqslant 2N} w(n) = \sigma^2 VN\left(e^{-\gamma}\frac{\log R}{\log z}\right)^k I(F)\left(1 + O\left(\frac{\log s}{s^2}\right)\right) + O\left(\frac{N}{(\log N)^{99k}}\right),$$

where

$$\sigma = \prod_{z<p\leqslant R} (1+k/p)^{-1} \sim \left(\frac{\log z}{\log R}\right)^k.$$

The final error term is negligible since

$$V \geqslant \prod_{p\leqslant 2k}\left(1-\frac{p-1}{p}\right)\prod_{2k<p\leqslant z}\left(1-\frac{k}{p}\right) \gg_k \frac{1}{(\log N)^k}. \tag{7.40}$$

(ii) From the symmetry of F, we see that $\zeta_m(\mathbf{r})$ is independent of m. By Lemma 7.10,

$$\sum_{\mathbf{r}} \frac{\zeta_1^2(\mathbf{r})}{r_1\cdots r_k} = \sigma^2 \sum_{r_2,\dots,r_k} \frac{1}{r_2\cdots r_k}\left(\sum_{r_1}\frac{1}{r_1}F\left(\frac{\log r_1}{\log R},\dots,\frac{\log r_k}{\log R}\right)\right)^2.$$

For each i set $x_i = \frac{\log r_i}{\log R}$. With $x_2,\dots,x_k$ fixed we have $0\leqslant x_1\leqslant 1-(x_2+\cdots+x_k)$.

By Lemma 7.13, the sum on r_1 is equal to

$$e^{-\gamma}\frac{\log R}{\log z}\left[\int_0^{1-x_2-\cdots-x_k} F(\mathbf{x})\,dx_1 + O_F\left(\frac{\log s}{s^2}\right)\right].$$

Lemma 7.13, applied to the function $f(x_2,\dots,x_k) = (\int F(\mathbf{x})\,dx_1)^2$, then implies that

$$\sum_{\mathbf{r}} \frac{\zeta_1^2(\mathbf{r})}{r_1\cdots r_k} \sim \sigma^2\left(e^{-\gamma}\frac{\log R}{\log z}\right)^{k+1} J(F).$$

Recalling Proposition 7.9, we have

$$\sum_{m=1}^k \sum_{N<n\leqslant 2N} w(n)\mathbb{1}_{a_mn+b_m \text{ prime}} \sim \frac{kVN/\log N}{\prod_{p\leqslant z}(1-1/p)}\sum_{\mathbf{r}\in\mathscr{D}}\frac{\zeta_1(\mathbf{r})^2}{r_1\cdots r_k}.$$

Now $\prod_{p\leqslant z}(1-1/p)\sim e^{-\gamma}/\log z$ by Mertens, and $\frac{\log R}{\log N}\sim\frac{\theta}{2}$, and the desired asymptotic follows. □

Armed with Proposition 7.14, we can now apply these results to prime gaps and prime k-tuples.

Theorem 7.15. *Fix $k\in\mathbb{N}$ and $m\in\mathbb{N}$, and assume Hypothesis EH(θ). There is a constant c_k such that the following holds. Suppose that there is a function $F\in\mathscr{F}_k$ and such that*

$$M_k(F) := \frac{kJ(F)}{I(F)} > \frac{2(m-1)}{\theta},$$

where $I(F), J(F)$ are defined in (7.38) *and* (7.39)*, respectively. Then, for all sufficiently large N, and all admissible tuples $(a_1n+b_1,\dots,a_kn+b_k)$ of linear forms, with*

$$1\leqslant a_i\leqslant \log^{100}N, \qquad -a_iN/2 < b_i \leqslant N\log^{100}N \quad (1\leqslant i\leqslant k),$$

there are $\gg N/(\log N)^{c_k}$ integers $n\in(N,2N]$ for which at least m of the numbers $a_1n+b_1,\dots,a_kn+b_k$ are prime.

In particular, unconditionally the conclusion holds provided that $M_k(F)>4(m-1)$.

Proof. Let

$$S = \sum_{N<n\leqslant 2N} v(n)w(n), \qquad v(n) := \sum_{j=1}^k \mathbb{1}_{a_jn+b_j \text{ prime}} - (m-1).$$

By the two parts of Proposition 7.14, we have

$$S \sim VN\left(e^{-\gamma}\frac{\log z}{\log R}\right)^k\left[\frac{kJ(F)\theta}{2} - (m-1)I(F)\right].$$

By hypothesis, the bracketed expression is positive. From (7.40) we have

$$S' := \sum_{\substack{N<n\leqslant 2N\\ v(n)\geqslant 1}} w(n) \geqslant \sum_{N<n\leqslant 2N}\frac{v(n)w(n)}{k} = \frac{S}{k} \gg \frac{N}{(\log N)^k}.$$

It follows that there are many values of n with $v(n) > 0$. More precisely, by Cauchy-Schwarz,

$$(S')^2 \leqslant \#\{N < n \leqslant 2N : v(n) \geqslant 1\}\Big(\sum_{N<n\leqslant 2N} w(n)^2\Big).$$

Thus,

$$\#\{N < n \leqslant 2N : v(n) \geqslant 1\} \gg \frac{N^2}{(\log N)^{2k}}\Big(\sum_{N<n\leqslant 2N} w(n)^2\Big)^{-1} \tag{7.41}$$

By (7.12), we have the crude bound

$$0 \leqslant w(n) \leqslant \prod_{j=1}^{k}\tau(a_in+b_i)^2.$$

Hence, by Hölder's inequality and the bound $a_in + b_i \leqslant 3N(\log N)^{100}$,

$$\begin{aligned}\sum_{N<n\leqslant 2N} w(n)^2 &\leqslant \sum_{N<n\leqslant 2N}\prod_{j=1}^{k}\tau(a_in+b_i)^4\\ &\leqslant \prod_{j=1}^{k}\Big(\sum_{N<n\leqslant 2N}\tau(a_jn+b_j)^{4k}\Big)^{1/k}\\ &\leqslant \sum_{d\leqslant 3N(\log N)^{100}}\tau^{4k}(d)\\ &\ll N(\log N)^{100+2^{4k}-1}.\end{aligned}$$

Combined with (7.41), this completes the proof of the first claim, with $c_k = 2k + 100 + 2^{4k} - 1$.

The final statement follows from the Bombieri-Vinogradov Theorem, which implies Hypothesis EH(θ) for all $\theta < 1/2$. □

7.4. **Bounds on $M_k(F)$.**

Definition 6. *Let*

$$M_k := \sup_{F\in\mathscr{F}_k} M_k(F) = \sup_{F\in\mathscr{F}_k}\frac{kJ(F)}{I(F)}.$$

We begin with a simple upper bound for M_k from [**120**].

Theorem 7.16 ([**120**])**.** *We have* $M_k \leqslant \frac{k}{k-1}\log k = \log k + O(\frac{\log k}{k})$.

Proof. Let $x_2, \dots, x_k$ be given and set $y = 1 - (x_2 + \cdots + x_k)$. The integral over those $\mathbf{x}$ with $y = 0$ has measure zero and can be discarded. Thus, we may assume that $y > 0$. For any $A > 0$, by Cauchy-Schwarz we have

$$\begin{aligned}\left(\int_0^y F(\mathbf{x})\, dx_1 \right)^2 &\leqslant \int_0^y (1 + Ax_1) F(\mathbf{x})^2\, dx_1 \int_0^y \frac{dx_1}{1 + Ax_1} \\ &= \frac{\log(1 + Ay)}{A} \int_0^y (1 + Ax_1) F(\mathbf{x})^2\, dx_1.\end{aligned}$$

Put $A = B/y$ for some constant B to be determined, then integrate over $x_2, \dots, x_k$. From $y(1 + Ax_1) = y + Bx_1$ we get

$$J(F) \leqslant \frac{\log(1 + B)}{B} \int_{\mathcal{R}_k} (1 - x_2 - \cdots - x_k + Bx_1) F(\mathbf{x})^2\, d\mathbf{x}.$$

Since $F(\mathbf{x})$ is symmetric, we may interchange x_1 and x_i. Doing this for each i and summing gives

$$kJ(F) \leqslant \frac{\log(1 + B)}{B} \int_{\mathcal{R}_k} (k - (k-1) \sum x_i + B \sum x_i) F(\mathbf{x})^2\, d\mathbf{x}.$$

Taking $B = k - 1$, the integral above equals $kI(F)$ and this completes the proof. □

Theorem 7.17. *We have*

$$M_k \geqslant \log k - \log_2 k - O(1).$$

Proof. Let $\delta > 0$, and let $g : [0, \delta] \to [0, \infty)$ be a function with $|g(t)| \leqslant 1$ for all t. We'll specialize to functions of the type

$$F(\mathbf{x}) = g(x_1) \cdots g(x_k) \mathbb{1}(x_1 + \cdots + x_k \leqslant 1).$$

For short, let

$$m_1 = \int_0^\delta g(t)\, dt, \qquad m_2 = \int_0^\delta g^2(t)\, dt.$$

We interpret $I(F)$ and $J(F)$ probabilistically. Let $Z_1, \dots, Z_k$ be independent random variables with density function $\frac{1}{m_2} g^2(t)$, $0 \leqslant t \leqslant \delta$. Then

$$I(F) = \int_{\mathcal{R}_k} g^2(x_1) \cdots g^2(x_k)\, d\mathbf{x} = m_2^k\, \mathbb{P}(Z_1 + \cdots + Z_k \leqslant 1).$$

Also,

$$\begin{aligned}J(F) &\geqslant \int \cdots \int_{x_2 + \cdots + x_k \leqslant 1 - \delta} g^2(x_2) \cdots g^2(x_k) \left(\int_0^\delta g(x_1)\, dx_1 \right)^2 dx_2 \cdots dx_k \\ &= m_1^2 m_2^{k-1}\, \mathbb{P}(Z_2 + \cdots + Z_k \leqslant 1 - \delta).\end{aligned}$$

Thus,

$$M_k(F) \geqslant \frac{k m_1^2}{m_2} \cdot \frac{\mathbb{P}(Z_2 + \cdots + Z_k \leqslant 1 - \delta)}{\mathbb{P}(Z_1 + \cdots + Z_k \leqslant 1)} \geqslant \frac{k m_1^2}{m_2} \mathbb{P}(Z_2 + \cdots + Z_k \leqslant 1 - \delta). \tag{7.42}$$

We will choose g so that the probability on the RHS is about 1. If

$$\mu := \mathbb{E} Z_i = \frac{1}{m_2} \int_0^\delta t g^2(t)\, dt,$$

then we want

$$\mu < \frac{1-\delta}{k-1} \tag{7.43}$$

so that

$$\mathbb{E}(Z_2 + \cdots + Z_k) = (k-1)\mu < 1 - \delta.$$

We can bound the probability of a large deviation from the mean using Chebyshev's inequality and the calculated variance

$$\sigma^2 := \mathbb{E}(Z_i - \mu)^2 = \mathbb{E}Z_i^2 - \mu^2 = \frac{1}{m_2}\int_0^\delta t^2 g^2(t)\,dt - \mu^2.$$

Then

$$\mathbb{E}(Z_2 + \cdots + Z_k - (k-1)\mu)^2 = (k-1)\sigma^2.$$

Thus, under the assumption of (7.43), we have

$$\begin{aligned}\mathbb{P}\left(Z_2 + \cdots + Z_k \geqslant 1-\delta\right) &\leqslant \mathbb{P}\Big(|Z_2 + \cdots + Z_k - (k-1)\mu| \geqslant 1 - \delta - (k-1)\mu\Big)\\ &\leqslant \frac{(k-1)\sigma^2}{(1-\delta-(k-1)\mu)^2}.\end{aligned} \tag{7.44}$$

We will choose parameters so that the right side of (7.44) is very small.

It then remains to maximize m_1^2/m_2 subject to $\mu \lesssim 1/k$. Motivated by the proof of Theorem 7.16, for any $A > 0$, Cauchy-Schwarz gives

$$m_1^2 \leqslant \left(\int_0^\delta (1+At)g^2(t)\,dt\right)\left(\int_0^\delta \frac{dt}{1+At}\right) = (m_2 + Am_2\mu)\frac{\log(1+A\delta)}{A}.$$

Moreover, equality holds only for $g(t) = \frac{1}{1+At}$. That is,

$$\frac{m_1^2}{m_2} \leqslant \min_{A>0}\left(\frac{1}{A} + \mu\right)\log(1+A\delta). \tag{7.45}$$

The minimum occurs when

$$\left(\frac{1}{A} + \mu\right)\frac{\delta}{1+A\delta} = \frac{\log(1+A\delta)}{A^2},$$

and then taking $g(t) = \frac{1}{1+At}$ gives equality in (7.45). With this function $g(t)$ and assuming that $A\delta$ is large, we compute

$$m_1 = \frac{\log(1+A\delta)}{A}, \qquad m_2 = \frac{\delta}{1+A\delta} \approx \frac{1}{A},$$

and

$$\mu = \frac{1}{m_2 A^2}\left(\log(1+A\delta) - 1 + \frac{1}{1+A\delta}\right) \approx \frac{\log(1+A\delta)}{A}.$$

We will take a convenient choice of A, δ satisfying the simpler relation

$$A = k\log(1+A\delta). \tag{7.46}$$

Then

$$\mu = \frac{1}{k} - \frac{1}{A} + \frac{1}{kA\delta}, \quad m_1 = \frac{1}{k}, \quad 1/m_2 = A + 1/\delta,$$

and hence

$$\frac{m_1^2}{m_2} = \frac{1}{k^2}\left(A + \frac{1}{\delta}\right) > \frac{A}{k^2}.$$

Thus, by (7.42) and (7.44),

$$M_k(F) \geqslant \frac{A}{k}\left(1 - \frac{(k-1)\sigma^2}{(1-\delta-(k-1)\mu)^2}\right). \tag{7.47}$$

Crudely,

$$\sigma^2 \leqslant \mathbb{E}Z_i^2 \leqslant \frac{1}{m_2}\left(\frac{A\delta}{1+A\delta}\right)^2 \int_0^\delta \frac{t^2\,dt}{(At)^2} = \frac{\delta^2}{1+A\delta} < \frac{\delta}{A}.$$

Also, using (7.46), we compute (after some algebra)

$$1-\delta-(k-1)\mu = (k-e^{A/k})\left(\frac{1}{A} - \frac{1}{k(e^{A/k}-1)}\right).$$

In particular, for this to be positive, we need $A < k\log k$. Combined with (7.47), we conclude that

$$M_k(F) \geqslant \frac{A}{k}\left(1 - \frac{(k-1)(e^{A/k}-1)}{(k-e^{A/k})^2(1-\frac{A}{k(e^{A/k}-1)})^2}\right).$$

A good choice for A is

$$A = k(\log k - \log_2 k).$$

(this gives $\delta \sim \frac{1}{\log^2 k}$). Then $e^{A/k} = k/\log k$ and we see that

$$M_k \geqslant (\log k - \log_2 k)\left(1 - O\left(\frac{1}{\log k}\right)\right) = \log k - \log_2 k + O(1),$$

as required. □

Together with Theorem 7.15, Theorem 7.17 immediately implies Theorem 7.6. To obtain the claimed bounds on K_m, notice that by Theorem 7.17 there is a $k \ll me^{4m}$ with $M_k > 4(m-1)$.

The PolyMath8b project provides better bounds on M_k, as well as numerical bounds when k is small.

Theorem [**120**, Theorem 3.9 (vii), (xi)]. We have

(i) $M_{54} > 4$;
(ii) $M_k = \log k - O(1)$.

In what follows, it is convenient to adopt a definition of the minimum of numbers K_m that imply a weaker form of the conclusion Theorem 7.6.

Definition. For integers $m \geqslant 2$, $\widetilde{K}_m$ is the least number k so that for any system of admissible system of linear forms $(a_1n+b_1,\dots,a_kn+b_k)$, there are infinitely many n such that at least k of the forms are prime.

The prime $k-$tuples conjecture implies that $\widetilde{K}_m = m$. By a modification of the analysis in this section, it was proved in [**120**] that $\widetilde{K}_2 \leqslant 50$.

Under the assumption of the Elliott-Halberstam Conjecture (Conjecture 7.3), Maynard [**108**] showed that $M_5 > 2$ and hence $\widetilde{K}_2 \leqslant 5$, improving the bound $\widetilde{K}_2 \leqslant 6$ proved by Goldston, Pintz and Yıldırım[**64**] under the same hypothesis. A generalized version of the Elliot-Halberstam conjecture (in the notation of [**120**], this is the statement that GEH[θ] holds for all $\theta < 1$), implies that $\widetilde{K}_2 \leqslant 3$ [**120**, Theorem 16 (xii)].

We next show some applications using linear forms a_in+b_i with $a_i \neq 1$. Central to the topic is the following special case of the prime k-tuples conjecture.

Hypothesis $\mathscr{P}(a,b)$: there are infinitely many $n \in \mathbb{N}$ such that both $an+1$ an $bn+1$ are prime.

We note that for any set $\{a_1,\dots,a_k\}$ of k positive integers, the k-tuple of forms $(a_1n+b_1,\dots,a_kn+b_k)$ is admissible, since for any prime p, if $p|n$ then $p \nmid (a_1n+1)\cdots(a_kn+1)$. We then have

Theorem 7.18. *If $k \geqslant \widetilde{K}_2$ and $\{a_1,\dots,a_k\}$ is an set of k positive integers, then there are $i \neq j$ such that Hypothesis $\mathscr{P}(a_i,a_j)$ holds.*

7.5. **Consecutive integers with large prime factors.** One special case of the prime k-tuples conjecture, Conjecture 1.1 states that for infinitely many primes p, $p+1=2q$ for a prime q. In this direction, we prove

Theorem 7.19 (K. Ford and J. Maynard, 2014; unpublished)**.** *There is a constant B so that for infinitely many $n \in \mathbb{N}$, $P^+(n) \geqslant n/B$ and $P^+(n+1) \geqslant (n+1)/B$.*

Lemma 7.20 (Heath-Brown [**86**])**.** *For any k there is a set of k positive integers $a_1,\dots,a_k$ such that*

$$a_i - a_j = (a_i,a_j) \qquad (i>j). \tag{7.48}$$

Examples are

$$\begin{aligned} k=4: &\quad \{6,8,9,12\}\\ k=5: &\quad \{60,63,64,66,72\}. \end{aligned}$$

The condition (7.48) is equivalent to $(a_i-a_j)|a_i$ for all $i \neq j$. Thus, Lemma 7.20 is one of the assertions in Lemma 1 of Heath-Brown [**86**]. As we do not require the other properties of the a_i from [**86**], our proof is much shorter.

Proof. By induction. If $\{a_1,\dots,a_k\}$ satisfies (7.48), then so does the $k+1$ element set

$$\{M, M-a_1,\dots,M-a_k\},$$

where M is the least common multiple of the numbers $a_1,\dots,a_k$ and the differences $|a_i-a_j|$ for $i<j$. □

Proof of Theorem 7.19. Let $k \geqslant \widetilde{K}_2$ and $\{a_1,\dots,a_k\}$ satisfy (7.48). By Theorem 7.18, Hypothesis $\mathscr{P}(a_i,a_j)$ holds for some $i<j$. Suppose that $n \in \mathbb{N}$, and a_in+1 and a_jn+1 are both prime. By (7.48), $a_j-a_i=(a_i,a_j)$ and hence

$$\frac{a_j}{(a_i,a_j)}(a_in+1) \quad \text{and} \quad \frac{a_i}{(a_i,a_j)}(a_jn+1)$$

are consecutive integers. Thus the theorem follows with

$$B = \max_{i\neq j} \frac{a_i}{(a_i,a_j)}.$$

□

7.6. **Locally repeated values of Euler's function.** We partially solve a longstanding conjecture about the solubility of

$$\phi(n+k)=\phi(n), \tag{7.49}$$

where ϕ is Euler's function and k is a fixed positive integer.

Hypothesis $\mathcal{S}_k$. The equation (7.49) holds for infinitely many n.

Erdős conjectured in 1945 that for any m, the simultaneous equations

$$\phi(n)=\phi(n+1)=\cdots=\phi(n+m-1) \tag{7.50}$$

has infinitely many solutions n. If true, this would immediately imply hypothesis $\mathcal{S}_k$ for every k. However, there is only one solution of (7.50) known when $m\geqslant 3$, namely $n=5186$, $m=3$. In 1956, Sierpiński [**138**] showed that for any k, (7.49) has at least one solution n (e.g. take $n=(p-1)k$, where p is the smallest prime not dividing k). This was extended by Schinzel [**129**] and by Schinzel and Wakulicz [**132**], who showed that for any $k\leqslant 2\cdot 10^{58}$ there are at least two solutions of (7.49). In 1958, Schinzel [**129**] explicitly conjectured that $\mathcal{S}_k$ is true for every $k\in\mathbb{N}$. There is good numerical evidence for $\mathcal{S}_k$, at least when $k=1$ or k is even [**68**]. Information about solutions for $k\in\{1,2,3,4,5,6,7,8,9,10,11,12\}$ can also be found in OEIS [**115**] sequences A001274, A001494, A330251, A179186, A179187, A179188, A179189, A179202, A330429, A276503, A276504 and A217139, respectively. Below 10^{11} there are very few solutions of (7.49) when $k\equiv 3 \pmod 6$ [**68**], e.g. only the two solutions $n\in\{3,5\}$ for $k=3$ are known. A further search by G. Resta (see [**115**], sequence A330251) reveals 17 more solutions in $[10^{12},10^{15}]$.

There is a close connection between Hypothesis $\mathcal{S}_k$ for even k and generalized prime twins.

Schinzel [**129**] observed that Hypothesis $\mathcal{P}(1,2)$ implies $\mathcal{S}_k$ for every even k. The proof is simple: if $n+1$ and $2n+1$ are prime and larger than k, then

$$\phi(k(2n+1))=\phi((n+1)2k)=2n\phi(k).$$

Graham, Holt and Pomerance [**68**] generalized this idea, showing the following.

Lemma 7.21 ([**68**, Theorem 1])**.** *For any k and any number j such that j and $j+k$ have the same prime factors, Hypothesis $\mathscr{P}(\frac{j}{(j,j+k)},\frac{j+k}{(j,j+k)})$ implies $\mathcal{S}_k$.*

This also has an easy proof: if $\frac{j}{(j,j+k)}r+1$ and $\frac{j+k}{(j,j+k)}r+1$ are both prime, then $n=j(\frac{j+k}{(j,j+k)}r+1)$ satisfies (7.49). Note that for odd k there are no such numbers j, and for each even k there are finitely many such j (see [**68**], Section 3). Extending a bound of Erdős, Pomerance and Sárkőzy [**45**] in the case $k=1$, Graham, Holt and Pomerance showed that the solutions of (7.49) *not* generated from Lemma 7.21 are very rare, with counting function $O_k(x\exp\{-(\log x)^{1/3}\})$. Yamada [**151**] sharpened this bound to $O_k(x\exp\{-(1/\sqrt{2}+o(1))\sqrt{\log x\log\log\log x}\})$. Assuming the Hardy-Littlewood conjectures [**78**], when k is even we conclude that there are $\sim C_k x/\log^2 x$ solutions $n\leqslant x$ of (7.49), where $C_k>0$.

Using the fact that $\widetilde{K}_2\leqslant 50$, we will prove the following.

Theorem 7.22 (Ford [**52**], 2020)**.** *We have*

(a) *For any k that is a multiple of* 44720643463713815200, *$\mathcal{S}_k$ is true;*

(b) *There is some even $\ell\leqslant 3570$ such that $\mathcal{S}_k$ is true whenever $\ell|k$; consequently, the number of $k\leqslant x$ for which $\mathcal{S}_k$ is true is at least $x/3570$.*

Improvements to $\widetilde{K}_2$ allow us to improve significantly on Theorem 7.22.

Theorem 7.23 (Ford [**52**], 2020)**.** *If $\widetilde{K}_2 \leqslant 5$, then $\mathcal{S}_k$ is true for all k with $30|k$. If $\widetilde{K}_2 \leqslant 4$, then $\mathcal{S}_k$ is true for all k with $6|k$. In particular, the Elliot-Halberstam conjecture implies $\mathcal{S}_k$ for all k with $30|k$, and the Generalized Elliot-Halberstam conjecture implies $\mathcal{S}_k$ for all k with $6|k$.*

Incidentally, the conclusion of Theorem 7.23 when $\widetilde{K}_2 \leqslant 4$ is not improved if we have the stronger bound $\widetilde{K}_2 \leqslant 3$.

Using Theorem 7.6, we also make progress toward Erdős' conjecture that (7.50) has infinitely many solutions.

Theorem 7.24 (Ford [**52**], 2020)**.** *For any $m \geqslant 3$ there is a tuple of distinct positive integers $h_1, \dots, h_m$ so that for any $\ell \in \mathbb{N}$, the simultaneous equations*

$$\phi(n+\ell h_1) = \phi(n+\ell h_2) = \cdots = \phi(n+\ell h_m)$$

have infinitely many solutions n.

Now we get to the proofs. Throughout, $1 \leqslant a < b$ are integers. We first show that $\mathscr{P}(a,b)$ implies $\mathcal{S}_k$ for certain k, inverting Lemma 7.21. Define

$$\kappa(a,b) = (b'-a') \prod_{p|a'b'} p, \qquad a' = \frac{a}{(a,b)},\ b' = \frac{b}{(a,b)}. \tag{7.51}$$

We observe that $\kappa(a,b)$ is always even.

Lemma 7.25. *Assume $\mathscr{P}(a,b)$. Then $\mathcal{S}_k$ holds for every k which is a multiple of $\kappa(a,b)$.*

Proof. Define $a' = \frac{a}{(a,b)}$, $b' = \frac{b}{(a,b)}$ and observe that $\mathscr{P}(a,b) \Rightarrow \mathscr{P}(a',b')$. Let $s = \prod_{p|a'b'} p$, and suppose that $r \in NN$ is such that $a'r+1$ and $b'r+1$ are both prime. Let $\ell \in \mathbb{N}$ and set

$$m_1 = b'\ell s(a'r+1), \quad m_2 = a'\ell s(b'r+1).$$

As all of the prime factors of $a'b'$ divide ℓs, we have $\phi(b'\ell s) = b'\phi(\ell s)$ and $\phi(a'\ell s) = a'\phi(\ell s)$. If $r > a'b'\ell$, it follows than $\phi(m_1) = \phi(m_2)$. Finally, $m_1 - m_2 = (b'-a')\ell s = \ell\kappa(a,b)$. □

Proof of Theorem 7.22. Let

$$\{a_1, \dots, a_{50}\} = \{1, 2, 4, 5, 6, \dots, 48, 49, 52, 56\},$$

By $\widetilde{K}_2 \leqslant 50$ and Theorem 7.18, for some i, j with $1 \leqslant i < j \leqslant 50$, $\mathscr{P}(a_i, a_j)$ is true. We compute

$$\operatorname{lcm}\{\kappa(a_i,a_j) : 1 \leqslant i < j \leqslant 50\} = 44272064346371381 5200 = 2^5 3^3 5^2 \prod_{7 \leqslant p \leqslant 47} p,$$

and thus (a) follows from Lemma 7.25.

For part (b), we take

$$\begin{aligned}\{a_1, \dots, a_{50}\} = \{&15, 20, 30, 36, 40, 45, 60, 72, 75, 80, 90, 96, 100, 108, 120, 135, 144, 150, 180, 192, 200,\\ &216, 225, 240, 250, 270, 288, 300, 320, 324, 360, 375, 384, 400, 405, 450, 480, 500, 540, 600,\\ &720, 750, 810, 900, 960, 1080, 1200, 1440, 1500, 1800\},\end{aligned}$$

numbers that only have prime factors $2, 3, 5$. We also compute that

$$\max_{1\leqslant i<j\leqslant 50} \kappa(a_i, a_j) = 3570,$$

and again invoke Lemma 7.25. This proves (b). □

We believe that 3570 is the smallest number than can be produced for Theorem 7.22 (b), and have conducted extensive computer searches for sets $\{a_1, \dots, a_{50}\}$ with smaller value of $\max_{i\neq j} \kappa(a_i, a_j)$.

Proof of Theorem 7.23. Same as the proof of Theorem 7.22 (a), but take $\{a_1, a_2, a_3, a_4, a_5\} = \{1, 2, 3, 4, 6\}$ if $\widetilde{K}_2 \leqslant 5$ and $\{a_1, \dots, a_4\} = \{1, 2, 3, 4\}$ if $\widetilde{K}_2 \leqslant 4$. □

Proof of Theorem 7.24. Let $m \geqslant 2$, $k = \widetilde{K}_m$ and consider any set $\{a_1, a_2, \dots, a_k\}$ of k positive integers. Then there are $1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant k$ such that for infinitely many r, the m numbers $a_{i_1}r + 1, \dots, a_{i_m}r + 1$ are all prime. Let r be such a number. Define

$$h_j = \frac{(a_{i_1} \cdots a_{i_m})^2}{a_{i_j}} \qquad (1 \leqslant j \leqslant m).$$

Let $\ell \in \mathbb{N}$ and set $n = \ell(a_{i_1} \cdots a_{i_m})^2 r$. Then, since $a_{i_j} | h_j$ for all j, it follows that if r is sufficiently large then for any j,

$$\phi(n + \ell h_j) = \phi(\ell h_j(a_{i_j}r + 1)) = \phi(\ell h_j)a_{i_j}r = \phi(\ell h_j a_{i_j})r.$$ □

7.7. **Locally repeated values of σ.** One can ask analogous questions about the sum of divisors function $\sigma(n)$. As $\sigma(p) = p + 1$ vs $\phi(p) = p - 1$, oftentimes one can port theorems about ϕ over to σ. This is not the case here, since our results depend heavily on the existence of solutions of

$$a\phi(b) = b\phi(a),$$

which is true if and only if a and b have the same set of prime factors. The analogous equation

$$a\sigma(b) = b\sigma(a) \quad \Leftrightarrow \quad \frac{\sigma(a)}{a} = \frac{\sigma(b)}{b}$$

has more sporadic solutions, e.g. if a, b are both perfect numbers or multiply perfect numbers.

Theorem 7.26 (Ford [**52**], 2020)**.** *For a positive proportion of all $k \in \mathbb{N}$, the equation*

$$\sigma(n) = \sigma(n + k)$$

has infinitely many solutions n.

As we shall see from the proof, there is a specific number A and a finite set $\mathcal{B}$ such that for some element $b \in \mathcal{B}$, the equation $\sigma(n) = \sigma(n + k)$ has infinitely many solutions for all numbers $k = \ell b$ where $(\ell, A) = 1$. Unfortunately, our methods cannot specify any particular k for which the conclusion holds. Our method requires finding, for $t = \widetilde{K}_2$, numbers $a_1, \dots, a_t$ so that

$$\frac{\sigma(a_1)}{a_1} = \cdots = \frac{\sigma(a_t)}{a_t} = y. \tag{7.52}$$

Such collections of numbers are sometimes referred to as "friends" in the literature, e.g. [**118**]. Finding larger collections of a_i satisfying (7.52) leads to stronger conclusions.

Theorem 7.27 (Ford [**52**], 2020)**.** *Let $m \geqslant 2$, let $t = \widetilde{K}_m$ and assume that there is a y and positive integers $a_1, \dots, a_t$ satisfying* (7.52)*. Then there are positive integers $h_1 < h_2 < \cdots < h_m$ so that for a positive proportion of integers ℓ, there are infinitely many solutions of*

$$\sigma(n + \ell h_1) = \cdots = \sigma(n + \ell h_m).$$

It is known [**114**] that for $y = 9$, there is a set of 2095 integers satisfying (7.52). The number $y = 9$ has the largest known multiplicity of $\sigma(n)/n$. Also $\widetilde{K}_2 \leqslant 50$ [**120**], and hence Theorem 7.26 follows from the case $m = 2$ of Theorem 7.27. We cannot make the conclusion unconditional when $m \geqslant 3$, since the best known bounds for $\widetilde{K}_3$ is $\widetilde{K}_3 \leqslant 35410$ [**120**, Theorem 3.2 (ii)]. When $y = 2$, the numbers a_i are perfect numbers, and it is a famous problem since antiquity to determine if there are infinitely many perfect numbers. As of April, 2023, there are 51 known perfect numbers [**70**].

Conjecture 7.28. *For any t, there is an y such that $\sigma(a)/a = y$ has at least t solutions. That is, there are arbitrarily large circles of friends.*

Clearly, Conjecture 7.28 implies the conclusion of Theorem 7.27 for all m. In [**40**], Erdős mentions Conjecture A and states that he doesn't know of any argument that would lead to its resolution. In the opposite direction, Wirsing [**147**] showed that the number of $n \leqslant z$ with $\sigma(n)/n = y$ is is $O(z^{c/\log\log z})$ for some c, uniformly in y. Pollack and Pomerance [**118**] studied the solutions of (7.52), gathering data on pairs, triples and quadruples of friends, but did not address Conjecture A.

Using (7.52) and prime pairs $an-1$ and $bn-1$, one can generate many solutions of $\sigma(n) = \sigma(n+k)$, analogous to Lemma 7.21; see Yamada [**151**, Theorem 1.1]. For example, one can generate solutions with $k = 1$ if there is an integer m with $\sigma(m)/m = \sigma(m+1)/(m+1) = y$ (the ratios need not be integers as claimed in [**151**]). Indeed, if $r > m + 1$, and $rm - 1$ and $r(m+1) - 1$ are both prime, then

$$\sigma(m(r(m+1)-1)) = \sigma(m)r(m+1) = rm(m+1)y,$$
$$\sigma((m+1)(mr-1)) = \sigma(m+1)rm = rm(m+1)y.$$

Yamada [**151**, Theorem 1.2] showed that there are $\ll x \exp\{-(1/\sqrt{2} + o(1))\sqrt{\log x \log\log\log x}\}$ solutions $n \leqslant x$ not generated in this way.

Proof of Theorem 7.27. Let $t = \widetilde{K}_m$ and $a_1, \dots, a_t$ satisfy (7.52). Put $A = \operatorname{lcm}[a_1, \dots, a_t]$ and for each i define $b_i = A/a_i$. By Theorem 7.6 applied to the collection of linear forms $b_i n - 1$, $1 \leqslant i \leqslant t$, there exist $i_1, \dots, i_m$ such that for infinitely many $r \in \mathbb{N}$, the m numbers $b_{i_j} r - 1$ are all prime. Let r be such a number, and let $\ell \in \mathbb{N}$ such that $(\ell, A) = 1$ (this holds for a positive proportion of all ℓ). Let

$$t_j = \ell a_{i_j}(b_{i_j} r - 1) = A\ell r - \ell a_{i_j} \qquad (1 \leqslant j \leqslant m).$$

By (7.52), if r is sufficiently large then for every j we have

$$\sigma(t_j) = \sigma(\ell)\sigma(a_{i_j})b_{i_j} r = \sigma(\ell) y a_{i_j} b_{i_j} r = y\sigma(\ell) A r. \qquad \square$$

8. Large gaps bewteen consecutive primes

8.1. **Introduction.** Let p_n denote the n^{th} prime, and let

$$G(X) := \max_{p_{n+1}\leqslant X}(p_{n+1}-p_n)$$

denote the the maximum gap between consecutive primes less than X. It is clear from the prime number theorem that

$$G(X) \geqslant (1+o(1))\log X,$$

as the *average* gap between the prime numbers which are $\leqslant X$ is $\sim \log X$. Explicit gaps of (at least) this size may be constructed by observing that $P(n)+2, P(n)+3, \dots, P(n)+n$ is a sequence of $n-1$ composite numbers, where $P(n)$ is the product of the primes $\leqslant n$; by the prime number theorem, $P(n) = e^{(1+o(1))n}$.

In 1931, Westzynthius [**146**] proved that infinitely often the gap between consecutive prime numbers can be an arbitrarily large multiple of the average gap. His bound is[5]

$$G(X) \gg \frac{\log X \log_3 X}{\log_4 X}.$$

In 1934, Ricci [**126**] slightly improved this to $G(X) \gg \log X \log_3 X$. In 1935 Erdős [**33**] made a much larger improvement, showing

$$G(X) \gg \frac{\log X \log_2 X}{(\log_3 X)^2}$$

and in 1938 Rankin [**124**] added a quadruple-log, showing

$$G(X) \geqslant (c+o(1))\frac{\log X \log_2 X \log_4 X}{(\log_3 X)^2} \tag{8.1}$$

with $c = \frac{1}{3}$. The constant c was increased several times [**134, 125, 107, 117**], ultimately getting to $c = 2e^{\gamma}$ by Pintz [**117**].

It was a well-known open problem (and Erdős prize problem) to show that one could take $c \to \infty$. In August 2014, in two independent papers of Ford-Green-Konyagin-Tao [**53**] and Maynard [**110**], it was shown that c could be taken to be arbitrarily large. The methods of proof in [**53**] and [**110**] are quite different. The arguments in [**53**] used recent results [**72**] on the number of solutions to linear equations in primes, whereas the arguments in [**110**] instead relied on a multidimensional prime-detecting sieve of the type introduced in [**110**] (but more complicated). Later, in [**54**], a quantitative improvement to Rankin's bound was given, namely

Theorem 8.1 (Ford-Green-Konyagin-Maynard-Tao [**54**, Theorem 1])**.** *For sufficiently large X, one has*

$$G(X) \gg \frac{\log X \log_2 X \log_4 X}{\log_3 X}.$$

Here we present a proof of the original theorem from [**53**] and [**110**] that the c in (8.1) may be taken arbitrarily large (without a specific rate of growth), which is much simpler than that in either of the two original papers. It is based on a hybrid approach, utilizing ideas worked out in [**54**], but without the need to make various estimates quantitative.

[5]Recall that $\log_2 x = \log\log x$, $\log_3 x = \log\log\log x$, and so on.

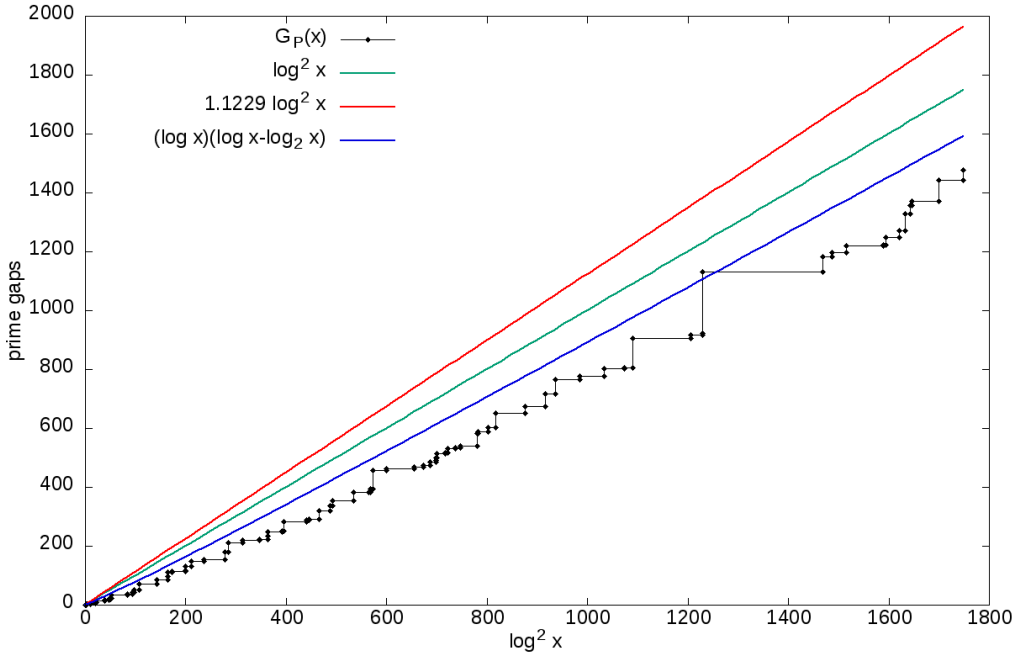

FIGURE 1. $G(x)$ vs. various approximations

Theorem 8.2. *The bound* (8.1) *holds for any real number $c > 0$, if x is large enough, depending on c.*

Based on a probabilistic model of primes, Cramér [**25**] conjectured that

$$\limsup_{X\to\infty} \frac{G(X)}{\log^2 X} = 1.$$

Granville [**69**] offered a refinement of Cramér's model and has conjectured that the lim sup above is in fact at least $2e^{-\gamma} = 1.1229\ldots$. Recently, Banks, Ford and Tao [**9**] made a specific conjecture about the largest gap which is dependent on a problem called the *interval sieve*. We will discuss this in much more detail in Section 9. These conjectures are well beyond the reach of our methods. Numerical evidence is inconclusive, in fact $\max_{X\leqslant 4\cdot 10^{18}} G(X)/\log^2 X \approx 0.9206$, slightly below the predictions of Cramér and Granville. This bound is a consequence of the gap of size 1132 following the prime 1693182318746371. See also Figure 1 for a plot of $G(x)$ versus various approximations.

Unconditional upper bounds for $G(X)$ are far from the conjectured truth, the best being $G(X) \ll X^{0.525}$ and due to Baker, Harman and Pintz [**8**]. Even the Riemann Hypothesis only[6] furnishes the weak bound $G(X) \ll X^{1/2}\log X$ [**24**].

8.1.1. *Notational conventions.* Our arguments will rely heavily on probabilistic arguments using discrete random variables. We will use boldface symbols such as $\mathbf{X}$ or $\mathbf{a}$ to denote random variables (and non-boldface symbols such as X or a to denote deterministic counterparts of these variables). Vector-valued random variables will be denoted in arrowed boldface, e.g. $\vec{\mathbf{a}} = (\mathbf{a}_s)_{s\in\mathcal{S}}$ might denote

[6]Some improvements in the logarithmic factor are available if one also assumes some form of the pair correlation conjecture for zeros of $\zeta(s)$; see [**84**].

a random tuple of random variables $\mathbf{a}_s$ indexed by some index set $\mathcal{S}$. The letters p and q will always denote primes.

8.2. **Overall plan of the proof.** We construct many consecutive integers in $(X/2, X]$, each of which has a "very small" prime factor. By the Chinese Remainder Theorem, this is equivalent to sieving out an interval by progressions to small primes.

Definition 7 (Jacobsthal's function)**.** *Let x be a positive integer. Define $Y(x)$ to be the largest gap in the integers with no prime factor $\leqslant x$. Equivalently, $Y(x)$ is the largest integer y for which one may select residue classes $a_p \mod p$, one for each prime $p \leqslant x$, which together "sieve out" (cover) a whole interval of integers of length y.*

The equivalence of the two definitions of $Y(x)$ is easy by the Chinese remainder theorem. Let $P(x) = \prod_{p\leqslant x} p$. One consequence is that $(P(x), 2P(x)]$ contains $Y(x)$ consecutive numbers that have a prime factor $\leqslant x$, and thus these numbers are all composite. We conclude that

$$G(2P(x)) \geqslant Y(x). \tag{8.2}$$

Since $\log P(x) \sim x$ by the Prime Number Theorem, this essentially says that

$$G(X) \gtrsim Y(\log X).$$

All known lower bounds in $G(X)$ are based on lower bounds for $Y(x)$.

Fix a positive real number c, let x be a large integer and define

$$y := \left\lfloor cx \frac{\log x \log_3 x}{(\log_2 x)^2} \right\rfloor, \tag{8.3}$$

Also let

$$z := x^{\log_3 x/(5\log_2 x)}, \tag{8.4}$$

and introduce the three disjoint sets of primes

$$\begin{aligned} \mathcal{S} &:= \{s \text{ prime} : \log^{20} x < s \leqslant z\}, \\ \mathcal{P} &:= \{p \text{ prime} : x/2 < p \leqslant x\}, \\ \mathcal{Q} &:= \{q \text{ prime} : x < q \leqslant y\}. \end{aligned}$$

We will show that $Y(x) \geqslant y - x$ by covering $(x, y]$ with residue classes modulo primes $\leqslant x$. Now $P(x) = e^{(1+o(1)x}$ by the prime number theorem, so (8.2) and (8.3) imply that

$$G(e^{(1+o(1))x}) \geqslant (1+o(1))cx \log x \frac{\log_3 x}{(\log_2 x)^2}.$$

AS G is monotone, Theorem 8.2 follows upon taking c arbitrarily large.

We first reduce the problem to finding residue classes modulo the primes in $\mathcal{S} \cup \mathcal{P}$ which cover most of the primes in $\mathcal{Q}$. For residue classes $\vec{a} = (a_s \mod s)_{s\in\mathcal{S}}$ and $\vec{b} = (b_p \mod p)_{p\in\mathcal{P}}$, define the sifted sets

$$S(\vec{a}) := \{n \in \mathbb{Z} : n \not\equiv a_s \pmod s \text{ for all } s \in \mathcal{S}\}$$

and likewise

$$T(\vec{b}) := \{n \in \mathbb{Z} : n \not\equiv b_p \pmod p \text{ for all } p \in \mathcal{P}\}.$$

Theorem 8.3 (Sieving primes)**.** *Let $c > 0$ be arbitrary, let x be sufficiently large and suppose that y obeys* (8.3)*. Then there are vectors $\vec{a} = (a_s \mod s)_{s\in\mathcal{S}}$ and $\vec{b} = (b_p \mod p)_{p\in\mathcal{P}}$, such that*

$$\left|\mathcal{Q} \cap S(\vec{a}) \cap T(\vec{b})\right| \leqslant \frac{x}{5\log x}. \tag{8.5}$$

Proof the Theorem 8.3 implies Theorem 8.2. We will first take

$$a_p = 0 \qquad (p \leqslant \log^{20} x, z < p \leqslant x/4).$$

This is a "big gun" and eliminates all numbers from $(x, y]$ except primes and a negligible set of z-smooth numbers. It dates from work of Westzynthius [**146**], Erdős [**33**] and Rankin [**124**]. Let $\vec{a}$ and $\vec{b}$ be as in Theorem 8.3, and consider the sifted set

$$\mathcal{U} := (x, y] \cap S(\vec{a}) \cap T(\vec{b}) \cap V,$$

where

$$V = \{n \in \mathbb{Z} : n \not\equiv 0 \pmod p \text{ for all } p \leqslant \log^{20} x \text{ and } z < p \leqslant x/4\}.$$

The elements of $\mathcal{U}$, by construction, are not divisible by any prime in $(0, \log^{20} x] \cup (z, x/4]$. Thus, each element must either be a z-smooth number (i.e., a number with all prime factors at most z), or must consist of a prime greater than $x/4$, possibly multiplied by some additional primes that are all at least $\log^{20} x$. However, from (8.3) we know that $y = o(x \log x)$. Thus, we see that an element of $\mathcal{U}$ is either a z-smooth number or a prime in $(x/4, y]$. In the second case, the element lies in $\mathcal{Q} \cap S(\vec{a}) \cap T(\vec{b})$. Conversely, every element of $\mathcal{Q} \cap S(\vec{a}) \cap T(\vec{b})$ lies in $\mathcal{U}$. Thus, $\mathcal{U}$ only differs from $\mathcal{Q} \cap S(\vec{a}) \cap T(\vec{b})$ by a set $\mathcal{R}$ consisting of z-smooth numbers in $[1, y]$. Let $u := \frac{\log y}{\log z} \sim 5\frac{\log_2 x}{\log_3 x}$. By Theorem 5.1 and (8.3),

$$|\mathcal{R}| \leqslant \Psi(y, z) \ll \frac{y}{(\log x)^{5+o(1)}} \ll \frac{x}{\log^2 x}. \tag{8.6}$$

Thus, we find that

$$|\mathcal{U}| \leqslant (1 + o(1))\frac{x}{5\log x}.$$

By matching each of these surviving elements to a distinct prime in $(x/4, x/2]$ and choosing congruence classes appropriately, we thus find congruence classes $a_p \mod p$ for $p \leqslant x$ which cover *all* of the integers in $(x, y]$. This proves $Y(x) \geqslant y - x = (1 - o(1))y$, and hence Theorem 8.2. □

8.2.1. *The Westzynthius-Erdős-Rankin argument giving* (8.1) *for some $c > 0$.* It is very easy to show that (8.5) holds for some value of $c > 0$. Indeed, by the pigeonhole principle, for any finite set $\mathcal{R}$ and prime p there is a residue class a_p so that $|\mathcal{R} \cap (a_p \mod p)| \geqslant |\mathcal{R}|/p$. Consequently, there are chocies of a_s for $s \in \mathcal{S}$ and b_p for $p \in \mathcal{P}$ such that

$$\begin{aligned}\left|\mathcal{Q} \cap S(\vec{a}) \cap T(\vec{b})\right| &\leqslant |\mathcal{Q}| \prod_{s\in\mathcal{S}} \left(1 - \frac{1}{s}\right) \prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p}\right) \\ &\ll |\mathcal{Q}| \frac{\log_2 x}{\log z} \\ &\ll \frac{y(\log_2 x)^2}{(\log x)^2 \log_3 x} \leqslant \frac{cx}{\log x}.\end{aligned}$$

Hence, if c is small enough, then (8.5) holds.

8.2.2. *Improving the Westzynthius-Erdős-Rankin method.* We argue in two stages:

Stage 1. For each prime $s \in \mathcal{S}$, we select each $\mathbf{a}_s \mod s$ uniformly at random from $\mathbb{Z}/s\mathbb{Z}$, independently for each s. Let $\vec{\mathbf{a}} := (\mathbf{a}_s \mod s)_{s\in\mathcal{S}}$;

Stage 2. For each prime $p \in \mathcal{P}$, we select the residue class $b_p \mod p$ also at random, but in a strategic way which is dependent on the choice of $\vec{\mathbf{a}}$; this is done in such a way so that each $b_p \mod p$ covers many primes in $\mathcal{Q}$ left uncovered by Stage 1.

What we gain in Stage 1 is only what is typical for residue classes for these primes, but the random choice will be advantagoeus for Stage 2, which is much more complicated. Greatly improving upon arguments from Maier and Pomerance [**107**] and Pintz [**117**], we show in Stage 2 that for any fixed r, there are residue classes $b_p \mod p$ for $p \in \mathcal{P}$ which cover an average of at least r numbers left uncovered by Stage 1.

8.3. **Concentration of $S(\vec{\mathbf{a}})$.** The sifted set $S(\vec{\mathbf{a}})$ is a random periodic subset of $\mathbb{Z}$, each element surviving with probability

$$\sigma := \prod_{s\in\mathcal{S}} \left(1 - \frac{1}{s}\right) = \prod_{\log^{20} x < s \leqslant z} \left(1 - \frac{1}{s}\right).$$

From Mertens' bound and (8.4),

$$\sigma \sim \frac{\log(\log^{20} x)}{\log z} = \frac{100(\log_2 x)^2}{\log x \log_3 x}. \tag{8.7}$$

In particular, we have

$$\mathbb{E}\big|\mathcal{Q} \cap S(\vec{\mathbf{a}})\big| = \sum_{q\in\mathcal{Q}} \mathbb{P}(q \in S(\vec{\mathbf{a}})) = \sigma|\mathcal{Q}| \sim 100c\frac{x}{\log x}. \tag{8.8}$$

We now show that $|\mathcal{Q}\cap S(\vec{\mathbf{a}})|$ is concentrated about its mean. To accomplish this, we first show that for any reasonably sized integers $n_1, \dots, n_t$, the events $n_i \in S(\vec{\mathbf{a}})$ are close to being independent.

Lemma 8.4. *Let $t \leqslant \log x$, and let $n_1, \dots, n_t$ be distinct integers in $[-x^2, x^2]$. Then*

$$\mathbb{P}(n_1, \dots, n_t \in S(\vec{\mathbf{a}})) = \left(1 + O\left(\frac{1}{\log^{16} x}\right)\right)\sigma^t.$$

Proof. For each $s \in \mathcal{S}$, the probability that $n_1, \dots, n_t$ are all avoid $\mathbf{a}_s \mod s$ is equal to $1 - \frac{m(s)}{t}$, where $m(s)$ is the number of distinct residue classes modulo s occupied by $n_1, \dots, n_t$. We have $m(s) = t$ unless s divides one of the numbers $n_i - n_j$ for $1 \leqslant i < j \leqslant t$. Since $|n_i - n_j| \leqslant 2x^2$, each difference $n_i - n_j$ has $O(\log x)$ prime factors. Therefore, $m(s) < t$ occurs for at most $O(t^2 \log x) = O(\log^3 x)$ primes $s \in \mathcal{S}$. Byt the independence of the choices $\mathbf{a}_s$ for $s \in \mathcal{S}$, the events "$\{n_1, \dots, n_t\}$

doesn't intersect $\mathbf{a}_s \mod s$" are independent. Thus,

$$\begin{aligned}\mathbb{P}(n_1,\dots,n_t\in S(\vec{\mathbf{a}})) &= \prod_{s\in\mathcal{S}}\left(1-\frac{m(s)}{s}\right)\\ &= \prod_{s\in\mathcal{S}}\left(1-\frac{t}{s}\right)\prod_{\substack{s\in\mathcal{S}\\ m(s)<t}}\left(1-\frac{t}{s}\right)^{-1}\left(1-\frac{m(s)}{s}\right)\\ &= \left(1+O\left(\frac{1}{\log^{19}x}\right)\right)^{O(\log^3 x)}\prod_{s\in\mathcal{S}}\left(1-\frac{t}{s}\right)\\ &= \left(1+O\left(\frac{1}{\log^{16}x}\right)\right)\sigma^t\prod_{s\in\mathcal{S}}\left(1+O\left(\frac{t^2}{s^2}\right)\right)\\ &= \left(1+O\left(\frac{1}{\log^{16}x}\right)\right)\sigma^t.\end{aligned}$$

□

Corollary 8.5. *With probability $\geqslant 1-O(1/\log^8 x)$, we have*

$$\big|\mathcal{Q}\cap S(\vec{\mathbf{a}})\big| = \left(1+O\left(\frac{1}{\log^4 x}\right)\right)\sigma|\mathcal{Q}| \sim 100c\frac{x}{\log x}. \tag{8.9}$$

Proof. From Lemma 8.4, we have

$$\begin{aligned}\mathbb{E}\big|(\mathcal{Q}\cap S(\vec{\mathbf{a}}))\big|^2 &= \sum_{q_1,q_2\in\mathcal{Q}}\mathbb{P}(q_1,q_2\in S(\vec{\mathbf{a}}))\\ &= \underbrace{\sigma|\mathcal{Q}|}_{q_1=q_2} + \underbrace{\left(1+O\left(\frac{1}{\log^{16}x}\right)\right)\sigma^2|\mathcal{Q}|\cdot(|\mathcal{Q}|-1)}_{q_1\neq q_2}\\ &= \sigma^2|\mathcal{Q}|^2\left(1+O\left(\frac{1}{\log^{16}x}\right)\right),\end{aligned}$$

using that $\sigma|\mathcal{Q}| \gg (x/\log x)^2$. Thus, by (8.8), (8.3), (8.7),

$$\mathbb{E}\big(\big|\mathcal{Q}\cap S(\vec{\mathbf{a}})\big|-\sigma|\mathcal{Q}|\big)^2 = \mathbb{E}\big|\mathcal{Q}\cap S(\vec{\mathbf{a}})\big|^2-(\sigma|\mathcal{Q}|)^2 \ll \frac{(\sigma|\mathcal{Q}|)^2}{\log^{16}x}.$$

Combining this last estimate with Chebyshev's inequality, we conclude that

$$\mathbb{P}\left(\big|\,|\mathcal{Q}\cap S(\vec{\mathbf{a}})|-\sigma|\mathcal{Q}|\big| \geqslant \frac{\sigma|\mathcal{Q}|}{\log^4 x}\right) \ll \frac{(\sigma|\mathcal{Q}|)^2/\log^{16}x}{(\sigma|\mathcal{Q}|)^2(\log x)^{-8}} = \frac{1}{\log^8 x}.$$

□

8.4. **The weight function $w^*(p,n)$.** We will also choose the vector $\vec{b}$ at random, but in a strategic way which is *dependent* on $\vec{\mathbf{a}}$. Each residue class $b_p \mod p$ must cover many primes in $\mathcal{Q}$ (in fact, many primes in $\mathcal{Q}\cap S(\vec{\mathbf{a}})$), and accomplishing this is the key to success of the whole enterprise. This is done via showing the existence of a good "weight" function $w^*(p,n)$.

Let k be a fixed positive integer, large in terms of the constant c (in fact, $k=\lfloor\exp(c^2)\rfloor$ will do), and let $(h_1,\dots,h_k)$ be a fixed *admissible k-tuple* of positive integers (that is, for any prime, $(h_1,\dots,h_k)$ does not cover $(\mathbb{Z}/p\mathbb{Z})$). For instance, $(1,3^2,5^2,\dots,(2k-1)^2)$.

Theorem 8.6 (Existence of good sieve weight)**.** *Let k be a fixed, large positive integer and $(h_1, \dots, h_k)$ a fixed admissible $k-$tuple of distinct positive integers. Suppose x is large and y is defined by (8.3), with $c > 0$ fixed. Then there are quantities τ, u satisfying*

$$\tau = x^{o(1)}, \qquad u \gg \log k \qquad (x \to \infty) \tag{8.10}$$

and a non-negative weight function $w^*(p, n)$ defined on $\mathcal{P} \times ([-y, y] \cap \mathbb{Z})$ satisfying

- *Uniformly for every $p \in \mathcal{P}$, one has*

$$\sum_{n \in \mathbb{Z}} w^*(p, n) \sim \tau \frac{y}{\log^k x}. \tag{8.11}$$

- *Uniformly for every $q \in \mathcal{Q}$ and $i = 1, \dots, k$, one has*

$$\sum_{p \in \mathcal{P}} w^*(p, q - h_i p) \sim \tau \frac{u}{k} \frac{x/2}{\log^k x}. \tag{8.12}$$

- *Uniformly for all $p \in \mathcal{P}$ and $n \in \mathbb{Z}$,*

$$w^*(p, n) \ll x^{o(1)} \qquad (x \to \infty). \tag{8.13}$$

Remark 2. One should think of $w^*(p, n)$ as being a smoothed out indicator function for the event that $n + h_1 p, \dots, n + h_k p$ are all almost primes in $[1, y]$. It is thus natural to bring into play the technology from Section 7.

Proof. We first recall the construction of sieve weights from Section 7. Set $\theta = 1/3$, let k be large and for large x define

$$s = \log_2 x, \quad R = x^{\frac{\theta}{2} - \frac{3/2}{s}}, \quad D = R^{1/s}, \quad z = D^{1/s} = R^{1/s^2},$$

and define $\mathscr{D}$ by (7.8). Let μ^+ be an upper bound sieve guaranteed by the Fundamental Lemma (Theorem 3.6) with parameters D, z. Fix $F \in \mathscr{F}_k$, so that

$$M_k(F) = \frac{kJ(F)}{I(F)} \geqslant 0.9 \log k,$$

which exists by Theorem 7.17 if k is large enough. Define $\xi(\mathbf{r})$ by (7.37) and define $\lambda(\mathbf{d})$ from Lemma 7.11. Note that with k and F fixed, $\lambda(\mathbf{d})$ depends only on R and z.

We will be applying Proposition 7.14 with $N = x/2$ and with $N = 2y$, and also with many different k-tuples of admissible forms $(a_1 n + b_1, \dots, a_k n + b_k)$. With these two choices of N, we see that (7.7) holds.

For $p \in \mathcal{P}$ and $-y \leqslant n \leqslant y$ we define

$$w^*(p, n) = \Bigg(\sum_{t | (n+h_1 p) \cdots (n+h_k p)} \mu^+(t) \Bigg) \Bigg(\sum_{\forall j : d_j | n + h_j p} \lambda(\mathbf{d}) \Bigg)^2 \quad (-y < n \leqslant y), \tag{8.14}$$

We are ready to deploy Proposition 7.14 (i), with $N = 2y$ and with the forms $m + (h_i p - 3y), 1 \leqslant i \leqslant k$, where now the variable m runs over $(N, 2N]$; that is, we set $m = n + 3y$. For this collection of forms, we have (recalling the definition (7.10))

$$E = \Bigg| p^{k(k-1)/2} \prod_{i<j} (h_j - h_i) \Bigg|.$$

Since all prime factors of E are either $O(1)$ or $> x/2 > R$, we see that for all $\mathbf{d} \in \mathscr{D}$ we have $\mathbf{d} \in \mathscr{E}$. Hence, setting $a_i = 1$ and $b_i = h_i p - 3y$ for $1 \leqslant i \leqslant k$, as have $w^*(p,n) = w(n+3y)$, where $w(m)$ is define in (7.12). We also have the required bounds (7.15) for a_i and b_i. Thus, Proposition 7.14 (i) implies that

$$\sum_{-y<n\leqslant y} w^*(p,n) = \sum_{N<m\leqslant 2N} w(m) \sim 2yV\left(e^{-\gamma}\frac{\log z}{\log R}\right)^k I(F),$$

where $I(F)$ is defined in (7.38) and

$$V = \prod_{\text{prime} v\leqslant z}\left(1-\frac{\rho(v)}{v}\right).$$

Here, for prime $v \leqslant z$,

$$\begin{aligned}\rho(v) &= \#\left\{n \mod v : \prod_{j=1}^k (n+h_j p) \equiv 0 \pmod v\right\}\\ &= \#\left\{n \mod v : \prod_{j=1}^k (n+h_j) \equiv 0 \pmod v\right\}\end{aligned}$$

is independent of p (since $v < p$). This proves (8.11) with

$$\tau = 2V\left(e^{-\gamma}(\log x)\frac{\log z}{\log R}\right)^k I(F) = x^{o(1)}. \tag{8.15}$$

Next, fix a prime $q \in \mathcal{Q}$ and an index $i \in \{1,\dots,k\}$. By (8.14) and the fact that q is prime $> z$,

$$\sum_{p\in\mathcal{P}} w^*(p, q-h_i p) = \sum_{x/2<n\leqslant x} \mathbb{1}_{n \text{ prime}}\left(\sum_{t\mid \prod_j (q+(h_j-h_i)n)} \mu^+(t)\right)\left(\sum_{d_j \mid q+(h_j-h_i)n\ \forall j} \lambda(\mathbf{d})\right)^2.$$

This corresponds exactly to the sum in Proposition 7.14 (ii) applied to the linear forms $a_j n + b_j$, $1 \leqslant j \leqslant k$, where $a_i = 1$, $b_i = 0$ and for $j \neq i$, $a_j = h_j - h_i$ and $b_i = q$. Here we have $N = x/2$. By (7.10), for these forms we have

$$E = \left|\prod_{j\neq i}(h_j - h_i)\cdot q^{k-1}\cdot \prod_{\substack{j_1<j_2\\ j_1\neq i, j_2\neq i}} (h_{j_1} - h_{j_2})q\right|,$$

which again has all of its prime factors $> x > R$ or $O(1)$. Thus, for all $\mathbf{d} \in \mathscr{D}$, $\mathbf{d} \in \mathscr{E}$ as well. We also have the required bounds (7.26) on a_i and b_i. Therefore, Proposition 7.14 (i) implies (8.10), where

$$u = \frac{\log R}{2\log x} M_k(F).$$

Since $M_k(F) \gg \log k$, $u \gg \log k$.

Finally, (8.13) holds since the definition of $w^*(p,n)$ implies that $w^*(p,n)$ is bounded by a divisor function, and $|\mu^+(t)| \leqslant 1$ and $|\lambda(\mathbf{d})| \leqslant 1$. □

8.5. **Strategic choice of $\vec{b}$.** Our goal is to first choose a "good" vector $\vec{a}$ and then choose residue classes b_p modulo $p \in \mathcal{P}$ such that $b_p \mod p$ contains many elements of $\mathcal{Q} \cap S(\vec{a})$.

For each $p \in \mathcal{P}$, let $\tilde{\mathbf{n}}_p$ denote the random integer with probability density

$$\mathbb{P}(\tilde{\mathbf{n}}_p = n) := \frac{w^*(p,n)}{\sum_{n' \in \mathbb{Z}} w^*(p,n')} \qquad (-y \leqslant n \leqslant y), \tag{8.16}$$

and chosen independently for each $p \in \mathcal{P}$. Notice that the random numbers $\tilde{\mathbf{n}}_p$ are chosen independently of the vector $\vec{\mathbf{a}}$. Roughly speaking, this gives more weight to n for which many of the numbers $n + h_i p$ are prime.

In order to capture the event that $b_p \mod p$ contains many elements of $S(\vec{\mathbf{a}})$, for each $p \in \mathcal{P}$ and fixed (non-random) $\vec{a}$, we consider the quantity

$$X_p(\vec{a}) := \mathbb{P}(\tilde{\mathbf{n}}_p + h_i p \in S(\vec{a}) \text{ for all } i = 1, \dots, k), \tag{8.17}$$

over the random integer $\tilde{\mathbf{n}}_p$. In light of Lemma 8.4, we expect that $X_p(\vec{a}) \sim \sigma^k$ for most choices of p and $\vec{a}$, and this will be confirmed below (Lemma 8.8). With this in mind, let $\mathcal{P}(\vec{a})$ denote the set of all the primes $p \in \mathcal{P}$ such that

$$\left| X_p(\vec{a}) - \sigma^k \right| \leqslant \frac{\sigma^k}{\log^3 x}. \tag{8.18}$$

We now define the random variables $\mathbf{n}_p$. Suppose we are in the event $\vec{\mathbf{a}} = \vec{a}$. If $p \in \mathcal{P} \backslash \mathcal{P}(\vec{a})$, we set $\mathbf{n}_p = 0$. Otherwise, if $p \in \mathcal{P}(\vec{a})$, we define $\mathbf{n}_p$ to be the random integer with conditional probability distribution

$$\mathbb{P}(\mathbf{n}_p = n | \vec{\mathbf{a}} = \vec{a}) := \frac{Z_p(\vec{a}; n)}{X_p(\vec{a})}, \tag{8.19}$$

where, for any $\vec{a}$, any $p \in \mathcal{P}$ and any n,

$$Z_p(\vec{a}; n) := \begin{cases} \mathbb{P}(\tilde{\mathbf{n}}_p = n) & \text{if } n + h_j p \in S(\vec{a}) \text{ for } j = 1, \dots, k \\ 0 & \text{otherwise,} \end{cases} \tag{8.20}$$

with the $\mathbf{n}_p$ ($p \in \mathcal{P}(\vec{a})$) jointly independent, conditionally on the event $\vec{\mathbf{a}} = \vec{a}$. From (8.17) we see that these random variables are well defined; indeed,

$$\sum_n Z_p(\vec{a}; n) = \mathbb{P}(\tilde{\mathbf{n}}_p + h_i p \in S(\vec{a}) \text{ for all } i = 1, \dots, k) = X_p(\vec{a}).$$

Observe that if $\vec{\mathbf{a}} = \vec{a}$ and $p \in \mathcal{P}(\vec{a})$, the support of $\mathbf{n}_p$ is contained in those n for which $n + h_i p \in S(\vec{a})$ for all i, that is, the residue class $n \mod p$ contains at least k elements of $S(\vec{a})$. Also, recalling Remark 2, we see that, conditionally on $\vec{\mathbf{a}} = \vec{a}$, $\mathbf{n}_p$ is very likely to be a number for which there are $\gg \log k$ of the numbers $n + h_1 p, \dots, n + h_k p$ which are prime (and hence in $\mathcal{Q}$). For a fixed vector $\vec{a}$ and $p \in \mathcal{P}$, define the random set

$$\mathbf{e}_p(\vec{a}) := \{\mathbf{n}_p + h_i p : 1 \leqslant i \leqslant k\} \cap \mathcal{Q} \cap S(\vec{a}). \tag{8.21}$$

Here the course of randomness is the random number $\mathbf{n}_p$. Thus, we expect that $\mathbf{e}_p(\vec{\mathbf{a}})$ will be a large set, which implies that the residue class $\mathbf{n}_p \mod p$ has many elements of $\mathcal{Q} \cap S(\vec{a})$, as desired. Unlike $\tilde{\mathbf{n}}_p$, the random numbers $\mathbf{n}_p$ are very much dependent on the vector $\vec{\mathbf{a}}$. However, conditional on $\vec{\mathbf{a}} = \vec{a}$, all of the $\mathbf{n}_p$ ($p \in \mathcal{P}$) are independent.

Our main objective is to show that $Z_p(\vec{\mathbf{a}}; q - h_i p)$ is often large.

Lemma 8.7. *Fix c, k, and let u be the value guaranteed by Theorem 8.6. Define $Z_p(\vec{a}; n)$ by (8.20). Then, as $x \to \infty$, with probability $1 - o(1)$ in the random vector $\vec{\mathbf{a}}$ we have*

$$\sigma^{-k} \sum_{i=1}^{k} \sum_{p \in \mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; q - h_i p) \geqslant \frac{u}{250c} \tag{8.22}$$

for all but at most $\frac{x}{100 \log x}$ of the primes $q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})$.

Proof of Theorem 8.3 from Lemma 8.7. By Corollary 8.5 and Lemma 8.7, there is some vector $\vec{a}$ such that

$$|\mathcal{Q} \cap S(\vec{a})| \leqslant 110c \frac{x}{\log x} \tag{8.23}$$

and also (8.22) holds for all but $o(x/\log x)$ primes $q \in \mathcal{Q} \cap S(\vec{a})$. Fix this vector $\vec{a}$ (it is no longer random), and choose random integers $\mathbf{n}_p$ according to the laws (8.19). Let $\mathcal{Q}'$ denote the set of primes $q \in \mathcal{Q} \cap S(\vec{a})$ for which (8.22) holds. By Lemma 8.7,

$$\big|(\mathcal{Q} \cap S(\vec{a})) \setminus \mathcal{Q}'\big| = o\left(\frac{x}{\log x}\right) \quad (x \to \infty). \tag{8.24}$$

Substituting definition (8.19) into the left hand side of of (8.22), using (8.18), and observing that $q = \mathbf{n}_p + h_i p$ is only possible if $p \in \mathcal{P}(\vec{a})$ (since otherwise $\mathbf{n}_p = 0$ and $q = h_i p$ is not possible; here we use the fact that $h_i > 0$), we find that

$$\begin{aligned}
\sigma^{-k} \sum_{i=1}^{k} \sum_{p \in \mathcal{P}(\vec{a})} Z_p(\vec{a}; q - h_i p) &= \sigma^{-k} \sum_{i=1}^{k} \sum_{p \in \mathcal{P}(\vec{a})} X_p(\vec{a}) \mathbb{P}(\mathbf{n}_p = q - h_i p | \vec{\mathbf{a}} = \vec{a}) \\
&= \left(1 + O\left(\frac{1}{\log^3 x}\right)\right) \sum_{i=1}^{k} \sum_{p \in \mathcal{P}(\vec{a})} \mathbb{P}(\mathbf{n}_p = q - h_i p | \vec{\mathbf{a}} = \vec{a}) \\
&= \left(1 + O\left(\frac{1}{\log^3 x}\right)\right) \sum_{p \in \mathcal{P}} \mathbb{P}(q \in \mathbf{e}_p(\vec{a})).
\end{aligned}$$

By (8.22), we conclude that if x is large enough then

$$\sum_{p \in \mathcal{P}} \mathbb{P}(q \in \mathbf{e}_p(\vec{a})) \geqslant \frac{u}{300c}. \tag{8.25}$$

Observe that each random set $\mathbf{e}_p$ is contained in a single arithmetic progression modulo p. Now we set $\mathbf{b}_p = \mathbf{n}_p \mod p$ for each $p \in \mathcal{P}$. Then, for any $q \in \mathcal{Q}'$, by the independence of the $\mathbf{n}_p$,

$$\begin{aligned}
\mathbb{P}(q \in T(\vec{\mathbf{b}})) &= \mathbb{P}\left(q \not\equiv \mathbf{b}_p \pmod p \ \forall p \in \mathcal{P}\right) \\
&= \prod_{p \in \mathcal{P}} \left(1 - \mathbb{P}(q \equiv \mathbf{b}_p \pmod p)\right) \\
&\leqslant \prod_{p \in \mathcal{P}} \left(1 - \mathbb{P}(q \in \mathbf{e}_p(\vec{a}))\right) \\
&\leqslant \exp\Big\{ - \sum_{p \in \mathcal{P}} \mathbb{P}(q \in \mathbf{e}_p(\vec{a}))\Big\} \\
&= \exp\Big\{ - \frac{u}{300c}\Big\}.
\end{aligned} \tag{8.26}$$

Hence, by (8.23) and (8.24), if x is large enough then

$$\begin{aligned}\mathbb{E}\big|\mathcal{Q}\cap S(\vec{a})\cap T(\vec{\mathbf{b}})\big| &\leqslant \big|(\mathcal{Q}\cap S(\vec{a}))\setminus \mathcal{Q}'\big| + \sum_{q\in\mathcal{Q}'} \exp\Big\{-\frac{u}{300c}\Big\}\\ &\leqslant \frac{x}{\log x}\bigg(\frac{1}{100} + 110c\exp\Big\{-\frac{u}{300c}\Big\}\bigg).\end{aligned}$$

The right hand side is $\leqslant \frac{x}{5\log x}$ if k is large enough, thanks to (8.10) which states $u \gg \log k$. Therefore, there is some choice of the vector $\vec{b}$ so that (8.5) holds, and this completes the proof of Theorem 8.3. □

It remains to prove Lemma 8.7. We first confirm that $\mathcal{P}\backslash\mathcal{P}(\vec{\mathbf{a}})$ is small with high probability.

Lemma 8.8. *We have*

$$\mathbb{E}|\mathcal{P}(\vec{\mathbf{a}})| = |\mathcal{P}| + O\bigg(\frac{x}{(\log x)^{11}}\bigg) = |\mathcal{P}|\Big(1+O\Big(\frac{1}{\log^{10} x}\Big)\Big).$$

Proof. It suffices to show that for each $p\in\mathcal{P}$, we have

$$\mathbb{P}\left(p\in\mathcal{P}(\vec{\mathbf{a}})\right) = 1 - O\bigg(\frac{1}{\log^{10} x}\bigg). \tag{8.27}$$

From (8.27), we get

$$\mathbb{E}\big|\mathcal{P}\backslash\mathcal{P}(\vec{\mathbf{a}})\big| = \sum_{p\in\mathcal{P}} \mathbb{P}\left(p\notin\mathcal{P}(\vec{\mathbf{a}})\right) \ll \frac{|\mathcal{P}|}{\log^{10} x} \ll \frac{x}{\log^{11} x},$$

and the claim follows. We will prove (8.27) by computing first and second moments of $X_p(\vec{\mathbf{a}})$. Recall that, by (8.17),

$$X_p(\vec{\mathbf{a}}) := \sum_n 1_{n+h_ip\in S(\vec{\mathbf{a}})\,\forall i}\,\mathbb{P}(\tilde{\mathbf{n}}_p = n).$$

Since the quantities $\mathbb{P}(\tilde{\mathbf{n}}_p = n)$, for $-y\leqslant n\leqslant y$, are independent of $\vec{\mathbf{a}}$, Lemma 8.4 implies that

$$\begin{aligned}\mathbb{E}X_p(\vec{\mathbf{a}}) &= \sum_n \mathbb{P}(n+h_ip\in S(\vec{\mathbf{a}}) \text{ for all } i=1,\dots,k)\,\mathbb{P}(\tilde{\mathbf{n}}_p=n)\\ &= \bigg(1+O\bigg(\frac{1}{\log^{16} x}\bigg)\bigg)\sigma^k\sum_n \mathbb{P}(\tilde{\mathbf{n}}_p=n) = \bigg(1+O\bigg(\frac{1}{\log^{16} x}\bigg)\bigg)\sigma^k.\end{aligned} \tag{8.28}$$

Similarly,

$$\mathbb{E}X_p(\vec{\mathbf{a}})^2 = \sum_{n_1,n_2} \mathbb{P}(n_j+h_ip\in S(\vec{\mathbf{a}}) \text{ for } 1\leqslant i\leqslant k; j=1,2)\mathbb{P}(\tilde{\mathbf{n}}_p=n_1)\mathbb{P}(\tilde{\mathbf{n}}_p=n_2).$$

When $n_1\not\equiv n_2 \pmod p$, we have

$$\#\{n_j+h_ip : i=1,\dots,k; j=1,2\} = 2k$$

and Lemma 8.4 implies that

$$\mathbb{P}(n_j+h_ip\in S(\vec{\mathbf{a}}) \text{ for } 1\leqslant i\leqslant k; j=1,2) = \bigg(1+O\Big(\frac{1}{\log^{16} x}\Big)\bigg)\sigma^{2k}.$$

There are $O(y(y/p)) = O(y^2/x) = O(x\log^2 x)$ pairs n_1, n_2 with $n_1 \equiv n_2 \pmod p$. Therefore,

$$\mathbb{E}X_p(\vec{\mathbf{a}})^2 = \left(1 + O\left(\frac{1}{\log^{16} x}\right)\right)\sigma^{2k} + O\left((x\log^2 x)\max_n \left(\mathbb{P}(\tilde{\mathbf{n}}_p = n)\right)^2\right).$$

From (8.16), (8.11), (8.13), and (8.10) one has $\mathbb{P}(\tilde{\mathbf{n}}_p = n) \ll x^{-0.99}$ for all $p \in \mathcal{P}$ and $n \in \mathbb{Z}$. Also, $\sigma \gg 1/\log x$ by (8.7), hence

$$\mathbb{E}X_p(\vec{\mathbf{a}})^2 = \left(1 + O\left(\frac{1}{\log^{16} x}\right)\right)\sigma^{2k}. \tag{8.29}$$

Combining (8.28) and (8.29), we compute

$$\mathbb{E}\left|X_p(\vec{\mathbf{a}}) - \sigma^k\right|^2 \ll \frac{\sigma^{2k}}{\log^{16} x}.$$

By Chebyshev's inequality, for any $p \in \mathcal{P}$,

$$\mathbb{P}\left(p \notin \mathcal{P}(\vec{\mathbf{a}})\right) = \mathbb{P}\left(\left|X_p(\vec{\mathbf{a}}) - \sigma^k\right| \geqslant \frac{1}{\log^3 x}\right) \ll \frac{1}{\log^{10} x},$$

which proves (8.27). □

Proof of Lemma 8.7. We first show that replacing $\mathcal{P}(\vec{\mathbf{a}})$ with $\mathcal{P}$ has negligible effect on the sum, with high probability. The purpose is to set up an application of (8.12) and to decouple some expressions involving $\vec{\mathbf{a}}$.

By Lemma 8.4 and the definition (8.20) of $Z_p(\vec{a}; n)$, we have

$$\begin{aligned}\mathbb{E}\sum_n \sigma^{-k}\sum_{p\in\mathcal{P}} Z_p(\vec{\mathbf{a}}; n) &= \sigma^{-k}\sum_{p\in\mathcal{P}}\sum_n \mathbb{P}(\tilde{\mathbf{n}}_p = n)\mathbb{P}(n + h_j p \in S(\vec{\mathbf{a}}) \text{ for } j = 1, \dots, k)\\ &= \left(1 + O\left(\frac{1}{\log^{16} x}\right)\right)\sum_{p\in\mathcal{P}}\sum_n \mathbb{P}(\tilde{\mathbf{n}}_p = n)\\ &= \left(1 + O\left(\frac{1}{\log^{16} x}\right)\right)|\mathcal{P}|.\end{aligned}$$

Next, by (8.18) and Lemma 8.8 we have

$$\mathbb{E}\sum_n \sigma^{-k}\sum_{p\in\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; n) = \sigma^{-k}\sum_{\vec{a}} \mathbb{P}(\vec{\mathbf{a}} = \vec{a})\sum_{p\in\mathcal{P}(\vec{a})} X_p(\vec{a})\underbrace{\sum_n \mathbb{P}(\mathbf{n}_p = n|\vec{\mathbf{a}} = \vec{a})}_{\text{equals } 1}$$

$$\begin{aligned}&= \left(1 + O\left(\frac{1}{\log^3 x}\right)\right)\sum_{\vec{a}} \mathbb{P}(\vec{\mathbf{a}} = \vec{a})\sum_{p\in\mathcal{P}(\vec{a})} 1\\ &= \left(1 + O\left(\frac{1}{\log^3 x}\right)\right)\mathbb{E}\,|\mathcal{P}(\vec{\mathbf{a}})| = \left(1 + O\left(\frac{1}{\log^3 x}\right)\right)|\mathcal{P}|.\end{aligned}$$

Subtracting these two expectations, we conclude that

$$\mathbb{E}\sum_n \sigma^{-k}\sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; n) \ll \frac{|\mathcal{P}|}{\log^3 x} \ll \frac{x}{\log^4 x}. \tag{8.30}$$

We substitute $n = q - h_i p$ and find that

$$\begin{aligned}\mathbb{E} \sum_{q\in\mathcal{Q}\cap S(\vec{\mathbf{a}})} \sigma^{-k} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; q-h_ip) &\leqslant \mathbb{E} \sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} \sum_{i=1}^{k} \sigma^{-k} \sum_{n} Z_p(\vec{\mathbf{a}}; n)\\ &= k\cdot \mathbb{E} \sum_{n} \sigma^{-k} \sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; n)\\ &\ll \frac{x}{\log^4 x}.\end{aligned}$$

By Markov's inequality, with probability $1 - O(1/\log x)$, we have

$$\sum_{q\in\mathcal{Q}\cap S(\vec{\mathbf{a}})} \sigma^{-k} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; q-h_ip) \leqslant \frac{x}{\log^3 x}.$$

Hence, with probability $1 - O(1/\log x)$, we have

$$\sigma^{-k} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}\setminus\mathcal{P}(\vec{\mathbf{a}})} Z_p(\vec{\mathbf{a}}; q-h_ip) \leqslant \frac{1}{\log x}$$

for all but at most $x/\log^2 x$ primes $q \in \mathcal{Q}\cap S(\vec{\mathbf{a}})$. Recalling our goal (8.22), it therefore suffices to show that with probability $1 - o(1)$, for all but at most $x/(200\log x)$ primes $q \in \mathcal{Q}\cap S(\vec{\mathbf{a}})$, one has

$$F(q;\vec{\mathbf{a}}) := \sigma^{-k} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}} Z_p(\vec{\mathbf{a}}; q-h_ip) \geqslant \frac{u}{240c}. \tag{8.31}$$

We accomplish this with another first-second moment calculation.

First, we note that from Theorem 8.6, (8.11) and (8.12), we have

$$\sum_{p\in\mathcal{P}} \mathbb{P}(q = \tilde{\mathbf{n}}_p + h_ip) = \frac{\sum_{p\in\mathcal{P}} w^*(p, q-h_ip)}{\sum_m w^*(p,m)} \sim \frac{u}{k}\frac{x}{2y} \qquad (q\in\mathcal{Q}, 1\leqslant i\leqslant k). \tag{8.32}$$

Hence, combining (8.32) with Lemma 8.4 and (8.16), we get

$$\begin{aligned}\mathbb{E} \sum_{q\in\mathcal{Q}\cap S(\vec{\mathbf{a}})} F(q;\vec{\mathbf{a}}) &= \sigma^{-k} \sum_{q\in\mathcal{Q}} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}} \mathbb{P}(q+(h_j-h_i)p\in S(\vec{\mathbf{a}})\ \forall j)\mathbb{P}(\tilde{\mathbf{n}}_p = q-h_ip)\\ &= \left(1+O\left(\frac{1}{\log^{16} x}\right)\right) \sum_{q\in\mathcal{Q}} \sum_{i=1}^{k} \sum_{p\in\mathcal{P}} \mathbb{P}(\tilde{\mathbf{n}}_p = q-h_ip)\\ &\sim \sum_{q\in\mathcal{Q}} \sum_{i=1}^{k} \frac{ux}{2ky} \sim \frac{\sigma y}{\log x}\left(\frac{ux}{2\sigma y}\right).\end{aligned}$$

This shows that $F(q;\vec{\mathbf{a}})$ is about $ux/(2\sigma y)$ on average, since by Corollary 8.5, $|\mathcal{Q}\cap S(\vec{\mathbf{a}})| \sim \sigma y/\log x$ with high probability. Similarly,

$$\begin{aligned}\mathbb{E} \sum_{q\in\mathcal{Q}\cap S(\vec{\mathbf{a}})} F(q;\vec{\mathbf{a}})^2 = \sigma^{-2k} \sum_{q\in\mathcal{Q}} \sum_{p_1,p_2\in\mathcal{P}} \sum_{i_1,i_2} &\mathbb{P}(q+(h_j-h_{i_\ell})p_\ell \in S(\vec{\mathbf{a}}) \text{ for } j=1,\dots,k; \ell=1,2)\\ &\times \mathbb{P}(\tilde{\mathbf{n}}_{p_1} = q-h_{i_1}p_1)\mathbb{P}(\tilde{\mathbf{n}}_{p_2} = q-h_{i_2}p_2).\end{aligned}$$

The terms with $p_1 = p_2$ contribute negligibly; indeed, since $\mathbb{P}(\tilde{\mathbf{n}} = n) \ll x^{-0.99}$, these terms contribute an amount which is

$$\ll \sigma^{-2k}|\mathcal{Q}| \cdot |\mathcal{P}|k^2(x^{-0.99})^2 \ll x^{0.03}.$$

When $p_1 \neq p_2$ and $q \in \mathcal{Q}$, there are $2k-1$ distinct numbers $q + (h_j - h_{i_\ell})p_\ell$, $1 \leqslant j \leqslant k$, $1 \leqslant \ell \leqslant 2$, since the terms $j = i_1, \ell = 1$ and $j = i_2, \ell = 2$ are both equal to q and no other terms are equal; this is the case whether or not $i_1 = i_2$. Therefore, by Lemma 8.4

$$\mathbb{E} \sum_{q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})} F(q; \vec{\mathbf{a}})^2 \sim \frac{\sigma y}{\log x} \left(\frac{ux}{2\sigma y}\right)^2.$$

Putting the first and second moment bounds together, and using (8.8), we get

$$\begin{aligned}\mathbb{E} \sum_{q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})} \left(F(q; \vec{\mathbf{a}}) - \frac{xu}{2\sigma y}\right)^2 &= \mathbb{E} \sum_{q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})} F(q; \vec{\mathbf{a}})^2 - 2\frac{xu}{2\sigma y} \mathbb{E} \sum_{q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})} F(q; \vec{\mathbf{a}}) + \left(\frac{xu}{2\sigma y}\right)^2 \mathbb{E}\big|\mathcal{Q} \cap S(\vec{\mathbf{a}})\big| \\ &= o\left(\frac{\sigma y}{\log x}\left(\frac{xu}{2\sigma y}\right)^2\right).\end{aligned}$$

Using Chebyshev's inequality, we find that the left side is $o(\frac{\sigma y}{\log x}(\frac{xu}{2\sigma y})^2)$ with probability $1 - o(1)$. In this event, $F(q; \vec{\mathbf{a}}) \sim \frac{xu}{2\sigma y}$ for all but $o(\sigma y/\log x)$ primes $q \in \mathcal{Q} \cap S(\vec{\mathbf{a}})$. Now $\sigma y/\log x \asymp x/\log x$ and

$$\frac{xu}{2\sigma y} \sim \frac{u}{200c}$$

and Lemma 8.7 follows. □

8.6. **Remarks: quantitative improvements.** The proof described above may be adapted to the case where $k \to \infty$ ($x \to \infty$). The relevant estimates needed for the analog of Theorem 8.6 may be found in [**109**]. The error terms corresponding to bounds for primes in progressions from the sum (8.12) need to be of the form $x/(\log x)^{100k^2}$, and this is achieved (as in [**109**]) by a careful analysis of Siegel exceptional zeros. With $k \approx (\log x)^{1/5}$, the analog of Theorem 8.6 will have a factor $u \gg \log\log x$ in (8.12). Inserted into the final estimate (8.26) in the proof of Theorem 2, we find that to succeed we need

$$u \gg c \log c,$$

which means that we can improve $Y(x)$ be a factor $\gg u/\log u$ (and hence $G(x)$ improved by a factor $\gg \frac{\log u}{\log\log u}$. In light of (8.25), however, we may expect an improvement of order $\gg u$ for $Y(x)$, and of order $\gg \log u$ for $G(x)$. This is achieved by dealing effectively with overlaps among the random sets $\mathbf{e}_p(\vec{a})$, using techniques from hypergraph covering (§4-5 of [**54**]).

8.7. **The influence of Siegel exceptional zeros on large gaps.** In this subsection, we show that the existence of exceptional zeros of a certain type implies a lower bound for $G(x)$ which is larger than the bound in Theorem 8.1. More generally, we derive a similar conclusion whenever there are arithmetic progressions containing few primes. We recall that $\pi(x;q,b)$ the number of primes $p \leqslant x$ satisfying $p \equiv b \pmod q$.

All of the methods used to prove lower bounds on $G(x)$ utilize a simple connection between $G(x)$ and Jacobsthal's function $Y(u)$, defined in Definition 7. Recall the simple relation (8.2) giving a lower bound for $G(x)$ in terms of $Y(x)$, and that $P(x) = e^{(1+o(1))x}$ by the prime number theorem.

Theorem 8.9 (Ford, 2019 [**50**])**.** *Suppose that x is large, $q > b \geqslant 1$, $(b,q) = 1$, $x \geqslant 2q$, and $\pi(x;q,b) \leqslant \frac{\delta x}{\phi(q)}$ with $0 \leqslant \delta \leqslant 1$. Then*

$$G(e^{2u}) \geqslant Y(u) \geqslant \frac{x-b}{q} \geqslant \frac{x}{2q},$$

where u is the smallest integer satisfying $u > 2\sqrt{x}$ and $\frac{u}{\log u} \geqslant \frac{6\delta x}{q}$.

Proof. Let u be as in the theorem, and let

$$y = \frac{x-b}{q}. \tag{8.33}$$

Too show that $Y(u) \geqslant y$, it suffices to find residue classes $a_p \mod p$, one for each prime $p \leqslant u$, which together cover $[0,y]$. For each prime $p \leqslant u/2$ with $p \nmid q$, define a_p by

$$qa_p + b \equiv 0 \pmod p.$$

Recall that $(b,q) = 1$. In this way, if $0 \leqslant n \leqslant y$ and $n \not\equiv a_p \pmod p$ for all such p, then $m = qn+b$ has no prime factor $\leqslant u/2$. Also, $x = qy + b < (u/2)^2$ by hypothesis, and thus m is prime. Let

$$\mathcal{N} = \{0 \leqslant n \leqslant y : n \not\equiv a_p \pmod p, \forall p \leqslant u/2 \text{ with } p \nmid q\}.$$

It follows from the hypothesis of the theorem that

$$|\mathcal{N}| \leqslant \pi(qy+b;q,b) = \pi(x;q,b) \leqslant \frac{\delta x}{\phi(q)}.$$

Next, we choose residue classes a_p for primes $p|q$ with $p \leqslant u/2$ using a greedy algorithm, successively selecting for each p a residue class $a_p \mod p$ which covers at least a proportion $1/p$ of the elements remaining uncovered. As $u > 2\sqrt{x} > 2\sqrt{q}$, there is at most one prime $p|q$ satisfying $p > u/2$. Letting $\mathcal{N}'$ denote the set of $n \in [0,y]$ not covered by $\{a_p \mod p : p \leqslant u/2\}$, we have

$$|\mathcal{N}'| \leqslant |\mathcal{N}| \prod_{p|q, p\leqslant u/2} \left(1 - \frac{1}{p}\right) \leqslant 2|\mathcal{N}|\frac{\phi(q)}{q} \leqslant \frac{2\delta x}{q}.$$

By hypothesis,

$$|\mathcal{N}'| \leqslant \frac{u}{3\log u},$$

which, by the prime number theorem, is less than the number of primes in $(u/2, u]$ for u large enough (as $u > 2\sqrt{x}$, this happens if x is large enough). Thus, we may associate each number $n \in \mathcal{N}'$ with a distinct prime in $p_n \in (u/2, u)$. Choosing $a_{p_n} \equiv n \pmod{p_n}$ for each $n \in \mathcal{N}'$ then ensures that $\{a_p \mod p : p \leqslant u\}$ covers all of $[0,y]$, as desired. □

An immediate corollary gives a lower bound on $G(x)$ assuming a lower bound on $L(q,b)$, the least prime in the progression $b \mod q$. We take $\delta = 0$ and $u = \lceil 2\sqrt{x} \rceil$.

Corollary 8.10 (Ford [**50**]). *Suppose that $(q,b) = 1$, $x \geqslant 2q$ and $L(q,b) > x$. Then*

$$G(e^{4\sqrt{x}+2}) \geqslant \frac{x}{2q}.$$

Theorem 8.9 is a partial converse to a theorem of Pomerance [**122**, Theorem 1], which provides a lower bound on $\max_{(b,q)=1} L(q,b)$ given a lower bound on the maximal gap between numbers coprime to m, where $(m,q) = 1$ and $m \leqslant q^{1-o(1)}$.

Linnik's theorem [**104**] states that for $(q,b) = 1$, $L(q,b) \ll q^L$ for some constant L; the best published result of this kind is due to Xylouris [**149**], who showed that the bound holds with $L = 5.18$ [7]. Assuming the Extended Riemann Hypothesis (ERH) for Dirichlet L-functions, we obtain a stronger bound $L(q,b) \ll_\varepsilon q^{2+\varepsilon}$ for every $\varepsilon > 0$.

Corollary 8.11 (Ford [**50**]). *Suppose that $c > 2$ and there are infinitely many pairs (q,b) with $L(q,b) \geqslant q^c$ (a violation of ERH). Then*

$$\limsup_{X\to\infty} \frac{G(X)}{(\log X)^{2-\frac{2}{c}}} > 0.$$

Proof. Apply Theorem 8.9 with $x = q^c$ and $X = e^{4\sqrt{x}+2}$. Then

$$G(X) \geqslant \frac{x}{2q} \gg q^{c-1} \gg (\log X)^{2-2/c}. \qquad \square$$

It is, however, conjectured that $L(q,b) \ll \phi(q)\log^2 q$; see [**102**] for a precise version of this conjecture and for the best known lower bounds on $\max_{(b,q)=1} L(q,b)$.

We may also exceed the bound in Theorem 8.1 under the assumption that exceptional zeros of Dirichlet L-functions exist. Roughly speaking, an exceptional zero of $L(s,\chi)$ is a zero which is real and very close to 1. As such, their existence violates ERH for $L(s,\chi)$. Classical results (see [**26**, §14]) imply that if $c_0 > 0$ is small enough, and $q \geqslant 3$, then there is at most one character χ modulo q for which $L(s,\chi)$ has a zero in the region

$$\{\sigma + it \in \mathbb{C} : \sigma \geqslant 1 - c_0/\log(qt)\},$$

and moreover the character is real and the zero is real. We shall refer to such zeros as "exceptional zeros" with respect to c_0. Moreover, by reducing c_0 if necessary, it is known that moduli q for which an exceptional zero exists are very rare.

Siegel's theorem [**26**, Sec. 21] implies that

$$1 - \beta_q \gg_\varepsilon q^{-\varepsilon} \qquad (\forall\, \varepsilon > 0), \tag{8.34}$$

for (hypothetical) exceptional zeros β_q, although we cannot say any rate at which this occurs (the bound is *ineffective*). The exceptional zeros are also known as Siegel zeros or Landau-Siegel zeros in the literature. Their existence implies a great irregularity in the distribution of primes modulo q, given by Gallagher's Prime Number Theorem [**62**].

[7] In his Ph.D. thesis [**150**], Xyloris claims a better bound $L = 5$.

Proposition 8.12. *For some absolute constant $B > 1$, we have the following. Suppose that χ is a real character with conductor q and $L(1-\delta,\chi) = 0$ for some $\frac{1}{2} \leqslant \delta < 1$. Then, for all b with $\chi_q(b) = 1$ and all $x \geqslant q^B$, we have*

$$\pi(x;q,b) \ll \frac{\min(1,\delta\log x)x}{\phi(q)\log x} \leqslant \frac{\delta x}{\phi(q)}.$$

Proof. We prove this with $B = 12$ using the main result from Thorner and Zaman [**143**]. By hypothesis, $x \geqslant q^{12}$ and therefore we may apply [**143**, Corollary 1.4], obtaining

$$\begin{aligned}\pi(x;q,b) &\leqslant \sqrt{x} + \frac{2}{\log x} \sum_{\substack{\sqrt{x}<p\leqslant x\\ p\equiv 1 \pmod q}} \log p\\ &\ll \sqrt{x} + \frac{\lambda x}{\phi(q)\log x},\end{aligned}$$

where

$$\lambda := 1 - x^{-\delta}/(1-\delta) \ll \min(1,\delta\log x).$$

We have the effective estimate $\delta \gg q^{-2/3}$, hence

$$\frac{\delta x}{\phi(q)} \geqslant \frac{\delta x}{q} \gg q^{-5/3}x \gg x^{3/4}$$

and the proposition follows. □

One can leverage this irregularity to prove *regularity* results about primes that are out of reach otherwise, the most spectacular application being Heath-Brown's [**85**] deduction of the twin prime conjecture from the existence of exceptional zeros (for an appropriate c_0). See Iwaniec's survey article [**97**] for background on attempts to prove the non-existence of exceptional zeros and discussion about other applications of their existence. There are also a variety of problems where one argues in different ways depending on whether or not exceptional zeros exist, a principal example being Linnik's Theorem on primes in arithmetic progressions (see, e.g., [**61**, Ch. 24]).

Apply Proposition 8.12 with $x = q^B$. Recalling (8.34), we see that the quantity u in Theorem 8.9 satisfies

$$u \asymp \frac{\delta x\log x}{q} = q^{B-1+o(1)} \quad (q\to\infty)$$

and consequently that $\log u \asymp \log q$. Setting $X = e^{2u}$, we conclude the following:

Theorem 8.13 (Ford [**50**])**.** *Suppose that χ is a real character with conductor q and that $L(1-\delta,\chi) = 0$ for some $\frac{1}{2} \leqslant \delta < 1$. Then there is some X with $\log\log X \asymp \log q$ and*

$$G(X) \gg \frac{\log X}{\delta\log_2 X}. \tag{8.35}$$

For example, if k is fixed and there exist infinitely many exceptional zeros $\delta = \delta_q$ satisfying $\delta_q \leqslant (\log q)^{-k}$, we see that there is an unbounded set of X for which

$$G(X) \gg_k (\log X)(\log_2 X)^{k-1}.$$

this improves upon Theorem 8.1 for $k \geqslant 2$. Similarly, if $m \geqslant 2$ and there is an infinite set of q satisfying $\delta = \delta_q = q^{-1/\log_m q}$, then for an unbounded set of X,

$$G(X) > (\log X) \exp\left\{ c_m \frac{\log_2 X}{\log_{m+1} X} \right\},$$

for some constant $c_m > 0$.

Remark. We have made no use in the proof of estimates for numbers lacking large prime factors, as in the proof of Theorem 8.2. There does not seem to be any advantage to this in our argument.

9. Random models for primes, gaps, and correlations

9.1. **Cramér's model.** In 1936, Cramér [**25**] introduced a random model of primes and used it to predict that $G(x) \ll \log^2 x$. His model is based on the fact that if a random number is sampled near x, it has a likelihood of about $1/\log x$ of being prime. Now let $X_3, X_4, X_5, \dots$ be independent random variables taking values $0, 1$ and with

$$\mathbb{P}(X_n = 1) = \frac{1}{\log n}.$$

Then let $\mathcal{C} = \{n : X_n = 1\}$ be a random set of positive integers. Cramér argued that $\mathcal{C}$ behaves globally like the primes and ought to behave locally like the primes as well. We will analyze the global and local behavior, showing that with probability 1, the count of elements of $\mathcal{C}$ is similar to $\pi(x)$ under the Riemann Hypothesis, but that local correlations do not match those of primes. We will then show almost surely that the largest gap, $G_{\mathcal{C}}(x)$, between elements of $\mathcal{C}$ up to x, satisfies $G_{\mathcal{C}}(x) \sim \log^2 x$ as $x \to \infty$.

We recall the Borel-Cantelli lemma from probability theory.

Lemma 9.1 (Borel-Cantelli)**.** *Let $E_1, E_2, \dots$ be events in a probability space.*

(i) If $\sum_{j=1}^{\infty} \mathbb{P}(E_j) < \infty$, then with probability 1, only finitely many of the E_i occur;

(ii) If $\sum_{j=1}^{\infty} \mathbb{P}(E_j) = \infty$ and the the E_j are mutually independent, then with probability 1, infinitely many of the E_i occur.

Theorem 9.2. *Fix $c > 3/2$. With probability 1, we have*

$$\#\{n \leqslant x : n \in \mathcal{C}\} = \mathrm{li}(x) + O(x^{1/2} \log^c x).$$

Proof. Fix $c > 3/2$. For any integers $u \geqslant 2$ and $v \geqslant 0$, we let

$$N(u, v) := \#\{u < n \leqslant v : n \in \mathscr{C}\}$$

If $u < v \leqslant 2u$ then

$$\mathbb{E}N(u, v) = \sum_{u < n \leqslant v} \frac{1}{\log n} = \int_u^v \frac{dt}{\log t} + O(1)$$

and

$$\begin{aligned} EN(u, v)^2 &= \sum_{u < n_1, n_2 \leqslant v} \mathbb{P}\big(n_1 \in \mathcal{C} \text{ and } n_2 \in \mathcal{C}\big) \\ &= \sum_{\substack{u < n_1, n_2 \leqslant v \\ n_1 \neq n_2}} \frac{1}{(\log n_1)(\log n_2)} + \sum_{u < n \leqslant v} \frac{1}{\log n} \\ &= \bigg(\sum_{u < n \leqslant v} \frac{1}{\log n} \bigg)^2 + \sum_{u < n \leqslant v} \bigg(\frac{1}{\log n} - \frac{1}{\log^2 n} \bigg) \\ &= \big(\mathbb{E}N(u, v)\big)^2 + O\bigg(\frac{v - u}{\log u} \bigg). \end{aligned}$$

Now let

$$\Delta(u, v) = N(u, v) - \int_u^v \frac{dt}{\log t}.$$

For $u < v \leqslant 2u$ it follows that

$$\mathbb{E}\Delta(u,v)^2 \ll \frac{v-u}{\log u}. \tag{9.1}$$

Let x be a large power of two. For integers h, m with $\sqrt{x} \leqslant 2^m \leqslant x$ and $0 \leqslant h \leqslant x/2^m - 1$, let $G_{m,h}$ be the event that

$$\big|\Delta(x + h\cdot 2^m, x + (h+1)2^m)\big| \leqslant x^{1/2}(\log x)^{c-1}.$$

For large x, (9.1) and Chebyshev's inequality imply that

$$\mathbb{P}\big(\text{not } G_{h,m}\big) \ll \frac{2^m}{x(\log x)^{2c-1}}.$$

Let F_x denote the event that $G_{h,m}$ holds for all such h, m. By a union bound, we see that

$$\mathbb{P}F_x = 1 - O((\log x)^{2-2c}).$$

Since $2c - 2 > 1$, the Borel-Cantelli lemma implies that with probability one, F_{2^s} is true for all large integers s, say $s \geqslant s_0$. On this event, for all $s \geqslant s_0$, $x = 2^s$ and $1 \leqslant y \leqslant x$ we have

$$\begin{aligned}
|\Delta(x, x+y)| &= \Big| \sum_{2\sqrt{x} \leqslant 2^m \leqslant y} \Delta\Big(x + \lfloor y/2^{m+1}\rfloor 2^{m+1}, x + \lfloor y/2^m\rfloor 2^m\Big)\Big| + O(\sqrt{x}) \\
&\leqslant \sum_{2\sqrt{x} \leqslant 2^m \leqslant y} x^{1/2}(\log x)^{c-1} + O(\sqrt{x}) \\
&\ll x^{1/2}(\log x)^c.
\end{aligned}$$

Therefore, for any u satisfying $2^t \leqslant u < 2^{t+1}$ with $t \geqslant s_0$, we have

$$\#\{n \leqslant x : n \in \mathscr{C}\} = O(2^{s_0}) + \sum_{s_0 \leqslant s \leqslant t} 2^{s/2} s^c \ll x^{1/2}\log^c x. \qquad \square$$

Using the Law of the Iterated Logarithm from probability theory (e.g., Chapter VIII.5 in [**46**]), one can be more precise about the almost sure behavior of the counting function $\#\{n \leqslant x : n \in \mathcal{C}\}$, and we get

$$\limsup_{x\to\infty} \left| \frac{\#\{n \leqslant x : n \in \mathcal{C}\} - \int_2^x dt/\log t}{\left(\frac{2x\log_2 x}{\log x}\right)^{1/2}} \right| = 1.$$

Now we analyze the largest gap below x in the random set $\mathcal{C}$, which we denote by $G_{\mathcal{C}}(x)$.

Theorem 9.3. *Suppose that $\ell(n)$ and $u(n)$ are increasing sequences of positive integers such that for large n,*

$$\frac{1}{2}\log^2 n \leqslant \ell(n) \leqslant 2\log^2 n, \qquad \ell(n + \lceil 5\log^2 n\rceil) \leqslant \ell(n) + 1.$$

Suppose further that

$$\sum_{n=2}^{\infty} \frac{e^{-\ell(n)/\log n}}{\log n} = \infty, \qquad \sum_{n=2}^{\infty} \frac{e^{-u(n)/\log n}}{\log n} < \infty.$$

With probability 1, we have $G_{\mathcal{C}}(x) \geqslant \ell(x)$ for infinitely many integers x and $G_{\mathcal{C}}(x) \leqslant u(x)$ for all sufficiently large x.

For example, fix an integer $t \geqslant 4$, and the following sequences satisfy the conditions of Theorem 9.3 for sufficiently large n:

$$\ell(n) = \lfloor (\log n)(\log n + \log_3 n + \log_4 n + \cdots + \log_t n) \rfloor,$$
$$u(n) = \big\lfloor (\log n)(\log n + \log_3 n + \log_4 n + \cdots + \log_{t-1} n + 2\log_t n) \big\rfloor.$$

Proof. For $n \geqslant 3$ and $k \geqslant 1$, let $Y_{n,k}$ be the event that $n \in \mathcal{C}$ and none of the numbers $n+1, n+2, \dots, n+k$ lies in $\mathcal{C}$. That is, there is a gap of size $\geqslant k$ following n. By the independence,

$$\mathbb{P}\, Y_{n,k} = \frac{1}{\log n}\left(1 - \frac{1}{\log(n+1)}\right)\cdots\left(1 - \frac{1}{\log(n+k)}\right).$$

In particular, if $k \leqslant 10\log^2 n$ then for $1 \leqslant j \leqslant k$ we have

$$\frac{1}{\log(n+j)} = \frac{1}{\log n + O\left(\frac{\log^2 n}{n}\right)} = \frac{1}{\log n} + O\left(\frac{1}{n}\right)$$

and it follows that

$$\begin{aligned} \mathbb{P}\, Y_{n,k} &= \frac{1}{\log n}\left(1 - \frac{1}{\log n} + O\left(\frac{1}{n}\right)\right)^k \\ &= \left(\frac{1}{\log n}\right)\exp\left\{k\left(-\frac{1}{\log n} + O\left(\frac{1}{2\log^2 n}\right)\right)\right\} \\ &\asymp \frac{e^{-k/\log n}}{\log n}. \end{aligned} \tag{9.2}$$

The rough idea is that for fixed $\varepsilon > 0$ and $k = \big\lfloor (1+\varepsilon)\log^2 n \big\rfloor$, $\mathbb{P}\, Y_{n,k} \ll n^{-1-\varepsilon}$ and $\sum_n n^{-1-\varepsilon}$ converges, therefore with probability 1 only finitely many such events occur; on the other hand, if $k = \big\lfloor (1-\varepsilon)\log^2 n \big\rfloor$ then $\mathbb{P}\, Y_{n,k} \gg n^{-1+\varepsilon}(\log n)^{-1}$ and $\sum n^{-1+\varepsilon}(\log n)$ diverges, so we expect that almost surely infinitely many such events occur (one must be somewhat careful here, since Y_{n_1,k_1} and Y_{n_2,k_2} are independent only if the intervals $[n_1+1, n_1+k_1]$ and $[n_2+1, n_2+k_2]$ are disjoint).

We now make this kind of argument completely rigorous and more precise. Fix an integer $t \geqslant 4$ and for large enough n (in terms of t) let $k = u(n)$. By (9.2), for sufficiently large n_0 we have

$$\sum_{n=n_0}^{\infty} \mathbb{P}\, Y_{n,u(n)} \ll \sum_{n=n_0}^{\infty} \frac{e^{-u(n)/\log n}}{\log n} \ll 1.$$

By Borel-Cantelli, with probability 1, $Y_{n,u(n)}$ fails for all sufficiently large n. On this event, for all sufficiently large x, $G(x) \leqslant u(x)$ since $u(x)$ is increasing. This proves the second part of the theorem.

For the lower bound, define the event

$$W_{n,k} = \big(Y_{n,k} \text{ or } Y_{n+1,k} \text{ or } \cdots \text{ or } Y_{n+k-1,k}\big).$$

The events $Y_{n,k}, \dots, , Y_{n+k-1,k}$ are mutually disjoint, hence (9.2) implies that if $k \leqslant 2\log^2 n$ then

$$\mathbb{P}\, W_{n,k} \asymp \frac{k e^{-k/\log n}}{\log n}. \tag{9.3}$$

Let $n_1, n_2, \dots$ and $k_1, k_2, \dots$ be sequences of positive integers such that $n_{j+1} \geqslant n_j + 2k_j$ for each j. Since the event W_{n_j,k_j} depends only on whether or not the integers $n_j, \dots, n_j + 2k_1 - 1$ are in $\mathcal{C}$, we see that the events W_{n_j,k_j} are independent. In particular, we may take

$$n_1 = 10, \quad n_{j+1} = n_j + \lceil 5 \log^2 n_j \rceil \quad (j \geqslant 1)$$

and

$$k_j = \ell(n_{j+1}) \qquad (j \geqslant 1).$$

We have $n_{j+1} \sim n_j$ as $j \to \infty$ and for large enough j we have

$$n_j + 2k_j \leqslant n_j + 4 \log^2 n_{j+1} < n_j + 5 \log^2 n_j \leqslant n_{j+1}.$$

Thus, for some j_0, the events $W_{n_j,k_j}, j \geqslant j_0$, are mutually independent. With these definitions, if j is sufficiently large and W_{n_j,k_j} holds then there is some $x \leqslant n_{j+1}$ with $Y_{x,\ell(n_{j+1})}$ holding, and hence

$$G_{\mathcal{C}}(n_{j+1}) \geqslant \ell(n_{j+1}).$$

To prove the first part of the theorem, it suffices to show that with probability 1, there are infinitely many j with W_{n_j,k_j} holding. By Corel-Cantelli, it thus suffices to prove that $\sum_j \mathbb{P}\, W_{n_j,k_j}$ diverges.

By hypothesis, when $n_j \leqslant n \leqslant n_{j+1}$ we have $\ell(n) = \ell(n_{j+1}) + O(1)$ and

$$\frac{1}{\log n} = \frac{1}{\log n_{j+1}} + O\left(\frac{1}{n_{j+1}}\right).$$

Thus, by (9.3) and the assumed lower bound $\ell(n_{j+1}) \geqslant \frac{1}{2} \log^2 n_{j+1} \geqslant \frac{1}{2} \log^2 n_j$, we have for sufficiently large j_0 the bounds

$$\begin{aligned}
\sum_{j \geqslant j_0} \mathbb{P}\, W_{n_j,k_j} &\gg \sum_{j \geqslant j_0} \frac{\ell(n_{j+1}) e^{-\ell(n_{j+1})/\log n_j}}{\log n_j} \\
&\gg \sum_{j \geqslant j_0} (\log n_j) e^{-\ell(n_{j+1})/\log n_j} \\
&\gg \sum_{j \geqslant j_0} \sum_{n_j \leqslant n < n_{j+1}} \frac{e^{-\ell(n)/\log n}}{\log n} \\
&= \sum_{n \geqslant n_{j_0}} \frac{e^{-\ell(n)/\log n}}{\log n}.
\end{aligned}$$

By hypothesis, the right side diverges and this completes the proof. □

The Cramér model has a number of flaws, that is, statistical behavior which is known to be different from that of the primes. In particular, if $(h_1, \dots, h_k)$ is any tuple of distinct integers then

$$\mathbb{E} \#\{n \leqslant x : n + h_i \in \mathcal{C}\ \forall i\} \sim \sum_{n \leqslant x} \frac{1}{\log^k n} \sim \frac{x}{\log^k x}$$

as $x \to \infty$. Using a 2nd moment argument similar to that in Theorem 9.2, one can show that with probability 1, $\#\{n \leqslant x : n + h_i \in \mathcal{C}\ \forall i\} \sim x/\log^k x$. This not only has the wrong predictive constant (missing $\mathfrak{S}(\mathcal{H})$ factor) compared to the prime k-tuples conjecture (Conjecture 1.1), but when $(h_1, \dots, h_k)$ is inadmissible it predicts the wrong order of magnitude; for actual primes the analog of the left side is $O(1)$. For example, the Cramér model $\mathcal{C}$ has $\sim x/\log^2 x$ pairs of consecutive elements below x. As these correlations are "local statistics" and the gap function $G_{\mathcal{C}}(x)$ is also a

"local statistic", it casts some doubt in the validity that the maximal prime gap function $G(x)$ has similar behavior to $G_{\mathcal{C}}(x)$.

9.2. **Granville's refinement of Cramér's model.** Unlike the random Cramér set $\mathcal{C}$, the actual primes lie in one progression modulo 2 (with one exception), two progressions modulo 3 (with one exception), and so on. In 1990[8], Granville [**69**] proposed a variation of Cramér's random construction which corrects these flaws, by selecting only integers which are not divisible by small primes p.

For simplicity we describe Granville's model within a dyadic interval $(x/2, x]$. Let $T = \frac{\log x}{\log_2 x}$, and let

$$P = \prod_{p \leqslant T} p, \quad \theta = \prod_{p \leqslant T} (1 - 1/p).$$

The integers coprime to P form a periodic set, $\mathcal{S}$, with period $P = x^{o(1)}$ and density θ, and there are $x^{1-o(1)}$ complete periods inside $[1, x]$. We now for a random set $\mathcal{G}$ by choosing $n \in (x/2, x]$ to be in $\mathcal{G}$ with probability

$$\begin{cases} 0 & \text{if } (n, P) > 1 \\ \frac{1/\theta}{\log n} & \text{if } (n, P) = 1; \text{that is, } n \in \mathcal{S}. \end{cases}$$

As with the Cramér model, the choices are independent for different n. It is easy to compute the expected size of $\mathcal{G}$:

$$\begin{aligned} \mathbb{E}|\mathcal{G}| &= \sum_{\substack{3 \leqslant n \leqslant x \\ (n,P)=1}} \frac{1/\theta}{\log n} \\ &= \frac{1}{\theta} \sum_{\substack{1 \leqslant m \leqslant P \\ m \in \mathcal{S}}} \sum_{\substack{x/2 < n \leqslant x \\ n \equiv m \pmod P}} \frac{1}{\log n}. \end{aligned}$$

By Euler's summation formula, the inner sum is $\frac{1}{P}(\mathrm{li}(x) - \mathrm{li}(x/2)) + O(1)$ and thus

$$\mathbb{E}|\mathcal{G}| = \frac{\mathrm{li}(x) - \mathrm{li}(x/2)}{P\theta} \sum_{\substack{1 \leqslant m \leqslant P \\ m \in \mathcal{S}}} 1 + O(P/\theta) = \mathrm{li}(x) - \mathrm{li}(x/2) + O(x^{o(1)}). \tag{9.4}$$

Using a second moment method similarly to the proof of Theorem 9.2, one can show that $|\mathcal{G}|$ is concentrated near $\mathrm{li}(x)$, thus matching the global distribution of primes. Moreover, for any fixed k-tuple of non-negative integers $\mathcal{H} = (h_1, \dots, h_k)$ with $0 \leqslant h_1 < h_2 < \cdots < h_k$, let $\mathcal{G}(\mathcal{H})$ denote the number of $n \leqslant x$ for which $n + h_i \in \mathcal{G}$ for all i.

Let x be so large that $T > \max h_i$. It is clear that $\mathcal{G}(\mathcal{H})$ is empty if $\mathcal{H}$ is inadmissible, since there is a prime $\leqslant \max h_i$ for which $\mathcal{H}$ covers all residue classes modulo p and thus for any $n \in \mathbb{Z}$, $p|(n + h_i)$ for some i.

Now suppose that $\mathcal{H}$ is admissible and let $\rho(p)$ be the number of residue classes modulo p occupied by $\mathcal{H}$. Let $\mathcal{S}(\mathcal{H})$ be the set of integers n for which $n + h_i \in \mathcal{S}$ for all i; that is, $(n + h_i, P) = 1$ for

[8]Andrew Granville announced his new model at the 1990 Illinois Number Theory Conference, with KF in attendance

all i. This set is also periodic modulo P and has density

$$\nu = \prod_{p \leqslant T} \left(1 - \frac{\rho(p)}{p}\right).$$

By a similar computation to that leading to (9.4), we find

$$\begin{aligned}\mathbb{E}|\mathcal{G}(\mathcal{H})| &= \frac{1}{\theta^k} \sum_{\substack{1 \leqslant m \leqslant P \\ m \in \mathcal{S}(\mathcal{H})}} \sum_{\substack{x/2 < n \leqslant x - h_k \\ n \equiv m \pmod P}} \frac{1}{\log(n+h_1)\log(n+h_2)\cdots\log(n+h_k)} \\ &= \frac{\nu}{\theta^k} \int_{x/2}^{x} \frac{dt}{(\log t)^k} + O(x^{o(1)}).\end{aligned}$$

Since

$$\frac{\nu}{\theta^k} = \prod_{p \leqslant T} \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} = \mathfrak{S}(\mathcal{H})\left(1 + O_k\left(\frac{1}{T}\right)\right).$$

We conclude that

$$\mathbb{E}|\mathcal{G}(\mathcal{H})| \sim \mathfrak{S}(\mathcal{H})\,\frac{x/2}{\log^k x},$$

matching the conjectured asymptotic from the prime k-tuples conjecture (see Conjecture 1.1).

We now analyze the maximal gap, $G_{\mathcal{G}}(x)$, in Granville's random set. As in the proof of Theorem 9.3, we consider the events $Y_{n,k}$ that $n \in \mathcal{G}$ and none of $n+1, \dots, n+k$ are in $\mathcal{G}$. Then

$$\mathbb{P}\, Y_{n,k} = \frac{\mathbb{1}_{n \in \mathcal{S}}}{\theta \log n} \prod_{\substack{1 \leqslant \ell \leqslant k \\ n+\ell \in \mathcal{S}}} \left(1 - \frac{1}{\theta \log(n+\ell)}\right) \sim \frac{\mathbb{1}_{n \in \mathcal{S}}}{\theta \log n}\left(1 - \frac{1}{\theta \log n}\right)^{|\mathcal{S} \cap [n+1, n+k]|}.$$

Notice that if $n \in \mathcal{S}$ then the probability is greatly affected by the number of elements of $\mathcal{S}$ in he interval $[n+1, n+k]$. In particular, for any $m \in \mathbb{N}$, $mP+1 \in \mathcal{S}$ but none of $mP+2, \dots, mP+T$ are in $\mathcal{S}$ as each is divisible by some prime $p \leqslant T$. Then $\mathbb{P}\, Y_{mP+1,T-1} = 1/(\theta \log(mP+1))$. To make $\mathbb{P}\, Y_{n,k}$ large, we should choose an interval where there are few elements of $\mathcal{S}$. As $P = x^{o(1)}$, any such interval will be repeated $x^{1-o(1)}$ times in $[1, x]$. We are interested in gaps of size $k \asymp \log^2 x$. The average size of $\mathcal{S} \cap [n+1, \dots, n+k]$ equals $\theta k \sim \frac{k}{e^\gamma \log_2 x}$. However, if $k \approx (\log x)^2 = T^{2+o(1)}$, the set $\mathcal{S} \cap [2, k+1]$ contains only the primes in $(T, k+1]$ together with products $p_1 p_2$ where p_1, p_2 are primes and $T < p_1 < p_2, p_1 p_2 \leqslant k+1$. The number of primes is $\sim k/\log k \sim k/(2\log_2 x)$ and the number of products $p_1 p_2$ is

$$\leqslant \sum_{T < p_1 \leqslant \sqrt{k+1}} \pi\left(\frac{k+1}{p_1}\right) - \pi(p_1) \ll \frac{k}{\log k} \sum_{T < p_1 \leqslant \sqrt{k+1}} \frac{1}{p_1} \ll \frac{k \log_3 x}{(\log_2 x)^2}.$$

Thus, letting $A_k = |\mathcal{S} \cap [2, k+1]|$, we have

$$A_k \sim \frac{k}{2\log_2 x},$$

which is smaller than $\frac{k}{e^\gamma \log_2 x}$ by a factor $1/(2e^\gamma)$. Thus, for any $m \in \mathbb{Z}$,

$$|\mathcal{S} \cap [mP+2, mP+k+1]| = A_k$$

as well. When $k = \lfloor c(\log x)^2 \rfloor$ and $mP+1 \in (x/2, x-k]$, a similar argument as that in (9.2) yields

$$\mathbb{P}\, Y_{mP+1,k} \sim \frac{1}{\theta \log x}\left(1 - \frac{1}{\theta \log x}\right)^{A_k} \asymp \frac{e^{-A_k/(\theta \log x)}}{\theta \log x} = n^{-ce^\gamma/2+o(1)}.$$

If we choose $c < 2e^{-\gamma}$ then the exponent of n is > -1. As these events $Y_{mP+1,k}$ are independent (since $P > k$), the probability that none of them occur is

$$\big(1 - n^{-ce^\gamma/2+o(1)}\big)^{x/(2P)+O(1)} = \exp\{-n^{1-ce^\gamma/2+o(1)}\},$$

which is very tiny. Repeating the argument for $x = 2^m$, $m \in \mathbb{N}$, we see that by Borel Cantelli, almost surely we have

$$\limsup_{n\to\infty} \frac{G_{\mathcal{G}}(x)}{\log^2 x} \geqslant 2e^{-\gamma} = 1.122918\cdots. \tag{9.5}$$

This is larger than the largest almost sure gap in the Cramér model by a factor $2e^{-\gamma} = 1.229\ldots$.

Notice that this argument only provides a *lower* bound on the largest gap. Is it possible that there is another number $h \in [1, P]$ so that

$$|\mathcal{S} \cap [h+1, h+k]| \leqslant \lambda \frac{k}{\log_2 x}$$

with $\lambda < \frac{1}{2}$? If so, then this would produce a lower bound with a larger constant. The answer to this question is unknown, and is known as the "interval sieve problem", since we are effectively sieving $[h+1, h+k]$ by the primes below T. In 2019, Banks, Ford and Tao [**9**] showed that if exceptional zeros of Dirichlet L-functions exist then there are intervals with $\lambda \to 0$ as $x \to \infty$. This gives a lower bound on the maximal gap which is of a *larger order* than what Cramér's model gives.

Definition 8. *We say that* exceptional zeros exist *if there is an infinite set $\mathcal{E} \subset \mathbb{N}$, such that for every $q \in \mathcal{E}$ there is a real Dirichlet character χ_q and a zero $1 - \delta_q$ with $L(1-\delta_q, \chi_q) = 0$ and $\delta_q = o(1/\log q)$ as $q \to \infty$.*

Theorem 9.4 ([**9**], Section 2)**.** *Suppose that exceptional zeros exist. Then*

$$\limsup_{x\to\infty} \frac{G_{\mathcal{G}}(x)}{\log^2 x} = \infty.$$

Proof. By the argument leading to (9.5), it suffices to show that for every $c > 0$, if $k = \lfloor c \log^2 x \rfloor$ and x is large enough then there is a number $h \in [1, P] \cap \mathcal{S}$ so that

$$|\mathcal{S} \cap [h+1, h+k]| \leqslant \frac{\log^2 x}{2 \log_2 x}, \tag{9.6}$$

for then

$$\mathbb{P} Y_{mP+h} > n^{-e^\gamma/2+o(1)},$$

and the exponent is > -1.

We use Gallagher's prime number theorem, Proposition 8.12, in a similar way to the proof of Theorem 8.9. To show (9.6) it suffices to find reside classes $a_p \mod p$ for $p \leqslant T$ such that the number of elements of $[1, k]$ left uncovered by these residue classes is at most $\frac{\log^2 x}{2\log_2 x}$. From Definition

8, take a q with $\delta_q \leqslant \varepsilon/\log q$, where $\varepsilon > 0$ is a function of c. Let x be a large power of two such that

$$k \asymp q^{1/\sqrt{\varepsilon}}.$$

In this way, $T > q$. For $p \leqslant T$ and $p \nmid q$, choose a_p satisfying $qa_p + 1 \equiv 0 \pmod p$, and choose a_p for $p|q$ is a greedy way, similar to the proof of Theorem 8.9. Let $\mathcal{U}_1$ be the set of $n \in [1, k]$ with $n \not\equiv a_p \pmod p$ for all $p \leqslant T$, $p \nmid q$. For such $n \in \mathcal{U}_1$, $qn + 1 \leqslant qk + 1 \leqslant k^{1.1} \leqslant T^{2.5}$ and $qn + 1$ has no prime factor $< T$. Hence, $qn + 1$ is either prime or the product of two primes $> T$. By Proposition 8.12,

$$\pi(qk + 1; q, 1) \ll \frac{\delta_q qk}{\phi(q)},$$

Hence, letting $\mathcal{U}_2$ be the the set of $n \in [1, k]$ with $n \not\equiv a_p \pmod p$ for all $p \leqslant T$,

$$|\mathcal{U}_2| \leqslant \frac{\phi(q)}{q}\left[\pi(qk + 1; q, 1) + \sum_{T<p\leqslant\sqrt{qk+1}} \pi\left(\frac{qk + 1}{p}; q, \overline{p}\right)\right],$$

where $\overline{p}$ is the inverse of p modulo q. Applying the Brun-Titchmarsh theorem (Theorem 2.7) to the sum over p, and recalling that $k/p \gg \sqrt{k/q} > k^{1/3}$ if ε is small enough, we see that

$$\begin{aligned}\sum_{T<p\leqslant\sqrt{qk+1}} \pi\left(\frac{qk + 1}{p}; q, \overline{p}\right) &\ll \frac{qk}{\phi(q)\log k} \sum_{T<p\leqslant\sqrt{qk+1}} \frac{1}{p} \\ &\ll \frac{qk}{\phi(q)\log k} \cdot \frac{\log(\sqrt{qk}/T)}{\log T}.\end{aligned}$$

Since $q \asymp k^{\sqrt{\varepsilon}}$, we then conclude that

$$|\mathcal{U}_2| \ll \delta_q k + \frac{k}{\log k} \cdot \frac{\log(\sqrt{qk}/T)}{\log T} \ll \sqrt{\varepsilon}\frac{k}{\log k} \ll \sqrt{\varepsilon}\, \frac{c \log^2 x}{\log_2 x}.$$

Taking ε sufficiently small completes the proof of (9.6). □

9.3. **The Banks-Ford-Tao model of primes [9].** To be added.

9.4. **The Wintner-Montgomery model of primes based on zeros of $\zeta(s)$.** To be added.

References

[1] L. M. Adleman and D. R. Heath-Brown, *The first case of Fermat's last theorem.* Invent. Math. **79** (1985), no. 2, 409–416.

[2] M. Agrawal, N. Kayal and N. Saxena, *PRIMES is in P.* Ann. of Math. (2) **160** (2004), no. 2, 781–793; errata: Ann. of Math. (2) **189** (2019), no. 1, 317–318.

[3] W. R. Alford, Andrew Granville and Carl Pomerance, *There are infinitely many Carmichael numbers.* Ann. of Math. **139** (1994), 703–722.

[4] Antal Balog, *$p + a$ without large prime factors.* Seminar on number theory, 1983–1984 (Talence, 1983/1984), Exp. No. 31, 5 pp., Univ. Bordeaux I, Talence, 1984.

[5] Antal Balog, *Linear equations in primes.* Mathematika **39** (1992), no. 2, 367–378.

[6] Roger C. Baker and Glyn Harman, *The Brun-Titchmarsh theorem on average,* Analytic Number Theory (Proceedings in honor of Heini Halberstam), Birkhauser, Boston, 1996, 39–103.

[7] Roger C. Baker and Glyn Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–361.

[8] Roger C. Baker, Glyn Harman and Janos Pintz, *The difference between consecutive primes. II.*, Proc. London Math. Soc. (3) **83** (2001), no. 3, 532–562.

[9] William Banks, Kevin Ford and Terence Tao, *Large prime gaps and probabilistic models*, submitted. `Arxiv: 1908.08613`
[10] Paul T. Bateman, *The distribution of values of the Euler function.* Acta Arith. **21** (1972), 329–345.
[11] Paul T. Bateman and R. A. Horn, *Primes represented by irreducible polynomials in one variable*, Proc. Sympos. Pure Math., Vol. VIII, Amer. Math. Soc., Providence, R.I. (1965), 119–132.
[12] Patrick Billingsly, *On the distribution of large prime divisors*, Collection of articles dedicated to the memory of Alfréd Rényi, I. Period. Math. Hungar. **2** (1972), 283–289.
[13] Enrico Bombieri, *On the large sieve.* Mathematika **12** (1965), 201–225.
[14] Enrico Bombieri and Harold Davenport, *Small differences between prime numbers.* Proc. Roy. Soc. London Ser. A **293** (1966), 1–18.
[15] Enrico Bombieri, John Friedlander and Henryk Iwaniec, *Primes in arithmetic progressions to large moduli. II*, Math. Ann. **277** (1987), 361–393.
[16] V. Y. Bouniakowsky, *Nouveaux théorèmes relatifs à la distinction des nombres premiers et à la décomposition des entiers en facteurs*, Mém. Acad. Sci. St. Pétersbourg (6) (1857), 305–329.
[17] Viggo Brun, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Archiv for Math. og Naturvid. B **34** (1915), no. 8, 19pp.
[18] Viggo Brun, *Om fordelingen av primtallene i forskjellige talklasser. En ovre begrænsning*, Nyt Tidsskr. f. Math. **27** B (1916), 45–58.
[19] Viggo Brun, *La série 1/5+1/7+1/11+1/13+ 1/17+1/19+1/29+1/31+1/41+1/43+1/59+1/61+... où les dénominateurs sont "nombres premiers jumeaux" est convergente ou finie*, Bull. Sci. Math. (2) **43** (1919), 100–104, 124–128.
[20] Viggo Brun, *Le crible d'Eratosthène et le théorème de Goldbach*, Skr. Norske Vid.-Akad. Kristiania I. **1920**, no. 3, 36 pp.
[21] Robert D. Carmichael, *On Euler's ϕ-function* , Bull. Amer. Math. Soc. **13** (1907) , 241–243.
[22] Robert D. Carmichael, *Note on Euler's ϕ-function* , Bull. Amer. Math. Soc. **28** (1922) , 109–110.
[23] Alina C. Cojocaru and M. Ram Murty. *An introduction to sieve methods and their applications.* London Mathematical Society Student Texts, vol. **66**. Cambridge University Press, 2006.
[24] Harald Cramér, *Some theorems concerning prime numbers*, Ark. Mat. Astr. Fys. **15** (1920), 1–33.
[25] Harald Cramér, *On the order of magnitude of the difference between consecutive prime numbers,* Acta Arith. **2** (1936), 396–403.
[26] Harold Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Mathematics vol. 74, Springer-Verlag, New York, 2000.
[27] Harold Diamond, Heini Halberstam and William F. Galway, *A Higher-Dimensional Sieve Method: With Procedures for Computing Sieve Functions*, Cambridge Tracts in Mathematics, vol. 177, 2008.
[28] L. E. Dickson, *A new extension of Dirichlet's theorem on prime numbers*, Messenger Math., **33** (1904), 155–161.
[29] P. G. L. Dirichlet, P. G. L. *Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält* [Proof of the theorem that every unbounded arithmetic progression, whose first term and common difference are integers without common factors, contains infinitely many prime numbers], Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften zu Berlin, 48 (1837), 45–71.
[30] P. Donnelly and G. Grimmett, *On the asymptotic distribution of large prime factors*, J. London Math. Soc. (2) **47** (1993), 395–404.
[31] P. D. T. A. Elliott and Heini Halberstam, *A conjecture in prime number theory*, Symp. Math. **4** (1968), 59–72.
[32] Paul Erdős, *On the normal number of prime factors of $p-1$ and some related problems concerning Euler's ϕ-function* , Quart. J. Math. Oxford **6** (1935), 205–213.
[33] Paul Erdős, *On the difference of consecutive primes* , Quart. J. Math. Oxford Ser. **6** (1935), 124–128.
[34] Paul Erdős, *The difference of consecutive primes.* Duke Math. J. **6** (1940), 438–441.
[35] Paul Erdős, *Some remarks on Euler's ϕ-function and some related problems* , Bull. Amer. Math. Soc. **51** (1945), 540–544.
[36] Paul Erdős, *On integers of the form 2^k+p and some related problems*, Summa Brasil. Math. **2** (1950). 113–123.
[37] Paul Erdős, ———, *Some remarks on number theory*, Riveon Lematematika **9** (1955), 45–48, (Hebrew. English summary).
[38] Paul Erdős, *Some remarks on Euler's ϕ-function* , Acta Arith. **4** (1958), 10–19.
[39] Paul Erdős, *An asymptotic inequality in the theory of numbers*, Vestnik Leningrad. Univ. **15** (1960), no. 13, 41–49, (Russian).

[40] Paul Erdős, *On the distribution of numbers of the form $\sigma(n)/n$ and on some related questions,* Pacific J. Math. **52** (1974), 59–65.
[41] Paul Erdős and Richard R.Hall, *On the values of Euler's ϕ-function* , Acta Arith. **22** (1973), 201–206.
[42] Paul Erdős and Richard R.Hall, *Distinct values of Euler's ϕ-function* , Mathematika **23** (1976), 1–3.
[43] Paul Erdős and Marc Kac, *The Gaussian law of errors in the theory of additive number theoretic functions.* Amer. J. Math. **62** (1940), 738–742.
[44] Paul Erdős and Carl Pomerance, *On the normal number of prime factors of $\phi(n)$* , Rocky Mountain J. of Math. **15** (1985), 343–352.
[45] Paul Erdős, Carl Pomerance, and Andrá s Sárkőzy, *On locally repeated values of certain arithmetic functions. III.* Proc. Amer. Math. Soc. **101** (1987), no. 1, 1–7.
[46] William Feller. *An introduction to probability theory and its applications. Vol. I. Third edition.* John Wiley and Sons, Inc., New York–London–Sydney 1968, xviii+509 pp.
[47] Kevin Ford, *The distribution of totients*, Ramanujan J. (Paul Erdős memorial issue) **2** (1998), 67–151.
[48] Kevin Ford, *The number of solutions of $\phi(x) = m$*, Annals of Math. **150** (1999), 283–311.
[49] Kevin Ford, *The distribution of integers with a divisor in a given interval*, Ann. Math. (2) **168** (2008), 367–433.
[50] Kevin Ford, *Large prime gaps and progressions with few primes. Kevin Ford.* Rivista di Matematica della Universita di Parma , vol. 12, no. 1, (2021), 41–47. Proceedings of Second Symposium on Analytic Number Theory Cetraro, Italy, July 8–12, 2019.
[51] Kevin Ford, Lithuanian Math. J. **61** (3), (2021), 323–329. The "Kubilius 100" special volume.
[52] Kevin Ford, *Solutions of $\phi(n) = \phi(n + k)$ and $\sigma(n) = \sigma(n + k)$*, IMRN **2022** (2022), Issue 5, Pages 3561–3570.
[53] Kevin Ford, Ben Green, Sergei Konyagin, and Terence Tao, *Large gaps between consecutive prime numbers*, Ann. Math. **183** (2016), 935–974.
[54] Kevin Ford, Ben Green, Sergei Konyagin, James Maynard and Terence Tao, *Long gaps between consecutive prime numbers*, J. Amer. Math. Soc., **31** (2018), no. 1, 65–105.
[55] Kevin Ford and Heini Halberstam, *The Brun-Hooley sieve*, J. Number Th. **81** (2000), 335–350.
[56] Kevin Ford and Sergei Konyagin, *On two conjectures of Sierpiński concerning the arithmetic functions σ and ϕ*, Number Theory in Progress (Zakopane, Poland, 1997), vol. II, de Gruyter (1999), 795-803.
[57] Étienne Fouvry, *Théorème de Brun-Titchmarsh: application au théorème de Fermat.* (French) [The Brun-Titchmarsh theorem: application to the Fermat theorem], Invent. Math. **79** (1985), no. 2, 383–407.
[58] John Friedlander, *Shifted primes without large prime factors*, in Number theory and applications (Banff, AB, 1988) , Kluwer Acad. Publ., Dorbrecht (1989), 393–401.
[59] John Friedlander and Henryk Iwaniec, *The polynomial $X^2 + Y^4$ captures its primes* Ann. of Math. (2), **148** (3) (1998), 945–1040.
[60] John Friedlander and Henryk Iwaniec, *The asymptotic sieve*, Ann. of Math. (2), **148** (3) (1998), 1041–1065.
[61] John Friedlander and Henryk Iwaniec, *Opera de Cribro*, American Math. Soc., 2010.
[62] P. X. Gallagher,*A large sieve density estimate near $\sigma = 1$*, Invent. Math. **11** (1970), 329–339.
[63] M. Goldfeld, *On the number of primes p for which $p+a$ has a large prime factor*. Mathematika **16** (1969), 23–27.
[64] Daniel A. Goldston, Janos Pintz and Cem Y. Yıldırım, *Primes in tuples. I*, Ann. of Math. **170** (2009), no. 2, 819–862.
[65] Daniel A. Goldston, Janos Pintz and Cem Y. Yıldırım, *Primes in tuples. II.* Acta Math. **204** (2010), no. 1, 1–47.
[66] Daniel A. Goldston, Sidney W. Graham, Janos Pintz and Cem Y. Yıldırım, *Small gaps between primes or almost primes*, Trans. Amer. Math. Soc. **361** (2009), no. 10, 5285–5330.
[67] Daniel A. Goldston, Sidney W. Graham, Janos Pintz and Cem Y. Yıldırım, *Small gaps between products of two primes*, Proc. London Math. Soc. (3) **98** (2009), 741–774.
[68] Sidney W. Graham, J. J. Holt and Carl Pomerance, *On the solutions to $\phi(n) = \phi(n + k)$*. Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997), 867–882, de Gruyter, Berlin, 1999.
[69] Andrew Granville, *Harald Cramér and the distribution of prime numbers*, Scandanavian Actuarial J. **1** (1995), 12–28.
[70] The Great Internet Mersenne Prime Search, `https://www.mersenne.org/`
[71] Ben J. Green and Terence C. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. **167** (2008), 481–547.
[72] Ben J. Green and Terence C. Tao, *Linear equations in primes,* Annals of Math. **171** (2010), no. 3, 1753–1850.
[73] Ben J. Green and Terence C. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Annals of Math. **175** (2012), no. 2, 465–540.

[74] Ben J. Green, Terence C. Tao and Tamar Ziegler, *An inverse theorem for the Gowers U^4-norm,* Glasg. Math. J. **53** (2011), no. 1, 1–50.
[75] Ben J. Green, Terence C. Tao and Tamar Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Annals of Math. **176** (2012), 1231–1372.
[76] Heini Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[77] Richard R. Hall and Gérald Tenenbaum, *Divisors,* Cambridge Tracts in mathematics vol. **90**, 1988.
[78] G. H. Hardy and John E. Littlewood, *Some problems of 'partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math., **114** (3) (1923), 215–273.
[79] G. H. Hardy and John E. Littlewood, *Some problems of 'partitio numerorum'; VII.*, unpublished.
[80] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n* , Quart. J. Math. **48** (1917), 76–92.
[81] Glyn Harman, On the number of Carmichael numbers up to x. Bull. London Math. Soc. **37** (2005), no. 5, 641–650.
[82] Glyn Harman, *Prime detecting sieves*, London Math. Soc. monographs, 2006.
[83] Glyn Harman, Watt's mean value theorem and Carmichael numbers. Int. J. Number Theory **4** (2008), no. 2, 241–248.
[84] D. R. Heath-Brown, *Gaps between primes, and the pair correlation of zeros of the zeta function*, Acta Arith. **41** (1982), no. 1, 85–99.
[85] D. R. Heath-Brown, *Prime twins and Siegel zeros*, Proc. London Math. Soc. (3) **47** (1983), 193—224.
[86] D. R. Heath-Brown, *The divisor function at consecutive integers* Mathematika **31** (1984), no. 1, 141–149.
[87] D. R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford Ser. (2) **37** (1986), no. 145, 27–38.
[88] H. A. Helfgott, *The ternary Goldbach conjecture is true*, `arXiv: 1312.7748.`
[89] A. Hildebrand and Gé rald Tenenbaum , *Integers without large prime factors* , J. Théor. Nombres Bordeaux **5** (1993), 411–484.
[90] Christopher Hooley, *On Artin's conjecture*, J. Reine Angew. Math. **225** (1967), 209–220.
[91] Christopher Hooley, *On the Brun-Titchmarsh theorem.* J. Reine Angew. Math. **255** (1972), 60–79.
[92] Christopher Hooley, *On the largest prime factor of $p + a$.* Mathematika **20** (1973), 135–143.
[93] Christopher Hooley, *On the Brun-Titchmarsh theorem. II.* Proc. London Math. Soc. (3) **30** (1975), 114–128.
[94] Christopher Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge University Press, 1976.
[95] Christopher Hooley,*On an almost-pure sieve*, Acta Arith. **LXVI** (1994), 359–368.
[96] Henryk Iwaniec, *On the error term in the linear sieve.* Acta Arith. **19** (1971), 1–30.
[97] H. Iwaniec, *Conversations on the exceptional character*, in the book *Analytic number theory*, lectures given at the C.I.M.E. summer school in Cetraro, Italy, (A. Perelli, C. Viola, eds.), Lecture Notes in Mathematics vol. 1891, Springer-Verlag 2002, p. 97–132.
[98] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloquium Publications vol. 53, 2004.
[99] R. Jajte, *On random partitions of the segment*, Bull. Acad. Polon. Sci. **19** (1971), 231–233.
[100] Dimitrios Koukoulopoulos, *The distribution of prime numbers*, Graduate Studies in Math. vol. 203, Amer. Math. Soc., 2019.
[101] Eduard Landau, *Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes.* Mathematische Annalen. **56** (4) (1903), 645–670.
[102] Junxian Li, Kyle Pratt and George Shakan, *A lower bound for the least prime in an arithmetic progression.* Q. J. Math. Oxford **68** (2017), no. 3, 729–758.
[103] Jared Lichtman, *Primes in arithmetic progressions to large moduli and shifted primes without large prime factors*, preprint, November 2022. `arXiv:2211.09641`.
[104] Yuri V. Linnik, *On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon.*, Rec. Math. (Math. Sbornik), N. S. **15** (57), (1944), 347–368.
[105] Helmut Maier, *Small differences between prime numbers.* Michigan Math. J. **35** (1988), no. 3, 323–344.
[106] Helmut Maier and Carl Pomerance, *On the number of distinct values of Euler's ϕ-function* , Acta Arith. **49** (1988), 263–275.
[107] Helmut Maier and Carl Pomerance, *Unusually large gaps between consecutive primes.* Trans. Amer. Math. Soc. **322** (1990), no. 1, 201–237.
[108] James Maynard, *Small gaps between primes*, Annals of Math. **181** (2015), 383–413.
[109] James Maynard, *Dense clusters of primes in subsets*, Compos. Math. **152** (2016), no. 7, 1517–1554.

[110] James Maynard, *Large gaps between primes*, Ann. Math. **183** (2016), 915–933.
[111] Hugh L. Montgomery and Robert C. Vaughan, *On the large sieve*, Mathematika **20** (1973), 119–134.
[112] Yoichi Motohashi, *A note on the least prime in an arithmetic progression with a prime difference.* Acta Arith. **17** (1970), 283–285.
[113] Yoichi Motohashi and Janos Pintz, *A smoothed GPY sieve*, Bull. Lond. Math. Soc. **40** (2008), no. 2, 298–310.
[114] The Multiply Perfect Numbers Page, `http://wwwhomes.uni-bielefeld.de/achim/mpn.html`
[115] The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/`
[116] S. Pillai , *On some functions connected with* $\phi(n)$, Bull. Amer. Math. Soc. **35** (1929), 832–836.
[117] Janos Pintz, *Very large gaps between consecutive primes.* J. Number Theory **63** (1997), no. 2, 286–301.
[118] Paul Pollack and Carl Pomerance, *Some problems of Erdős on the sum-of-divisors function*, Trans. Amer. Math. Soc. Ser. B **3** (2016), 1–26.
[119] D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), 2067–2199.
[120] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded gaps between primes*, Research in the Mathematical Sciences 1:12 (2014).
[121] Carl Pomerance, *Popular values of Euler's function.* Mathematika **27** (1980), no. 1, 84–89.
[122] Carl Pomerance, *A note on the least prime in an arithmetic progression,* J. Number Theory **12** (1980) 218–223.
[123] Carl Pomerance, *On the distribution of the values of Euler's function* , Acta Arith. **47** (1986), 63–70.
[124] Robert A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. **13** (1938), 242–247.
[125] Robert A. Rankin, *The difference between consecutive prime numbers. V,* Proc. Edinburgh Math. Soc. (2) 13 (1962/63), 331–332.
[126] G. Ricci, *Ricerche arithmetiche sui polinomi II (Intorno a una proposizione non vera di Legendre)*, Rend. Circ. Mat. di Palermo **58** (1934), 190–208.
[127] N. P. Romanoff, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. **109** (1934), 668–678.
[128] Andzej Schinzel, *Sur l'equation* $\phi(x) = m$, Elem. Math. **11** 1956, 75–78.
[129] Andzej Schinzel, *Sur l'équation* $\phi(x + k) = \phi(x)$.. (French) Acta Arith **4** (1958), 181–184.
[130] Andzej Schinzel, *Remarks of the paper "Sur certaines hypothèses concernant les nombres premiers"* , Acta Arith. **7** (1961/62), 1–8.
[131] Andzej Schinzel and Wacław Sierpiński, Sur certaines hypothèses concernant les nombres premiers, Acta Arith., **4** (1958), 185–208. erratum, **5** (1959), p. 259.
[132] A. Schinzel and A. Wakulicz, *Sur l'équation* $\phi(x + k) = \phi(x)$. *II.* (French) Acta Arith. **5** (1959), 425–426.
[133] L. Shnirelmann, *Über additive Eigenschaften von Zahlen*, Math. Ann. **107** (1933), 649–690.
[134] A. Schönhage, *Eine Bemerkung zur Konstruktion grosser Primzahllücken*, Arch. Math. **14** (1963), 29–30.
[135] Atle Selberg, *The general sieve-method and its place in prime number theory.* Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 1, pp. 286–292. Amer. Math. Soc., Providence, R. I., 1952.
[136] Atle Selberg, *Lectures on sieves*, Atle Selberg's collected papers, vol. II, Springer-Verlag, 1991, pp. 65–247.
[137] Daniel Shanks, *On maximal gaps between successive primes.* Math. Comp. **18** (1964), 646–651.
[138] Wacław Sierpiński, *Sur une propriété de la fonction* $\phi(n)$. (French) Publ. Math. Debrecen **4** (1956), 184–185.
[139] T. Oliveira e Silva, S. Herzog, S. Pardi, *Empirical verification of the even Goldbach conjecture and computation of prime gaps up to* 4×10^{18}, Math. Comp. **83** (2014), 2033–2060.
[140] Gérald Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory, 3rd ed.*, Amer. Math Soc., 2015
[141] N. M. Timofeev, *Hardy-Ramanujan and Hálasz type inequalities for shifted prime numbers*, Math. Notes **57** (1995), 522–535.
[142] E. C. Titchmarsh, *A divisor problem*, Rend. Circ. Mat. Palermo **54** (1930), 414–429.
[143] J. Thorner and A. Zaman, *Refinements to the prime number theorem for arithmetic progressions*, preprint. `arXiv:2108.10878`.
[144] Christian Tudesq. Majoration de la loi locale de certaines fonctions additives. *Arch. Math. (Basel)*, 67(6):465–472, 1996.
[145] Ivan M. Vinogradov, *The method of trigonometric sums in the theory of numbers*, Trav. Inst. Math. Stekloff **23** (1947), 109 pp. Russian. English translation: *The method of trigonometrical sums in the theory of numbers. Translated, revised and annotated by K. F. Roth and Anne Davenport.* Interscience Publishers, London-New York, 1953. x+180 pp.
[146] E. Westzynthius, *Über die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind,* Commentationes Physico–Mathematicae, Societas Scientarium Fennica, Helsingfors **5**, no. 25, (1931) 1–37.

[147] Eduard Wirsing, Bemerkung zu der Arbeit über vollkommene Zahlen. (German) Math. Ann. **137** (1959), 316–318.
[148] Kent Wooldridge, *Values taken many times by Euler's phi-function.* Proc. Amer. Math. Soc. **76** (1979), no. 2, 229–234.
[149] T. Xylouris, *On the least prime in an arithmetic progression and estimates for the zeros of Dirichlet L-functions.* Acta Arith. **150** (2011), no. 1, 65–91.
[150] T. Xylouris, *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression.* (German) [The zeros of Dirichlet $L-$functions and the least prime in an arithmetic progression], Dissertation for the degree of Doctor of Mathematics and Natural Sciences at the University of Bonn, Bonn, 2011. Bonner Mathematische Schriften [Bonn Mathematical Publications], 404. Universität Bonn, Mathematisches Institut, Bonn, 2011. 110 pp.
[151] T. Yamada, *On equations* $\sigma(n) = \sigma(n+k)$ *and* $\phi(n) = \phi(n+k)$. J. Comb. Number Theory **9** (2017), no. 1, 15–21.
[152] Yitang Zhang, *Bounded gaps between primes.* Annals of Mathematics **179** (2014), 1121–1174.

Department of Mathematics, 1409 West Green Street, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
Email address: ford@math.uiuc.edu