Lectures on sieves

D.R. Heath-Brown
Mathematical Institute, Oxford

Preface

These are notes of a series of lectures on sieves, presented during the Special Activity in Analytic Number Theory, at the Max-Planck Institute for Mathematics in Bonn, during the period January–June 2002. The notes were taken by Boris Moroz, and it is a pleasure to thank him for writing them up in the current form. In addition, thanks are due to the institute for its hospitality and financial support, and also to the American Institute of Mathematics, where these notes were put into final form.

Being lecture notes they are not intended to have the formal style of a textbook. Nor is there any significant claim to novelty. In particular the material in §§3 & 4 owes much to the book of Halberstam and Richert [6]. In addition to the latter work the reader may find it helpful also to consult Greaves' exposition [5].

1 Introduction

Here are some typical questions which can be expressed as sieve problems:

(i) Is every even integer $n \geq 2$ a sum of two primes? (*Goldbach's conjecture*);

(ii) Are there infinitely many pairs of primes (p, q) with $q = p + 2$? (*the twin primes conjecture*);

(iii) Are there arbitrary long arithmetic progressions consisting only of primes ? ;

(iv) Are there infinitely many primes of the form $n^2 + 1$ with $n \in \mathbb{N}$? ;

(v) Is it true that, for every $n \in \mathbb{N}$, there is a prime p in the range $n^2 < p < (n+1)^2$? ;

(vi) For every $\varepsilon > 0$ is there an integer $N(\varepsilon)$ such that the interval $[N, N + N^{\varepsilon}]$ contains a square-free number, as soon as $N \geq N(\varepsilon)$? ;

(vii) Is it true that, for some $\gamma \leq 1/4$ and every sufficiently large x, there is a pair of natural numbers (m, n) such that $x \leq m^2 + n^2 < x + x^\gamma$?.

Exercise. Prove that

$$\{(m, n) : m, n \in \mathbb{N},\ x \leq m^2 + n^2 < x + 8x^{1/4}\} \neq \emptyset$$

for all sufficiently large x.

Let $\mathcal{A}$ be a finite subset of $\mathbb{N}$, and $\mathcal{P} \subset \mathbb{P}$, where $\mathbb{P}$ denotes the set of all primes. For any positive real z, set

$$S(\mathcal{A}, \mathcal{P}; z) := \#\ \{n \in \mathcal{A} :\ p|n \Rightarrow p \geq z \text{ for } p \in \mathcal{P}\}.$$

Example 1. Let $\mathcal{A} = \{n \in \mathbb{N} :\ n \leq x\}$ and $\mathcal{P} = \mathbb{P}$. Then

$$S(\mathcal{A}, \mathcal{P}; z) := \#\ \{n \in \mathbb{N} :\ n \leq x,\ p|n \Rightarrow p \geq z \text{ for } p \in \mathbb{P}\},$$

so that $S(\mathcal{A}, \mathcal{P}; z) = \pi(x) - \pi(z)$ for $\sqrt{x} < z \leq x$, where, as usual,

$$\pi(x) := \#\ \{p \in \mathbb{P} :\ p \leq x\}.$$

Example 2. Let $\mathcal{A} = \{n(2N - n) : n \in \mathbb{N},\ 2 \leq n \leq 2N - 2\}$ and $\mathcal{P} = \mathbb{P}$. Then

$$\begin{aligned} & S(\mathcal{A}, \mathcal{P}; \sqrt{2N}) \\ = \ & \#\ \{p \in \mathbb{P} :\ 2N - p \in \mathbb{P},\ \sqrt{2N} \leq p, 2N - p \leq 2N - 2\}. \end{aligned}$$

This relates to Goldbach's problem.

Example 3. Let

$$\mathcal{A} = \{n \in \mathbb{N} :\ n < x\}, \quad \mathcal{P} = \{p \in \mathbb{P} :\ p \equiv 3 \pmod 4\}.$$

Then

$$S(\mathcal{A}, \mathcal{P}; x) := \#\ \{n \in \mathbb{N} :\ n < x,\ p \nmid n \text{ for } p \in \mathbb{P},\ p \equiv 3 \pmod 4\},$$

so that

$$S(\mathcal{A}, \mathcal{P}; x) = \#\ \{n \in \mathbb{N} :\ n < x,\ n = l^2 + m^2 \text{ with } l, m \in \mathbb{N},\ \text{h.c.f.}(l, m) = 1\}.$$

We can therefore detect sums of two squares.

Let now $\Pi := \Pi(\mathcal{P}, z) := \prod_{p < z,\, p \in \mathcal{P}} p$, and let $\mathcal{A}_d := \{n \in \mathbb{N} :\ nd \in \mathcal{A}\}$. It follows that

$$S(\mathcal{A}, \mathcal{P}; z) = \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi) = 1}} 1 = \sum_{\substack{n \in \mathcal{A} \\ d | (n, \Pi)}} \mu(d) = \sum_{d | \Pi(\mathcal{P}, z)} \mu(d) \#\ \mathcal{A}_d. \tag{1.1}$$

Main assumption. *Suppose that*

$$\# \mathcal{A}_d = \frac{\omega(d)}{d} X + R_d, \tag{1.2}$$

with an absolutely multiplicative $\omega(d)$*, satisfying the conditions* $\omega(d) \geq 0$ *for* $d \in \mathbb{N}$ *and* $\omega(p) = 0$ *for* $p \in \mathbb{P} \setminus \mathcal{P}$*.*

In particular, $\# \mathcal{A} = X + R_1$. We think of the remainder R_d term as being small compared to the main term $\frac{\omega(d)}{d} X$.

In the notation of Example 2, one may write

$$\begin{aligned} \# \mathcal{A}_d &= \#\{n(2N-n) : n \in \mathbb{N},\ 2 \leq n \leq 2N-2,\ d|n(2N-n)\} \\ &= \sum_{\substack{m (\bmod d) \\ d|m(2N-m)}} \#\{n : 2 \leq n \leq 2N-2,\ n \equiv m \pmod d\} \\ &= \sum_{\substack{m (\bmod d) \\ d|m(2N-m)}} (\frac{2N}{d} + r(d,m)), \end{aligned}$$

where $r(d,m) \ll 1$. Thus if $\omega(d) = \# \{m (\bmod d) : d|m(2N-m)\}$, then

$$\# \mathcal{A}_d = \frac{\omega(d)}{d} X + R_d$$

with $X = 2N$ and so, for Example 2, we have

$$R_d = \sum_{\substack{m (\bmod d) \\ d|m(2N-m)}} r(d,m) \ll \omega(d).$$

In general it follows from (1.1) and (1.2) that

$$S(\mathcal{A}, \mathcal{P}; z) = X \sum_{d|\Pi(\mathcal{P},z)} \frac{\mu(d)\omega(d)}{d} + \sum_{d|\Pi(\mathcal{P},z)} \mu(d) R_d. \tag{1.3}$$

Writing $W(z;\omega) = \prod_{p<z,\, p \in \mathcal{P}} (1 - \frac{\omega(p)}{p})$ we deduce the following from (1.3).

Corollary 1.1 *We have*

$$|S(\mathcal{A}, \mathcal{P}; z) - XW(z;\omega)| \leq \sum_{d|\Pi(\mathcal{P},z)} |R_d|.$$

This corollary is known as the *Sieve of Eratosthenes –Legendre.*

Lemma 1.1 The Mertens Formula. *Let* $V(z) := \prod_{p<z,\, p\in\mathbb{P}} (1-1/p)$. *One has*

$$V(z) = \frac{e^{-\gamma}}{\log z} + O(\frac{1}{(\log z)^2}). \tag{1.4}$$

Proof. See, for instance, Prachar [16, pp. 80-81.]

Example 1 (continued). We have

$$\mathcal{A}_d = \{n \in \mathbb{N} :\ nd \le x\}, \quad \#\mathcal{A}_d = [\frac{x}{d}] = \frac{X}{d} + R_d$$

and $\mathcal{P} = \mathbb{P}$, so that we may take $X = x,\ \omega(d) = 1,\ |R_d| \le 1$. Therefore it follows from Corollary 1.1 and Lemma 1.1 that

$$S(\mathcal{A}, \mathcal{P}; z) = x\frac{e^{-\gamma}}{\log z}\{1 + O(1/(\log z))\} + O(2^{\pi(z)}),$$

since

$$\sum_{d|\Pi(\mathcal{P},z)} |R_d| \le \sum_{d|\Pi(\mathcal{P},z)} 1 = 2^{\pi(z)}.$$

Thus

$$S(\mathcal{A}, \mathcal{P}; z) \sim x\frac{e^{-\gamma}}{\log z}$$

as soon as $z \le \log x$, since

$$\begin{aligned}
2^{\pi(z)} &= e^{\pi(z)\log 2}\\
&= \exp(\frac{z}{\log z}\{1 + O(1/\log z)\}\log 2)\\
&\le \exp(\frac{\log x}{\log\log x}\{1 + O(1/\log\log x)\}\log 2)\\
&\le \exp(\frac{1}{2}\log x)\\
&= x^{1/2}\\
&= o(x/\log z)
\end{aligned}$$

for $z \le \log x,\ x \to \infty$.

Corollary 1.2 *For* $z \le \log x$ *we have*

$$\#\,\{n \in \mathbb{N} :\ n \le x,\ p|n \Rightarrow p \ge z \mathit{\ for\ } p \in \mathbb{P}\} \sim \frac{e^{-\gamma}x}{\log z}.$$

We would like to extend the admissible range of z in such a result. The proof of Corollary 1.2 uses the fact that

$$\sum_{d|n,\ d|\Pi(\mathcal{P},z)} \mu(d) = \begin{cases} 1, & (n,\Pi(\mathcal{P},z)) = 1, \\ 0, & (n,\Pi(\mathcal{P},z)) > 1. \end{cases}$$

However the sum here is over an uncomfortably large range. We therefore replace the above equality with two inequalities, and encounter the following.
Sieve problem. Find two real-valued functions $\mu^+(d)$ and $\mu^-(d)$, of suitably small support, satisfying the conditions

$$\sum_{d|n,\ d|\Pi(\mathcal{P},z)} \mu^-(d) \leq \begin{cases} 1, & (n,\Pi(\mathcal{P},z)) = 1, \\ 0, & (n,\Pi(\mathcal{P},z)) > 1, \end{cases} \tag{1.5}$$

and

$$\sum_{d|n,\ d|\Pi(\mathcal{P},z)} \mu^+(d) \geq \begin{cases} 1, & (n,\Pi(\mathcal{P},z)) = 1, \\ 0, & (n,\Pi(\mathcal{P},z)) > 1. \end{cases} \tag{1.6}$$

It follows from (1.6) for example, that

$$\begin{aligned} S(\mathcal{A},\mathcal{P};z) &= \sum_{n\in\mathcal{A},\ (n,\Pi(\mathcal{P},z))=1} 1 \\ &\leq \sum_{\substack{n\in\mathcal{A},\ d|n \\ d|\Pi(\mathcal{P},z)}} \mu^+(d) \\ &= \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)\#\,\mathcal{A}_d \\ &= X \sum_{d|\Pi(\mathcal{P},z)} \frac{\mu^+(d)\omega(d)}{d} + \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)R_d. \end{aligned} \tag{1.7}$$

Hence

$$S(\mathcal{A},\mathcal{P};z) \leq X \sum_{d|\Pi(\mathcal{P},z)} \frac{\mu^+(d)\omega(d)}{d} + \sum_{d|\Pi(\mathcal{P},z)} |\mu^+(d)||R_d|. \tag{1.8}$$

To minimise the right-hand side of (1.8), subject to the condition (1.6), is in general a challenging unsolved problem.

Some achievements of sieve methods.

(i) We have

$$\#\left\{(p,p') : p,p' \in \mathbb{P},\ p+p' = 2n\right\} \ll \frac{\sigma(n)}{n}\frac{n}{(\log n)^2}$$

where $\sigma(n)$ is the sum of divisors function. This is conjectured to be best possible, up to the value of the implied constant.

(ii) We have (Chen [3])

$$\#\left\{(p,p') : p \in \mathbb{P},\ p' \in \mathbb{P}_2,\ p+p' = 2n\right\} \gg \frac{n}{(\log n)^2},$$

where $\mathbb{P}_2$ is the set of positive integers which are either prime or a product of two primes.

(iii) We also have (Chen [3])

$$\#\left\{p \in \mathbb{P} :\ p \le x,\ p+2 \in \mathbb{P}_2\right\} \gg \frac{x}{(\log x)^2}.$$

(iv) We have (Iwaniec [11])

$$\#\left\{n \le x : n^2+1 \in \mathbb{P}_2\right\} \gg \frac{x}{\log x}.$$

(iv) We have (Heath-Brown [7])

$$\#\left\{n \equiv l \ (\mathrm{mod}\ k) :\ n \in \mathbb{P}_2,\ n \le k^2\right\} \gg \frac{k^2}{\phi(k)\log k}$$

for large enough k, if h.c.f.$(l,k) = 1$.

2 Selberg's sieve

To satisfy (1.6), let

$$\mu^+(d) = \sum_{d=[d_1,d_2]} \lambda_{d_1}\lambda_{d_2} \tag{2.1}$$

with $\lambda_1 = 1$ and $\lambda_d \in \mathbb{R}$. Clearly,

$$\sum_{d|n,\ d|\Pi(\mathcal{P},z)} \mu^+(d) = \sum_{[d_1,d_2]|(n,\Pi(\mathcal{P},z))} \lambda_{d_1}\lambda_{d_2}$$

$$= \sum_{d_1,d_2|(n,\Pi(\mathcal{P},z))} \lambda_{d_1}\lambda_{d_2} = \Big(\sum_{d|(n,\Pi(\mathcal{P},z))} \lambda_d\Big)^2 \ge 0.$$

Moreover, if $(n, \Pi(\mathcal{P}, z)) = 1$ then

$$\sum_{d|n,\ d|\Pi(\mathcal{P},z)} \mu^+(d) = \mu^+(1) = \lambda_1^2 = 1.$$

Hence $\mu^+(d)$ satisfies (1.6).

We shall minimise the main term on the right-hand side of (1.8) in the class of functions given by (2.1). Let

$$S_0 := \sum_{d|\Pi(\mathcal{P},z)} \frac{\omega(d)\mu^+(d)}{d}\,, \quad \tilde{R} := \sum |\mu^+(d)|\,|R_d|\,, \quad \xi := \sqrt{y}\,.$$

If $\lambda_d = 0$ for $d \geq \xi$, then $\mu^+(d) = 0$ for $d \geq y$. Now,

$$S_0 = \sum_{[d_1,\,d_2]|\Pi(\mathcal{P},z)} \frac{\omega([d_1, d_2])}{[d_1, d_2]} \lambda_{d_1}\lambda_{d_2}\,.$$

Suppose that $\lambda_d = 0$ for $d \geq \xi$, then

$$S_0 = \sum_{\substack{d_1,d_2|\Pi \\ d_1,d_2<\xi}}{}^{*} \frac{\omega(d_1)\lambda_{d_1}}{d_1}\frac{\omega(d_2)\lambda_{d_2}}{d_2}\,\frac{(d_1,d_2)}{\omega((d_1,d_2))}\,,$$

where $\sum^*$ omits the terms for which $\omega(d_1)\,\omega(d_2) = 0$.

We now introduce the assumption

$$0 \leq \omega(p) < p \quad \text{for} \quad p \in \mathcal{P} \qquad (\Omega_1),$$

which we shall refer to in future merely as **Condition** (Ω_1). Then if $\mu(k)\omega(k) \neq 0$ we find that

$$\begin{aligned}
\sum_{l|k} \mu\Big(\frac{k}{l}\Big)\,\frac{l}{\omega(l)} &= \sum_{l'|k} \mu(l')\,\frac{k/l'}{\omega(k/l')} \\
&= \frac{k}{\omega(k)} \sum_{l'|k} \mu(l')\,\frac{\omega(l')}{l'} \\
&= \frac{k}{\omega(k)} \prod_{p|k} \Big(1 - \frac{\omega(p)}{p}\Big) \\
&> 0.
\end{aligned}$$

Thus we may define a non-negative multiplicative function $g(k)$ by

$$\frac{1}{g(k)} = \sum_{l|k} \mu\Big(\frac{k}{l}\Big)\,\frac{l}{\omega(l)}.$$

The Möbius inversion formula then shows that

$$\sum_{k|d} \frac{1}{g(k)} = \frac{d}{\omega(d)}$$

for $\mu(d)\omega(d) \neq 0$. Under condition (Ω_1), it now follows that

$$\begin{aligned} S_0 &= \sideset{}{^*}\sum_{\substack{d_1, d_2 | \Pi \\ d_1, d_2 < \xi}} \frac{\omega(d_1)\lambda_{d_1}}{d_1} \frac{\omega(d_2)\lambda_{d_2}}{d_2} \sum_{l|(d_1,d_2)} \frac{1}{g(l)} \\ &= \sum_{\substack{l|\Pi(\mathcal{P}, z) \\ l<\xi}} \frac{1}{g(l)} \Big(\sum_{\substack{d|\Pi(\mathcal{P}, z) \\ l|d,\, d<\xi}} \frac{\omega(d)\lambda_d}{d} \Big)^2 \\ &= \sum_{\substack{l|\Pi(\mathcal{P}\, z) \\ l<\xi}} \frac{1}{g(l)} \, y_l^2 \end{aligned}$$

with

$$y_l := \sum_{\substack{d|\Pi(\mathcal{P}, z) \\ l|d,\, d<\xi}} \frac{\omega(d)\lambda_d}{d} \; .$$

Lemma 2.1 *We have*

$$\sum_{\substack{l|\Pi,\, d|l \\ l<\xi}} \mu(l)\, y_l = \frac{\omega(d)\lambda_d}{d} \, \mu(d) \; , \tag{2.2}$$

if $d < \xi$ *and* $d|\Pi$.

Proof. When $d|\Pi$ we have

$$\begin{aligned}
\sum_{\substack{l|\Pi,\, l<\xi \\ d|l}} \mu(l) y_l &= \sum_{\substack{l|\Pi,\, l<\xi \\ d|l}} \ \sum_{\substack{\delta|\Pi,\, \delta<\xi \\ l|\delta}} \mu(l) \frac{\omega(\delta)\lambda_\delta}{\delta} \\
&= \sum_{\substack{\delta|\Pi,\, \delta<\xi \\ d|\delta}} \frac{\omega(\delta)\lambda_\delta}{\delta} \sum_{d|l,\ l|\delta} \mu(l) \\
&= \sum_{\substack{\delta|\Pi,\, \delta<\xi \\ d|\delta}} \frac{\omega(\delta)\lambda_\delta}{\delta} \sum_{md|\delta} \mu(md) \\
&= \mu(d) \sum_{\substack{\delta<\xi,\, \delta|\Pi \\ d|\delta}} \frac{\omega(\delta)\lambda_\delta}{\delta} \sum_{m|\frac{\delta}{d}} \mu(m) \\
&= \frac{\mu(d)\, \omega(d)\, \lambda_d}{d} ,
\end{aligned}$$

as claimed. Here we have used the fact that Π is square-free, so that m and d are coprime for $md|\Pi$.

Since $\lambda_1 = 1$, it follows from (2.2) that

$$1 = \sum_{l|\Pi(\mathcal{P}, z),\, l<\xi} \mu(l)\, y_l = \sum_{l|\Pi(\mathcal{P}, z),\, l<\xi} \mu(l) \sqrt{g(l)}\, \frac{y_l}{\sqrt{g(l)}}$$

and therefore

$$1 \le \Big\{ \sum_{l|\Pi(\mathcal{P}, z),\ l<\xi} \mu(l)^2 g(l) \Big\} \Big\{ \sum_{l|\Pi(\mathcal{P}, z),\ l<\xi} y_l^2\, g(l)^{-1} \Big\} = G(\xi, z) S_0$$

by Cauchy's inequality, where

$$G(\xi,\, z) := \sum_{l|\Pi,\, l<\xi} \mu(l)^2 g(l) = \sum_{l|\Pi(\mathcal{P}, z),\, l<\xi} g(l) \ .$$

Thus $S_0 \ge G(\xi,\, z)^{-1}$, and $S_0 = G(\xi,\, z)^{-1}$ if and only if there is a constant c such that

$$\frac{y_l}{\sqrt{g(l)}} = c\mu(l) \sqrt{g(l)}$$

for every l, this being the condition for equality in Cauchy's inequality. For the optimal values $y_l = c\mu(l)\, g(l)$ Lemma 2.1 yields

$$1 \;=\; \frac{\omega(1)\lambda_1}{1}\mu(1) \;=\; \sum_{l|\Pi,\, l<\xi} \mu(l) y_l \;=\; c \sum_{l|\Pi,\, l<\xi} \mu(l)^2 g(l) \;=\; c\, G(\xi,\, z),$$

whence

$$c = \frac{1}{G(\xi,\, z)}\ .$$

Thus $S_0 = G(\xi,\, z)^{-1}$ providing that $y_l \;=\; \mu(l)\, g(l) G(\xi,\, z)^{-1}$, in which case (2.2) produces

$$\begin{aligned}
\lambda_d \;&=\; \mu(d)\, \frac{d}{\omega(d)} \sum_{l|\Pi,\, d|l,\, l<\xi} \mu(l)\, y_l \\
&=\; \frac{\mu(d)\, d}{\omega(d)\, G(\xi,\, z)} \sum_{\substack{l|\Pi,\, d|l \\ l<\xi}} \mu(l)^2 g(l) \\
&=\; \frac{\mu(d)\, d\, g(d)}{\omega(d)\, G(\xi,\, d)} \sum_{dj|\Pi(\mathcal{P},\, z),\, j<\frac{\xi}{d}} g(j)\ .
\end{aligned}$$

On recalling that

$$g(d) \;=\; \frac{\omega(d)}{d} \prod_{p|d} \big(1 - \frac{\omega(p)}{p}\big)^{-1}\ ,$$

one obtains the minimising condition

$$\lambda_d \;=\; \mu(d) \prod_{p|d} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} \frac{G_d(\frac{\xi}{d},\, z)}{G(\xi,\, z)} \quad \text{for} \quad d < \xi\ , \tag{2.3}$$

with

$$G_d(\xi,\, z) := \sum_{dj|\Pi(\mathcal{P},z),\, j<\xi} g(j)\ , \quad G_1(\xi,\, z) = G(\xi,\, z)\ .$$

The choice of variables (2.3), under the assumption

$$\lambda_d = 0 \quad \text{for} \quad d \geq \xi\ , \tag{2.4}$$

turns (1.8) into the inequality

$$S(\mathcal{A}, \mathcal{P}; z) \leq X\, G(\xi,\, z)^{-1} \;+ \sum_{d|\Pi(\mathcal{P},z),\, d<y} |\mu^+(d)|\, |R_d|, \tag{2.5}$$

since $\mu^+(d) = 0$ for $d \geq y$, in view of (2.4).

In order to bound λ_d we will require the following result.

Lemma 2.2 *Let $d|\Pi(\mathcal{P},z)$. Then*

$$G(\xi,\,z) \geq G_d\big(\frac{\xi}{d},\,z\big) \prod_{p|d}\big(1-\frac{\omega(p)}{p}\big)^{-1}. \tag{2.6}$$

Proof. Let $d|\Pi(\mathcal{P},z)$. Then

$$\begin{aligned} G(\xi,\,z) &= \sum_{l|d} \sum_{\substack{m<\xi \\ m|\Pi,\,(m,\,d)=l}} g(m) \\ &= \sum_{\substack{l|d,\,lh|\Pi \\ lh<\xi,\,(h,\,d/l)=(h,\,l)=1}} g(lh) \\ &= \sum_{l|d} g(l) \sum_{\substack{h<\xi/l,\,lh|\Pi \\ (h,\,d/l)=(h,\,l)=1}} g(h) \\ &= \sum_{l|d} g(l) \sum_{h<\xi/l,\,dh|\Pi} g(h) \\ &\geq \sum_{l|d} g(l) \sum_{h<\xi/d,\,dh|\Pi} g(h) \\ &= \sum_{l|d} g(l)\, G_d\big(\frac{\xi}{d},\,z\big)\,, \end{aligned}$$

since $\xi/d \leq \xi/l$ and $g(h) \geq 0$. The inequality (2.6) follows now from the identity

$$\sum_{l|d} g(l) = \prod_{p|d}(1+g(p)) = \prod_{p|d}\big\{1+\frac{\omega(p)}{p}\big(1-\frac{\omega(p)}{p}\big)^{-1}\big\} = \prod_{p|d}\big(1-\frac{\omega(p)}{p}\big)^{-1},$$

in view of the condition (Ω_1).

The "Fundamental Theorem for Selberg's sieve". *Assume (Ω_1). Then*

$$S(\mathcal{A},\mathcal{P};z) \leq \frac{X}{G(\xi,\,z)} + \sum_{d<y,\,d|\Pi(\mathcal{P},\,z)} 3^{\nu(d)}\,|R_d|$$

for $\xi \geq 1$, where

$$\xi := y^{\frac{1}{2}}\,, \quad \nu(d) := \sum_{p|d,\,p\in\mathbb{P}} 1\,, \quad G(\xi,\,z) := \sum_{l|\Pi(\mathcal{P},z),\,l<\xi} g(l)$$

with

$$g(d) = \frac{\omega(d)}{d} \prod_{p|d} \big(1 - \frac{\omega(p)}{p}\big)^{-1} .$$

Proof. It follows from the relations (2.3) and (2.6) that $|\lambda_d| \leq 1$. Therefore

$$|\mu^+(d)| = \Big| \sum_{[d_1,\, d_2]=d} \lambda_{d_1}\lambda_{d_2} \Big| \leq \sum_{[d_1,\, d_2]=d} 1 = \sum_{\lambda=0}^{\nu(d)} \binom{\nu(d)}{\lambda} 2^\lambda = 3^{\nu(d)} ,$$

for square-free d. Consequently, the assertion of the theorem follows from (2.5).

3 Some applications

We prove two corollaries of the Fundamental Theorem.

Corollary 3.1 *Let* $\mathcal{P} = \{p \in \mathbb{P} : p \nmid k\}$ *and let*

$$\omega(p) = \begin{cases} 0, & p|k \\ 1, & p \nmid k \end{cases} .$$

Then

$$S(\mathcal{A}, \mathcal{P}; z) \leq \frac{k}{\varphi(k)} \frac{X}{\log z} + \sum_{\substack{d|\Pi(\mathcal{P},z),\, d<z^2 \\ (d,\, k)=1}} 3^{\nu(d)} \, |R_d| \, . \tag{3.1}$$

Proof. When $l|\Pi(\mathcal{P}, z)$ we have

$$g(l) = \frac{\omega(l)}{l} \prod_{p|l} \Big(1 - \frac{\omega(p)}{p}\Big)^{-1} = \frac{1}{\varphi(l)}$$

so that

$$G(z,\, z) = \sum_{l|\Pi(\mathcal{P},z),\, l<z} g(l) = \sum_{\substack{(l,\, k)=1 \\ l<z}} \frac{\mu(l)^2}{\varphi(l)} =: H_k(z) \, .$$

Let $K(n)$ be the largest square-free ivisor of n. One then obtains

$$H_1(z) = \sum_{l<z} \frac{\mu(l)^2}{l} \prod_{p|l} \Big(1 - \frac{1}{p}\Big)^{-1} = \sum_{l=p_1 \dots p_h < z} \prod_{i=1}^{h} (p_i - 1)^{-1} =$$

$$\sum_{\substack{\alpha_i\geq 1\\ p_1\ldots p_h<z}} \frac{1}{p_1^{\alpha_1}\ldots p_h^{\alpha_h}} = \sum_{K(n)<z}\frac{1}{n} \geq \sum_{n<z}\frac{1}{n} \geq \log z\,.$$

On the other hand, for any square-free k we have

$$\begin{aligned}
H_1(z) &= \sum_{n<z}\frac{\mu(n)^2}{\varphi(n)} \\
&= \sum_{l|k}\sum_{\substack{n<z\\ (n,k)=l}}\frac{\mu(n)^2}{\varphi(n)} \\
&= \sum_{\substack{l|k,\ hl<z\\ (h,k/l)=1}}\frac{\mu(lh)^2}{\varphi(lh)} \\
&= \sum_{l|k}\frac{\mu(l)^2}{\varphi(l)}\sum_{\substack{h<z/l\\ (h,k)=1}}\frac{\mu(h)^2}{\varphi(h)} \\
&= \sum_{l|k}\frac{\mu(l)^2}{\varphi(l)}\,H_k\Big(\frac{z}{l}\Big) \\
&\leq H_k(z)\sum_{l|k}\frac{\mu(l)^2}{\varphi(l)} \\
&= H_k(z)\prod_{p|k}\Big(1+\frac{1}{p-1}\Big) \\
&= \frac{k}{\varphi(k)}\,H_k(z)\,.
\end{aligned}$$

Combining the last two estimates, one obtains

$$H_k(z) \geq \frac{\varphi(k)}{k}\,H_1(z) \geq \frac{\varphi(k)}{k}\log z\,. \tag{3.2}$$

Thus

$$G(z,z) \geq \frac{\phi(k)}{k}\log z$$

so that Corollary 3.1 follows from the Fundamental Theorem with $\xi = z$.

Corollary 3.2 *Let* $\mathcal{P} = \{p \in \mathbb{P} : p \nmid k\}$ *and let*

$$\omega(p) = \begin{cases} 0, & p|k, \\ \frac{p}{p-1}, & p \nmid k. \end{cases}$$

Then

$$S(\mathcal{A}, \mathcal{P}; z) \le \Big(\prod_{\substack{p|k \\ p\neq 2}} \frac{p-1}{p-2} \Big) C_2 \frac{X}{\log z} \Big\{ 1 + O\big(\frac{1}{\log z}\big) \Big\} + \sum_{\substack{d<z^2,\, d|\Pi(\mathcal{P},z) \\ (d,\,k)=1}} 3^{\nu(d)} |R_d|$$

for even k*, where*

$$C_2 := 2 \prod_{p\ge 3} \Big(1 - \frac{1}{(p-1)^2} \Big)$$

is the "twin prime constant".

Proof. We have $g(p) = 0$ if $p|k$, and

$$g(p) = \frac{\omega(p)}{p} \Big(1 - \frac{\omega(p)}{p} \Big)^{-1} = \frac{1}{(p-1)\Big(1 - \frac{1}{p-1}\Big)} = \frac{1}{p-2} = \frac{1}{\varphi(p)} \,(1 + g(p))$$

otherwise. Therefore

$$g(d) = \frac{1}{\varphi(d)} \sum_{l|d} g(l)\, \mu(l)^2$$

as soon as $(d,\, k) = 1$ and $\mu(d)^2 = 1$. Moreover, $g(d) = 0$ when $(d,\, k) \ne 1$. Thus

$$\begin{aligned} G(z,\, z) &= \sum_{d|\Pi(\mathcal{P},z),\, d<z} g(d) \\ &= \sum_{d<z,\, (d,\,k)=1\ ,l|d} \frac{\mu(d)^2}{\varphi(d)} g(l)\, \mu(l)^2 \\ &= \sum_{\substack{(ml,\,k)=1 \\ ml<z}} \frac{\mu(ml)^2}{\varphi(ml)} g(l)\, \mu(l)^2 \\ &= \sum_{l<z,\, (l,\,k)=1} \frac{\mu(l)^2\, g(l)}{\varphi(l)} \sum_{(m,\, lk)=1,\, m<\frac{z}{l}} \frac{\mu(m)^2}{\varphi(m)} \\ &= \sum_{l<z,\, (l,\,k)=1} \frac{\mu(l)^2\, g(l)}{\varphi(l)} H_{lk}\Big(\frac{z}{l}\Big)\, . \end{aligned}$$

In view of (3.2),

$$\begin{aligned}
G(z,z) &\geq \sum_{l<z,\,(l,k)=1} \frac{\mu(l)^2\, g(l)}{\varphi(l)} \frac{\varphi(kl)}{kl} \log\big(\frac{z}{l}\big) \\
&= \frac{\varphi(k)}{k} \sum_{l<z,\,(l,k)=1} \frac{\mu(l)^2\, g(l)}{l} \log\big(\frac{z}{l}\big) \\
&\geq \frac{\varphi(k)}{k} \sum_{\substack{l=1 \\ (l,k)=1}}^{\infty} \frac{\mu(l)^2\, g(l)}{l} \log\big(\frac{z}{l}\big) \\
&= \frac{\varphi(k)}{k}(\log z) \prod_{p\nmid k} \big(1 + \frac{1}{p\,(p-2)}\big) - \frac{\varphi(k)}{k} \sum_{\substack{l=1 \\ (l,k)=1}}^{\infty} \frac{\mu(l)^2\, g(l)}{l} \log l \\
&= \frac{\varphi(k)}{k} \prod_{p\nmid k} \big(1 + \frac{1}{p\,(p-2)}\big) \{\log z + O(1)\},
\end{aligned}$$

so that

$$\frac{1}{G(z,z)} \leq \frac{1}{\log z}\Big(1 + O\big(\frac{1}{\log z}\big)\Big) \Big(2 \prod_{p|k} \frac{p-1}{p-2}\Big) \prod_{p\geq 3} \big(1 - \frac{1}{(p-1)^2}\big)$$

since

$$\frac{\varphi(p)}{p} \big(1 + \frac{1}{p\,(p-2)}\big)^{-1} = \frac{p-2}{p-1}$$

and

$$\big(1 + \frac{1}{p(p-2)}\big)^{-1} = \big(1 - \frac{1}{(p-1)^2}\big)$$

for $p \neq 2$. Corollary 3.2 now follows from the Fundamental Theorem.

Our applications will require one further result.

Lemma 3.1 *Let $h \in \mathbb{N}$ and set*

$$S_1 = \sum_{d<x} \mu(d)^2\, h^{\nu(d)}, \quad S_2 = \sum_{d<x} \frac{\mu(d)^2}{d}\, h^{\nu(d)}\,.$$

We then have

$$S_1 \leq x\,(1+\log x)^h \quad \textit{and} \quad S_2 \leq (1+\log x)^h\,.$$

Proof. Clearly,

$$S_1 \le \sum_{d<x} \mu(d)^2 \, \frac{x}{d} \, h^{\nu(d)} \;=\; x\, S_2 \, .$$

Moreover,

$$\begin{aligned} S_2 \;&=\; \sum_{d<x} \frac{\mu(d)^2}{d} \sum_{\substack{(d_1,\dots,d_h) \\ d_1 \dots d_h = d}} 1 \\ &\le\; \sum_{d \ge 1} \frac{1}{d} \sum_{\substack{d_1 \dots d_h = d \\ d_1,\dots,d_h < x}} \mu(d_1)^2 \dots \mu(d_h)^2 \\ &=\; \Big(\sum_{d<x} \frac{\mu(d)^2}{d} \Big)^h \le\; (1 + \log x)^h , \end{aligned}$$

as asserted. Here we have used the fact that if $\mu(d)^2 = 1$ and $d = p_1 \dots p_h$, then

$$h^\nu = \#\{(d_1, \dots , d_h) \in \mathbb{N}^h : \, d_1 \dots d_h = d\}.$$

First application. Let $\mathcal{A} = \{\, n \,|\, x < n \le\, x + y \,\}$ and $\mathcal{P} = \mathbb{P}$, so that

$$\#\mathcal{A}_d = \frac{y}{d} \,+\, O(1), \quad \omega(d) = 1 \quad \text{for} \quad d \in \mathbb{N} \, .$$

Then

$$S(\mathcal{A}, \mathcal{P}; z) \;\le\; \frac{y}{\log z} \;+\; O\big(\sum_{d \le z^2} 3^{\nu(d)} \, \mu(d)^2 \big)$$

by Corollary 3.1. In view of Lemma 3.1, this gives

$$S(\mathcal{A}, \mathcal{P}; z) \;\le\; \frac{y}{\log z} \;+\; O\big(z^2 (1 + \log z)^3 \big) \, .$$

On the other hand,

$$S(\mathcal{A}, \mathcal{P}; z) \;=\#\{\, n \,:\, x < n \le\, x{+}y, \; p|n \;\Rightarrow\; p \ge z \,\} \;\ge\; \pi(x{+}y) - \pi(x) - z \, .$$

Taking $z = \frac{\sqrt{y}}{(\log y)^3}$, it follows that

$$\pi(x + y) - \pi(x) \le \frac{2y}{\log y} \;+\; O\Big(\frac{y \, \log \log y}{(\log y)^2} \Big) \quad \text{for } x, y \ge 2 \, . \tag{3.3}$$

In particular,

$$\pi(y) \le \frac{2y}{\log y} \;+\; O\Big(\frac{y \, \log \log y}{(\log y)^2} \Big) \, .$$

It has been proved (Heath-Brown [9]) that

$$\pi(x+y)-\pi(x) \sim \frac{y}{\log x} \tag{3.4}$$

for $y \geq x^{\frac{7}{12}}$. In contrast, (3.4) is false (Maier [14]) for $y \asymp (\log x)^A$, for any constant A. Thus (3.3) is useful for relatively small y. Montgomery and Vaughan [15] have removed the error term above and proved that

$$\pi(x+y)-\pi(x) \leq \frac{2y}{\log y} \quad \text{for } \ x\,,\, y \geq 2\,.$$

It has been conjectured that $\pi(x+y) \leq \pi(x)+\pi(y)$ for all $x, y \geq 2$. However Hensley and Richards [10] have proved that this would be incompatible with the k-tuples prime conjecture. It is not clear at the moment whether the factor 2 in (3.3) may be replaced by a smaller number. Indeed Erdős apparently believed that the constant may be taken as 1, while Selberg is reputed to have suggested that no constant below 2 is admissible.

Second application. Suppose that $(l,\, k)=1$ and let

$$\mathcal{A}=\{n \leq x:\ n \equiv l \pmod{k}\},$$

$$\pi(x;k,l) := \#\{\, p \in \mathbb{P}:\ p \leq x,\ p \equiv l \pmod{k}\,\},$$

$$\mathcal{P}=\{\, p\in \mathbb{P}:\ p \nmid k\,\}\,.$$

Clearly,

$$\pi(x;k,l) \leq\ S(\mathcal{A},\mathcal{P};z)+1+\frac{z}{k}\,. \tag{3.5}$$

Moreover,

$$\#\mathcal{A}_d=\frac{x}{k}\frac{\omega(d)}{d}\ +O(1)$$

with

$$\omega(d)=\begin{cases} 1, & \text{if } \ (d,k)=1, \\ 0, & \text{if } \ (d,k)\neq 1, \end{cases}$$

so that

$$\begin{aligned} S(\mathcal{A},\mathcal{P};z) \ &\leq\ \frac{k}{\varphi(k)}\,\frac{x}{k}\,\frac{1}{\log z}\ +\sum_{d\leq z^2} 3^{\nu(d)}\,\mu(d)^2 \\ &=\ \frac{x}{\varphi(k)\log z}\ +O\big(z^2(\log z)^3\big) \end{aligned}$$

by Corollary 3.1 and Lemma 3.1. Let $z=(x/k)^{1/2}(\log x/k)^{-3}$. Assuming that $x \geq 4k$, say, the estimate (3.5) yields the following result.

The Brun-Titchmarsh Theorem. *We have*

$$\pi(x;k,l) \le \frac{2x}{\varphi(k)\log x/k} + O\Big(\frac{x}{\varphi(k)}\frac{\log\log x/k}{(\log x/k)^2}\Big). \tag{3.6}$$

for $x \ge 4\varphi(k)$.

The Siegel-Walfisz Theorem gives

$$\pi(x;k,l) \sim \frac{x}{\varphi(k)\log x} \quad \text{for} \quad k \ll (\log x)^A$$

for any fixed A, so that the constant 2 in the Brun-Titchmarsh Theorem may be replaced by 1 if x is sufficiently large compared with x. Moreover Montgomery and Vaughan [15] have proved that

$$\pi(x;k,l) \le \frac{2x}{\varphi(k)\log x/k} \quad \text{for} \quad k < x.$$

The constant 2 in (3.6) is presumably hard to improve, for it is known that if one could replace 2 by $2-\delta$ with a positive constant δ, then it would follow that there are no "Siegel-Landau zeros".

Third application. Let

$$\mathcal{A} = \{\, 2N - p \,:\, p \in \mathcal{P},\ 3 \le p \le 2N-3 \,\}, \quad \mathcal{P} = \{\, p \in \mathbb{P} :\ (p, 2N) = 1 \,\},$$

so that

$$\begin{aligned} \mathcal{A}_d &= \#\{\, p \in \mathbb{P} :\ p \equiv 2N \pmod d,\ p \nmid 2N,\ p \le 2N-3 \,\} \\ &= \pi(2N; d, 2N) + O\big(1 + \nu(N)\big). \end{aligned}$$

When h.c.f.$(d, 2N) = 1$ we expect that

$$\pi(N; d, 2N) \sim \frac{\mathrm{Li}(2N)}{\varphi(d)}$$

for N large compared with d. We therefore take $X = \mathrm{Li}(2N)$ and $\omega(d) = \varphi(d)^{-1}d$. On writing

$$E(x;k,l) = \pi(x;k,l) - \frac{\mathrm{Li}(x)}{\varphi(k)}$$

it follows that

$$R_d = E(2N; d, 2N) + O\big(1 + \nu(N)\big).$$

On the other hand

$$r(2N) := \#\{\, (p,q) \,|\, p,q \in \mathbb{P},\ p+q = 2N \,\} \le S(\mathcal{A}, \mathcal{P}; z) + z$$

and we deduce from Corollary 3.2 that

$$r(2N) \le \Big(\prod_{\substack{p|2N \\ p\neq 2}} \frac{p-1}{p-2}\Big)\, C_2\, \frac{\mathrm{Li}(2N)}{\log z}\,\Big(1+O\big(\frac{1}{\log z}\big)\Big) \;+\; z$$

$$+\,O\big(\sum_{d\le z^2} 3^{\nu(d)}\,\mu(d)^2\,\{|E(2N;d,2N)|\;+\;1+\nu(N)\}\big)\,.$$

In order to estimate the remainder sum

$$S := \sum_{d\le z^2} 3^{\nu(d)}\mu(d)^2 E_1(2N,d),$$

where

$$E_1(x,k) := \max_{y\le x,\,(l,k)=1} |E(y;k,l)|\,,$$

we shall use the following well-known result.

The Bombieri-Vinogradov Theorem. *For every $c_1 > 0$, there is a positive constant c_2 such that*

$$\sum_{k\le x^{\frac{1}{2}}\,(\log x)^{-c_2}} E_1(x,k) \ll \frac{x}{(\log x)^{c_1}}.$$

Using Cauchy's inequality with Lemma 3.1 we find that

$$\begin{aligned} S^2 &\le \big\{\sum_{d<z^2}(3^{\nu(d)}\mu(d)^2)^2\,\frac{1}{d}\big\}\,\big\{\sum_{d\le z^2} d\,E_1(2N,d)^2\big\} \\ &\ll (\log z)^9 \sum_{d<z^2} d\,E_1(2N,d)^2. \end{aligned}$$

It is trivial that

$$\pi(x;k,l) \ll \frac{x}{k} \ll \frac{x}{\varphi(k)}$$

for $k\le x$, so that

$$E(x;k,l) \ll \pi(x;k,l) + \frac{x}{\varphi(k)} \ll \frac{x}{\varphi(k)} \ll \frac{x\log x}{k}.$$

Thus

$$\sum_{d<z^2} d\,E_1(2N,d)^2 \ll N(\log N)\sum_{d<z^2} E_1(2N,d) \ll N^2(\log N)^{-15},$$

by the Bombieri-Vinogradov Theorem with $c_1 = 16$, on taking

$$z^2 = \sqrt{N}\,(\log N)^{-c_2}.$$

We therefore conclude that

$$S^2 \ll \frac{N^2}{(\log N)^6}.$$

Thus

$$\sum_{d \le z^2} 3^{\nu(d)} \mu(d)^2 E_1(2N, d) = O(N\,(\log N)^{-3})\,.$$

Moreover, it follows from Lemma 3.1 that

$$\big(1+\nu(N)\big) \sum_{d \le z^2} 3^{\nu(d)} \mu(d)^2 \ll \big(1+\nu(N)\big)\sqrt{N}(\log N)^3 \ll \frac{N}{(\log N)^3}\,.$$

We may now deduce the following result from (3.1).

Theorem 3.1 *We have*

$$r(2N) \le \{4 + O(\frac{\log\log N}{\log N})\} a(N),$$

with

$$a(N) = \Big(\prod_{\substack{p|2N \\ p \ne 2}} \frac{p-1}{p-2}\Big)\, C_2\, \frac{2N}{(\log N)^2}.$$

It is conjectured that one may improve the Bombieri-Vinogradov Theorem to say that for any $\varepsilon > 0$ and any $c_1 > 0$ one has

$$\sum_{k \le x^{1-\varepsilon}} E_1(x, k) \ll \frac{x}{(\log x)^{c_1}}\,.$$

One would then obtain a bound

$$r(2N) \le (2 + o(1)) a(N)$$

in a completely analogous fashion. However the best unconditional result is due to Chen [4], in which the constant 4 is reduced to 3.9171. For comparison we note that it is conjectured that $r(2N) \sim a(N)$.

The following theorem can be proved in the same way as Theorem 3.1 (Exercise!).

Theorem 3.2 *For any positive integer k we have*

$$\#\{\, p \le x \,:\, p, p+2k \in \mathbb{P}\} \le 4\Big(\prod_{p|2k} \frac{p-1}{p-2}\Big)\, C_2\, \frac{x}{(\log x)^2}\, \Big(1 + O\Big(\frac{\log\log x}{\log x}\Big)\Big)\,.$$

Corollary 3.3 *(Viggo Brun)*
We have

$$\sum_{p,\, p+2 \in \mathbb{P}} \frac{1}{p} \,<\, \infty.$$

Bombieri, Friedlander and Iwaniec [2] have proved a variant of Theorem 3.2 with the constant 4 replaced by 7/2. Their method does not establish a result uniform in k and is therefore not applicable to Theorem 3.1. More complicated methods allow one to reduce the constant in Theorem 3.2 further slightly.

We proceed to discuss briefly some other applications of Selberg' sieve.
Definition. *Suppose that*

$$\sum_{w \le p < z} \frac{\omega(p)\, \log p}{p} \;=\; \kappa\, \log\Big(\frac{z}{w}\Big) \,+\, O(1) \quad \textit{for} \quad 2 \le w \le z. \tag{3.7}$$

Then the constant κ is called the **dimension** *of the sieve problem.*

One can get by with a slightly weaker assumption in fact. The above definition corresponds to a version of the condition $\Omega(\kappa, L)$ in the book by Halberstam and Richert [6, page 142].

Remark. Since

$$\sum_{w \le p < z} \frac{\log p}{p} \;=\; \log\Big(\frac{z}{w}\Big) \,+\, O(1)\,,$$

the dimension of the sieve problem coincides with the "average value" of $\omega(p)$. Note that in the two cases considered in this section we have $\kappa = 1$. This is clear for both

$$\omega(p) \;=\; \begin{cases} 1, & p \nmid k, \\ 0, & p|k, \end{cases}$$

and

$$\omega(p) \;=\; \begin{cases} 1 + \frac{1}{p-1}, & p \nmid k, \\ 0, & p|k, \end{cases}$$

In general, for the sieve problem of dimension κ, one obtains

$$\begin{aligned} G(z,z) &= \sum_{\substack{d\mid\Pi(\mathcal{P},z)\\ d\le z}} \frac{\omega(d)}{d} \prod_{p\mid d} \big(1-\frac{\omega(p)}{p}\big)^{-1} \\ &= \sum_{d\le z} \frac{\omega(d)\mu(d)^2}{d} \prod_{p\mid d} \big(1-\frac{\omega(p)}{p}\big)^{-1}. \end{aligned}$$

Though we shall not prove it, it turns out that

$$G(z,z) = \frac{1}{e^{\gamma\kappa}\,\Gamma(\kappa+1)} \prod_{p<z} \big(1-\frac{\omega(p)}{p}\big)^{-1} \big(1+O\big(\frac{1}{\log z}\big)\big)$$

see Halberstam and Richert [6, (5.3.1)]. It therefore follows from the Fundamental Theorem that

$$\begin{aligned} S(\mathcal{A},\mathcal{P};z) \;\le\; & X \prod_{p<z} \big(1-\frac{\omega(p)}{p}\big)\, e^{\gamma\kappa}\,\Gamma(\kappa+1)\, \big(1+O\big(\frac{1}{\log z}\big)\big) \\ & + \sum_{d<z^2,\, d\mid\Pi(\mathcal{P},z)} 3^{\nu(d)}|R_d|. \end{aligned} \tag{3.8}$$

Note that if $\omega(p)=1$, then

$$\prod_{p<z} \big(1-\frac{1}{p}\big) \sim \frac{e^{-\gamma}}{\log z}.$$

One therefore obtains

$$S(\mathcal{A},\mathbb{P};z) \le \frac{X}{\log z}\, \big(1+O\big(\frac{1}{\log z}\big)\big) \;+ \sum_{d<z^2,\, d\mid\Pi(\mathbb{P},z)} 3^{\nu(d)}|R_d|\,,$$

cf. (3.1).

Example. Let

$$\mathcal{A} = \{\, n(2N-n) \,:\, 3\le n\le 2N-3 \,\},$$

so that

$$\begin{aligned} \#\mathcal{A}_d &= \#\{\, n \mid 1\le n\le d,\ d\mid n(2N-n)\,\}\, \big(\frac{2N}{d}+O(1)\big) \\ &= \frac{\omega(d)}{d}\, 2N + O(\omega(d))\,, \end{aligned}$$

where

$$\omega(p) = \begin{cases} 1, & p \mid 2N, \\ 2, & p \nmid 2N. \end{cases}$$

In this case $\kappa = 2$ and the estimate (3.8) implies $r(2N) \le (8 + o(1))a(N)$ (Exercise!). This should be compared with Theorem 3.1, which was deduced using a sieve of dimension 1, together with the Bombieri-Vinogradov Theorem.

4 The parity phenomenon and limitations to sieve methods

The optimisation problem for the upper bound sieve amounts to the question of minimising the linear functional

$$l(\mu^*) := X \sum_{d \mid \Pi(\mathcal{P},z)} \frac{\mu^*(d)\,\omega(d)}{d} + \sum_{d \mid \Pi(\mathcal{P},z)} \mu^*(d) R_d$$

under the additional condition that

$$\sum_{d \mid n,\, d \mid \Pi(\mathcal{P},z)} \mu^*(d) \ge \begin{cases} 1, & (n, \Pi(\mathcal{P},z)) = 1, \\ 0, & \text{otherwise}, \end{cases}$$

(cf. (1.5)–(1.7)). We can view this as a linear programming problem. In the standard formulation of a linear programing problem one takes a real $m \times n$ matrix A and real column vectors $\mathbf{b}$ and $\mathbf{c}$ of lengths m and n respectively. One then seeks to minimize $\mathbf{c}^t\mathbf{x}$ over all column vectors $\mathbf{x} \in \mathbb{R}^n$, subject to the conditions that $\mathbf{x} \ge \mathbf{0}$ and $A\mathbf{x} \ge \mathbf{b}$. (Here $\mathbf{z} \ge \mathbf{w}$ means that $z_i \ge w_i$ for each index i.) In our problem the vector of values of $\mu^*(d)$ is not required to be non-negative, so we introduce two new functions $\mu^*_\pm(d)$ with $\mu^*_\pm(d) \ge 0$ for all d, and write $\mu^*(d) = \mu^*_+(d) - \mu^*_-(d)$. We can then produce a linear programming problem in standard form.

A great deal can be learnt about a linear programing problem by studying its "dual". For the problem described above, the dual problem is to maximize $\mathbf{y}^t\mathbf{b}$ over column vectors $\mathbf{y} \in \mathbb{R}^m$, subject to $\mathbf{y} \ge \mathbf{0}$ and $\mathbf{y}^t A \le \mathbf{c}^t$. Under these constraints one clearly has

$$\mathbf{c}^t\mathbf{x} \ge \mathbf{y}^t A\mathbf{x} \ge \mathbf{y}^t\mathbf{b} \tag{4.1}$$

and the Duality Theorem states that there exist vectors $\mathbf{x}, \mathbf{y}$ for which equality is attained.

In the context of the upper bound sieve problem, it transpires that the dual problem is essentially that of finding a sequence $\mathcal{A}$, with prescribed function $\omega(d)$, and with suitably small remainders R_d, for which $S(\mathcal{A},\mathcal{P};z)$ is as large as possible. We can interpret the inequalities (4.1) as saying that, for any vector $\mathbf{y}=\mathbf{y}_0$ which satisfies the relevant constraints, we must have

$$\inf \mathbf{c}^t\mathbf{x} \geq \mathbf{y}_0^t\mathbf{b},$$

and moreover, if we have vectors $\mathbf{x}$ and $\mathbf{y}$ which both satisfy the relevant constraints, and for which $\mathbf{c}^t\mathbf{x}=\mathbf{y}^t\mathbf{b}$, then both $\mathbf{x}$ and $\mathbf{y}$ must be extremal.

For the upper bound sieve problem, any sequence $\mathcal{A}$ defining a problem of dimension 1 will therefore produce a lower bound on the possible values of $l(\mu^*)$. Moreover if we can find a sequence $\mathcal{A}$ and a set of coefficients $\mu^*(d)$ for which $S(\mathcal{A},\mathcal{P};z)$ (corresponding to $\mathbf{y}^t\mathbf{b}$) and $l(\mu^*)$ (corresponding to $\mathbf{c}^t\mathbf{x}$) are approximately equal, then both must be essentially optimal.

We therefore examine in detail the following two sequences, first discussed by Selberg. Let $\Omega(n)$ be the number of prime factors of n, counted according to multiplicity, and define the *Liouville function* $\lambda(n)$ by

$$\lambda(n) = (-1)^{\Omega(n)}.$$

We then set

$$\mathcal{A}^+ = \{n \in \mathbb{N} :\, n \leq x,\ \lambda(n) = -1\,\},$$

and

$$\mathcal{A}^- = \{n \in \mathbb{N} :\, n \leq x,\ \lambda(n) = +1\,\},$$

which will relate to the upper bound and lower bound problems respectively. In the case of the first sequence we have

$$S(\mathcal{A}^+,\mathbb{P};z) = \pi(x)-\pi(z) = \frac{x}{\log x} + O\Big(\frac{x}{(\log x)^2}\Big) + O(z) \quad \text{for } z > x^{\frac{1}{3}}\,.$$

Let us see how this compares with the bound given by the Selberg sieve. To bound R_d for the sequences $\mathcal{A}^\pm$ we note that if

$$L(x) := \sum_{n\leq x} \lambda(n),$$

then $L(x) \ll E_c(x)$, where c is a suitable positive constant and

$$E_c(x) := \exp(-c\sqrt{\log x}).$$

(This follows by a similar analysis to that used for the summatory function of $\mu(n)$.) Now, if we let $X = \frac{x}{2}$ then

$$\begin{aligned}
\#\mathcal{A}_d^{\pm} &= \#\{m \le \frac{x}{d} :\ \lambda(md) = \mp 1\,\} \\
&= \#\{m \le \frac{x}{d} :\ \lambda(m) = \mp\lambda(d)\,\} \\
&= \frac{[x/d]}{2} \mp \lambda(d)\frac{L(x/d)}{2} \\
&= \frac{X}{d} + O\big(E_c\big(\frac{x}{d}\big)\big)\,,
\end{aligned}$$

for $d \le x$, and hence the remainder sum in the Fundamental Theorem is

$$\begin{aligned}
\sum_{d<y,\, d|\Pi(\mathcal{P},z)} 3^{\nu(d)}\,|R_d| &\ll \sum_{d<y,} 3^{\nu(d)}\,E_c\big(\frac{x}{d}\big) \\
&\le \Big(\sum_{d<y,} 9^{\nu(d)} d^{-1}\Big)^{\frac{1}{2}} \Big(\sum_{d<y} d\,E_c\big(\frac{x}{d}\big)^2\Big)^{\frac{1}{2}} \\
&\ll (\log y)^{\frac{9}{2}} \Big(\sum_{d<y} d\,E_c\big(\frac{x}{d}\big)^2\Big)^{\frac{1}{2}}
\end{aligned}$$

in view of Lemma 3.3. Furthermore,

$$\begin{aligned}
\sum_{d<y} dE_c\big(\frac{x}{d}\big)^2 &= x^2 \sum_{d<y} d^{-1}\exp(-2c\sqrt{\log x/d}) \\
&\le x^2 \sum_{d<y} d^{-1}\exp(-2c\sqrt{\log x/y}) \\
&\le x^2(\log y)\exp(-2c\sqrt{\log x/y})
\end{aligned}$$

for $y < x$. Thus

$$\sum_{d<y} 3^{\nu(d)}\,|R_d| \ll (\log x)^5\,x\,e^{-c\sqrt{\log x/y}} \ll x\,(\log x)^{-2} \tag{4.2}$$

for $y \le E_1(x)$, say. By Corollary 3.1, we have

$$S(\mathcal{A}^+, \mathbb{P}; z) \le \frac{X}{\log z} + \sum_{d<z^2} 3^{\nu(d)}|R_d|.$$

On taking $y = E_1(x)$ and $z = y^{1/2}$ one deduces from these estimates on recalling that $x = 2X$, that

$$S(\mathcal{A}^+, \mathbb{P}; z) \le \frac{X}{\log(\sqrt{E_1(x)})} + O\big(\frac{x}{(\log x)^2}\big) = \frac{x}{\log x} + O\big(\frac{x}{(\log x)^{3/2}}\big)\,.$$

Since

$$S(\mathcal{A}^+, \mathbb{P}; z) = \pi(x) - \pi(z) = \frac{x}{\log x} + O\Big(\frac{x}{(\log x)^2}\Big) + O(z) \quad \text{for } z > x^{1/3},$$

we conclude that the estimate (3.1) cannot be improved on for $z > x^{1/3}$.

Thus the Selberg sieve is best possible in this situation. Our remarks about linear programming then show that the Selberg sieve coefficients are an essentially optimal solution to the minimiztion problem for $l(\mu^*)$, and that the sequence $\mathcal{A}^+$ is a corresponding solution for the dual problem.

Turning to the lower bound sieve problem, we see that we can satisfy the relevant constraints

$$\sum_{d|n,\, d|\Pi(\mathcal{P},z)} \mu^*(d) \le \begin{cases} 1, & \text{if} \quad (n, \Pi(\mathcal{P}, z)) = 1, \\ 0, & \quad \text{otherwise,} \end{cases}$$

by taking $\mu^*(d)$ to be identically zero. For this choice we produce the trivial lower bound

$$S(\mathcal{A}, \mathcal{P}; z) \ge X \sum_{d|\Pi(\mathcal{P},z)} \frac{\mu^*(d)\omega(d)}{d} - \sum_{d|\Pi(\mathcal{P},z)} |\mu^*(d)|\, |R_d| = 0.$$

We now observe that for the sequence $\mathcal{A}^-$ we have $S(\mathcal{A}^-, \mathbb{P}, z) = 1$ for $z > x^{1/2}$, since only the integer $1 \in \mathcal{A}^-$ is counted. Thus the coefficients $\mu^*(d) = 0$ are essentially best possible for the linear programming problem in this situation, and the sequence $\mathcal{A}^-$ is essentially optimal for the corresponding dual problem.

Thus no set of lower bound sieve coefficients $\mu^*(d)$ with $|\mu^*(d)| \le 3^{\nu(d)}$ can produce

$$\sum_{d|\Pi(\mathbb{P},z),\, d<y} \frac{\mu^*(d)}{d} \gg \frac{1}{\log y}$$

when $z > x^{1/2}$. In particular one cannot show that $\pi(x) \gg x/\log x$ by sieve methods alone.

The two sequences $\mathcal{A}^+$ and $\mathcal{A}^-$ produce the same information input for the sieve. They have the same X, the same function $\omega(d)$, and their remainders R_d have the same order of magnitude. Thus there is no way that the sieve machinery can distinguish them. It is for this reason that the sieve encounters the *parity phenomenon*, since it is unable to distinguish integers for which $\Omega(n)$ is even, from those for which $\omega(n)$ is odd.

The sequences $\mathcal{A}^{\pm}$ have been shown to be essentially extremal for $z > x^{1/2}$, but it transpires that they are optimal for all z. To examine this fact we define

$$\begin{aligned} S^{\pm}(x,s): &= S\big(\mathcal{A}^{\pm},\mathbb{P};x^{1/s}\big) \\ &= \#\{n\in\mathbb{N}:\ n\le x,\ \lambda(n)=\mp 1,\ p|n\ \Rightarrow\ p\ge x^{1/s}\,\} \end{aligned}$$

for $s \geq 1$. We classify the integers n according to their smallest prime factor p. Then if

$$\eta_+ = 0, \quad \eta_- = 1$$

it follows that

$$\begin{aligned} S^{\pm}(x,s) &= \sum_{p\ge x^{1/s}} \#\{m:\ pm\le x,\ \lambda(m)=\pm 1,\ p'|m\ \Rightarrow\ p'\ge p\,\}\ +\ \eta_{\pm} \\ &= \sum_{x^{1/s}\le p\le x} S^{\mp}\big(\frac{x}{p},\,\frac{\log x}{\log p}-1\big)\ +\ \eta_{\pm}, \end{aligned}$$

since

$$p = \left(\frac{x}{p}\right)^{1/s'}$$

with

$$s' = \frac{\log x/p}{\log p}.$$

This leads to recursion formulae for $S^{\pm}(x,s)$. To produce appropriate starting values for the recursions we note that

$$S^{-}(x,s) = O(1), \quad (1\le s\le 2)$$

and

$$S^{+}(x,s) = \pi(x) - \pi(x^{1/s}) + O(1) \quad (1\le s\le 3).$$

Let us define continuous functions $F, f\,:\,\mathbb{R}_{>0}\to\mathbb{R}$ by the relations

$$F(s) = \frac{2\,e^{\gamma}}{s} \quad \text{for } 0<s\le 3, \qquad f(s)=0 \quad \text{for } 0<s\le 2,$$

and

$$(s\,F(s))' = f(s-1), \quad (s\,f(s))' = F(s-1) \quad \text{for } s>2\,.$$

We note that these definitions show that

$$f(s) = 2e^{\gamma}\frac{\log(s-1)}{s} \quad \text{for }\ 2\le s\le 4.$$

We shall now prove the following estimates.

Theorem 4.1 *Let $N \in \mathbb{N}$. Then we have*

$$S^+(x,s) = \frac{x/2}{e^\gamma \log(x^{1/s})} F(s) + O_N\Big(\frac{x}{(s-1)(\log x)^2}\Big) \tag{4.3}$$

and

$$S^-(x,s) = \frac{x/2}{e^\gamma \log(x^{1/s})} f(s) + O_N\Big(\frac{x}{(\log x)^2}\Big) \tag{4.4}$$

for $1 \le s \le N$.

Proof. Since

$$\frac{x/2}{e^\gamma \log(x^{1/s})} F(s) = \frac{x}{\log x} \quad (0 < s \le 3)$$

and

$$S^+(x,s) = \pi(x) - \pi(x^{1/s}) + O(1) \quad (0 < s \le 3)\,,$$

relation (4.3) holds for $N = 2$ and $N = 3$. Similarly, when $N = 2$ the equation (4.4) follows from the facts that $f(s) = 0$ and $S^-(x,s) = O(1)$ whenever $1 \le s \le 2$. We now prove (4.3) and (4.4) by induction on N. We shall consider only $S^+(x,s)$, leaving the discussion of $S^-(x,s)$ as an exercise. Thus we assume that (4.3) and (4.4) hold for N, and deduce that (4.3) holds for $N+1$. We therefore let $N < s \le N+1$, with $N \ge 3$. Since

$$0 < \frac{\log x}{\log p} - 1 \le 2 \quad \text{for} \quad p \ge x^{1/s}\,,$$

it follows that

$$\begin{aligned} S^+(x,s) &= \sum_{x^{1/s} \le p \le x} S^-\Big(\frac{x}{p}, \frac{\log x}{\log p} - 1\Big) \\ &= \sum_{x^{1/3} < p \le x} S^-\Big(\frac{x}{p}, \frac{\log x}{p} - 1\Big) + \sum_{x^{1/s} \le p \le x^{1/3}} S^-\Big(\frac{x}{p}, \frac{\log x}{\log p} - 1\Big) \\ &= \pi(x) - \pi(x^{1/3}) + O(1) + \sum_{x^{1/s} \le p < x^{1/3}} S^-\Big(\frac{x}{p}, \frac{\log x}{\log p} - 1\Big). \end{aligned}$$

Moreover, if $p \ge x^{1/s}$ then $\frac{\log x}{\log p} - 1 \le s - 1 \le N$. Therefore, by the inductive assumption, we have

$$S^-\Big(\frac{x}{p}, \frac{\log x}{\log p} - 1\Big) = \frac{(2e^\gamma)^{-1} x}{p \log p} f\Big(\frac{\log x}{\log p} - 1\Big) + O_N\Big(\frac{x}{p\,(\log x)^2}\Big) \tag{4.5}$$

for $x^{1/s} \le p \le x^{1/3}$. By partial summation we find that

$$\begin{aligned}\sum_{x^{1/s}\le p\le x^{1/3}} \frac{1}{p\,\log p} f\big(\frac{\log x}{\log p}-1\big) &= \int_{x^{1/s}}^{x^{1/3}} \frac{1}{t\log t} f\big(\frac{\log x}{\log t}-1\big)\frac{dt}{\log t} \\ &\qquad + O_N\big(\frac{x}{(\log x)^2}\big) \\ &= \frac{1}{\log x}\int_3^s f(v-1)\,dv + O_N\big(\frac{x}{(\log x)^2}\big) \\ &= \frac{1}{\log x}(sF(s)-3F(3)) + O_N\big(\frac{x}{(\log x)^2}\big),\end{aligned}$$

on substituting $t = x^{1/v}$. To handle the error term in (4.5) we note that

$$\sum_{x^{1/s}\le p\le x^{1/3}} \frac{x}{p(\log x)^2} \ll_N \frac{x}{(\log x)^2},$$

whence we conclude that

$$\begin{aligned}S^+(x,s) &= \frac{x}{2\,e^{\gamma}\log x}\,(sF(s)-3F(3)) + \frac{x}{\log x} + O_N\big(\frac{x}{(\log x)^2}\big) \\ &= \frac{x}{2\,e^{\gamma}\log x^{1/s}}F(s) + O_N\big(\frac{x}{(\log x)^2}\big),\end{aligned}$$

as required.

Remarks.

1) The properties of the functions f, F and their generalisations are discussed in detail in the books by Greaves [5] and Halberstam and Richert [6, Chapter 8].

2) If we set

$$W(z) := \prod_{p\le z}\big(1-\frac{1}{p}\big),$$

then the Mertens formula (1.4) gives

$$\frac{x/2}{e^{\gamma}\log(x^{1/s})} = XW(x^{1/s}) + O_N\big(\frac{x}{(\log x)^2}\big),$$

so that (4.3) and (4.4) imply that

$$S^+(x,s) = X\,W(x^{1/s})F(s) \;+\; O_N\big(\frac{x}{(s-1)(\log x)^2}\big)$$

and

$$S^-(x,s) = X\,W(x^{1/s})f(s) \;+\; O_N\big(\frac{x}{(\log x)^2}\big)$$

respectively.

5 The Rosser sieve

A *combinatorial sieve* is defined by choosing sets

$$T^+(y), T^-(y) \subseteq \{d \in \mathbb{N} : d < y,\ \mu(d)^2 = 1\}$$

and taking

$$\mu^{\pm}(d) = \begin{cases} \mu(d), & \text{if } d \in T^{\pm}(y), \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

The sets $T^{\pm}(y)$ have to be chosen so that $\mu^{\pm}(d)$ satisfy (1.5), and (1.6) respectively. As in the proof of (1.7) we have

$$S(\mathcal{A}, \mathcal{P}; z) \leq X \sum_{d|\Pi(\mathcal{P},z)} \frac{\mu^+(d)\omega(d)}{d} + \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)\, R_d\,. \tag{5.2}$$

It follows from (5.1) and (5.2) that

$$\Big| \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)\, R_d \Big| \leq \sum_{d\leq y,\, d|\Pi(\mathcal{P},z)} |R_d|\,.$$

We have to choose $T^+(y)$ so as to optimise the main term in (5.2) subject to the condition (1.6). It follows from (1.8) and its analogue for μ^- that

$$S(\mathcal{A}, \mathcal{P}; z) \geq X \sum_{\substack{d|\Pi(\mathcal{P},z) \\ d\in T^-(y)}} \frac{\mu^-(d)\omega(d)}{d} - \sum_{d\leq y,\, d|\Pi(\mathcal{P},z)} |R_d|$$

and

$$S(\mathcal{A}, \mathcal{P}; z) \leq X \sum_{\substack{d|\Pi(\mathcal{P},z) \\ d\in T^+(y)}} \frac{\mu^+(d)\omega(d)}{d} + \sum_{d\leq y,\, d|\Pi(\mathcal{P},z)} |R_d|\,.$$

One way to arrange for (1.5) and (1.6) to hold is as follows. Let

$$T := \{\, d \,|\, d \in \mathbb{N},\ \mu(d)^2 = 1 \,\},$$

write $d = p_1 p_2 p_3 \ldots$ with $p_1 > p_2 > p_3 \ldots$ for $d \in T$, and let

$$T_r = \{\, d \,|\, d \in T,\ (\nu(d) < r) \text{ or } (\nu(d) \geq r \,\&\, \mathfrak{P}(p_1, \ldots, p_r)) \,\}$$

for some predicate $\mathfrak{P}$ to be defined later. Let

$$T^+(y) = \bigcap_{u=1}^{\infty} T_{2u-1}, \quad T^-(y) = \bigcap_{u=1}^{\infty} T_{2u}\,.$$

With these definitions we have the following lemma.

Lemma 5.1 *For $m \in \mathbb{N}$ and $T^{\pm}(y)$ defined as above, we have*

$$\sum_{d|m} \mu^{-}(d) \leq \sum_{d|m} \mu(d) \leq \sum_{d|m} \mu^{+}(d) \,.$$

Proof. Let

$$B_{2u-1} = (T \setminus T_{2u-1}) \cap (\bigcap_{v<u} T_{2v-1}) \,,$$

then

$$T \setminus T^{+}(y) = \{d \in T : \exists v \,(d \notin T_{2v-1})\,\} = \bigcup_{u=1}^{\infty} B_{2u-1} \,.$$

Moreover we have $B_j \cap B_k = \emptyset$ if $j \neq k$. Therefore

$$\sum_{d|m} \mu(d) = \sum_{d|m,\, d \in T^{+}(y)} \mu(d) + U,$$

where

$$U := \sum_{u=1}^{\infty} \sum_{d|m,\, d \in B_{2u-1}} \mu(d) \,.$$

We set

$$C_r = \{d \in B_r : \nu(d) = r\,\},$$

and if $d > 1$ we write $p(d)$ for the smallest prime factor of d. We then define

$$Q(d) = \prod_{p \in \mathbb{P},\, p<p(d)} p, \quad (d > 1).$$

Now, if $d \in B_r$, then $d \notin T_r$. Thus $\nu(d) \geq r$ and $\mathfrak{P}(p_1, ..., p_r)$ does not hold. Write $d = ef$ with $e = p_1 ... p_r$ and $f | Q(e)$. Since $d \in B_r$, it follows that each property $\mathfrak{P}(p_1, ..., p_{r-2})$, $\mathfrak{P}(p_1, ..., p_{r-4})$, ... holds. Now let

$$C_r = \{d \in B_r : \nu(d) = r\,\} \,,$$

so that $e \in C_r$. Hence if $d \in B_r$ then we can write $d = ef$ with $e \in C_r$ and $f | Q(e)$. Clearly, the decomposition $d = ef$ with $e \in C_r$ and $f | Q(e)$ is unique.

Conversely, if $d = ef$ with $e \in C_r$ and $f | Q(e)$, then $d \in B_r$. Therefore

$$\begin{aligned}
\sum_{d|m,\, d\in B_{2u-1}} \mu(d) &= \sum_{\substack{ef|m,\, e\in C_{2u-1} \\ f|Q(e)}} \mu(ef) \\
&= \sum_{\substack{e\in C_{2u-1} \\ e|m}} \mu(e) \sum_{f|m,\, f|Q(e)} \mu(f) \\
&= \sum_{\substack{e|m,\, (m,Q(e))=1 \\ e\in C_{2u-1}}} \mu(e) \\
&= -\sum_{e|m,\, (m,Q(e))=1} 1 \\
&\le 0\,,
\end{aligned}$$

since $e \in C_{2u-1}$ implies that $\nu(e) = 2u - 1$ and hence that $\mu(e) = -1$. Thus

$$U = \sum_{u=1}^{\infty} \Big(\sum_{\substack{d\in B_{2u-1} \\ d|m}} \mu(d) \Big) \le 0\,,$$

so that

$$\begin{aligned}
\sum_{d|m} \mu^+(d) &= \sum_{d|m,\, d\in T^+(y)} \mu(d) \\
&= \sum_{d|m} \mu(d) - U \\
&\ge \sum_{d|m} \mu(d)
\end{aligned}$$

as claimed. The inequality

$$\sum_{d|m} \mu^-(d) \le \sum_{d|m} \mu(d)$$

can be proved in the same way.

This completes the proof of Lemma 5.1. However a useful alternative way of viewing the combinatoric facts used in the argument is as follows. We have

$$
\begin{aligned}
S(\mathcal{A},\mathcal{P};z) &= \sum_{n\in\mathcal{A}} \sum_{\substack{d|\Pi(\mathcal{P},z)\\ d|n}} \mu(d) \\
&= \sum_{n\in\mathcal{A}} \sum_{\substack{d|\Pi(\mathcal{P},z)\\ d|n}} \mu^+(d) - \sum_{u=1}^{\infty}\sum_{n\in\mathcal{A}} \sum_{\substack{d|n,\, d|\Pi(\mathcal{P},z)\\ d\in B_{2u-1}}} \mu(d) \\
&= \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)\#\mathcal{A}_d + \sum_{u=1}^{\infty}\sum_{n\in\mathcal{A}} \sum_{\substack{e|(n,\Pi(\mathcal{P},z)),\, e\in C_{2u-1}\\ (n,\Pi(\mathcal{P},z),Q(e))=1}} 1 \\
&= \sum_{d|\Pi(\mathcal{P},z)} \mu^+(d)\#\mathcal{A}_d + \sum_{u=1}^{\infty} \sum_{\substack{e\in C_{2u-1}\\ e|\Pi(\mathcal{P},z)}} S(\mathcal{A}_e,\mathcal{P};p(e)). \quad (5.3)
\end{aligned}
$$

Thus far, all we have said applies to any predicate $\mathfrak{P}$, and any sieve problem. We now specialize to a sieve problem of dimension 1, and examine (5.3) in the particular case $\mathcal{A}=\mathcal{A}^+$, which we expect to be extremal. Here we have

$$S(\mathcal{A}^+,\mathbb{P};z) = \frac{X}{e^\gamma\ \log(x^{1/s})}\,F(s)\ +\ O_N\Big(\frac{x}{(s-1)(\log x)^2}\Big)$$

for $1<s\le N$, by Theorem 4.1. Moreover the estimate (4.2) shows that

$$\sum_{d|\Pi(\mathbb{P},z)} \mu^+(d)\#\mathcal{A}_d^+ = X\sum_{d|\Pi(\mathbb{P},z)} \frac{\mu^+(d)}{d} + O(x(\log x)^{-2}).$$

We therefore conclude that

$$\sum_{d|\Pi(\mathbb{P},z)} \frac{\mu^+(d)}{d} = \frac{F(s)}{e^\gamma \log z} + \sum_{u=1}^{\infty}\sum_{e\in C_{2u-1}} S(\mathcal{A}_e^+,\mathbb{P};p(e)) + O_N((s-1)^{-1}(\log x)^{-2}),$$

where $y=E_1(x)$, $z=x^{1/s}$ and $1<s\le N$. If we replace $s=(\log x)/(\log z)$ by $s'=(\log y)/(\log z)$ then the right hand side above is

$$\frac{F(s')}{e^\gamma\log z} + \sum_{u=1}^{\infty} \sum_{\substack{e\in C_{2u-1}\\ e|\Pi(\mathcal{P},z)}} S(\mathcal{A}_e^+,\mathbb{P};p(e)) + O_N((s'-1)^{-1}(\log x)^{-3/2}) \quad (5.4)$$

for $1 < s' \le N$. We then re-define s as $(\log y)/(\log z)$. This produces an upper bound for the sum

$$\sum_{d|\Pi(\mathbb{P},z)} \frac{\mu^+(d)}{d} \tag{5.5}$$

which involves $F(s)$ together with information about the property $\mathfrak{P}$ incorporated in the definition of the sets C_r. Since our goal is to minimize the sum (5.5), we aim to choose $\mathfrak{P}$ so that $S(\mathcal{A}_e^+, \mathbb{P}; p(e))$ is as close to 0 as possible for $e \in C_{2u-1}$. However the relevant integers e all have $\nu(e) = 2u-1$, so that $\nu(m)$ is even for any $m \in \mathcal{A}_e^+$. Moreover, every such m satisfies $m \le x/e$. Hence we would have $S(\mathcal{A}_e^+, \mathbb{P}; p(e)) = 1$ providing that $x/e < p(e)^2$. Looking back at the definition of the set C_{2u-1} we see that we would want to have

$$p_1 p_2 \ldots p_{2u-3} p_{2u-1}^3 > x$$

whenever $\mathfrak{P}(p_1, \ldots, p_{2u-1})$ is false. Making a marginal adjustment to produce a condition which involves y rather than x we therefore take the property $\mathfrak{P}(p_1, \ldots, p_r)$ to say that

$$p_1 p_2 \ldots p_{r-1} p_r^3 < y,$$

whence

$$S^+(y) = \bigcap_{t=1}^{(r+1)/2} \{d \in \mathbb{N} : \ \mu(d)^2 = 1, \ (d = p_1 ... p_r \ \Rightarrow p_1 p_2 ... p_{2t-1}^3 < y)\} .$$

Although we have not made $S(\mathcal{A}_e^+, \mathbb{P}; p(e))$ completely vanish for e in C_{2u-1} it can be shown that this construction does indeed make the sum in (5.4) suitably small. We have therefore produced an admissible set of upper bound sieve coefficients $\mu^+(d)$ which match up with the Selberg sequence $\mathcal{A}^+$, and the linear programming argument then shows that both are optimal.

One can discuss the lower bound problem in exactly the same way, using the sequence $\mathcal{A}^-$, and leading to the choice

$$S^-(y) = \bigcap_{t=1}^{r/2} \{d \in \mathbb{N} : \ \mu(d)^2 = 1, \ (d = p_1 ... p_r \ \Rightarrow p_1 p_2 ... p_{2t}^3 < y)\} .$$

The construction of $\mu^\pm(d)$ we have been led to is known as the *Rosser-Iwaniec* sieve for dimension 1, there being variants in other dimensions. (The reader should note that, except for dimensions 1 and 1/2, the general Rosser-Iwaniec sieve is not known to be optimal. Indeed in many case it is known not to be optimal.) One noteworthy feature of the construction is that the definition of the weights $\mu^{\pm}(d)$ does not involve either the parameter z or the function $\omega(d)$.

Although our discussion has been concerned with the case $\omega(d) = 1$, the Rosser-Iwaniec weights may be applied to the general sieve problem of dimension 1. Thus if we set

$$M^{\pm}(z, y) = \sum_{d|\Pi(\mathcal{P},z)} \frac{\omega(d)\,\mu^{\pm}(d)}{d}$$

we will have

$$S(\mathcal{A}, \mathcal{P}; z) \leq M^{+}(z, y)X + \sum_{d|\Pi(\mathcal{P},z)} |R_d| \tag{5.6}$$

and

$$S(\mathcal{A}, \mathcal{P}; z) \geq M^{-}(z, y)X - \sum_{d|\Pi(\mathcal{P},z)} |R_d|. \tag{5.7}$$

Iwaniec [12] has established the following bounds for $M^{+}(z, y)$ and $M^{-}(z, y)$.

Theorem 5.1 *Suppose that*

$$\sum_{w\leq p<z} \frac{\omega(p)\log p}{p} \leq \log\frac{z}{w} + O(1)$$

and write, as usual,

$$W(z) = \prod_{p<z} \big(1 - \frac{\omega(p)}{p}\big)\,.$$

Then

$$M^{+}(z, y) \leq W(z)\{F(s) + O(e^{-s}(\log y)^{-1/3})\}$$

and

$$M^{-}(z, y) \leq W(z)\{f(s) + O(e^{-s}(\log y)^{-1/3})\}\,.$$

We shall not prove this theorem here.

Remark. Note that Iwaniec requires only a one-sided condition, in contrast to the two-sided condition (3.7) introduced in the context of the Selberg sieve. Thus Theorem 5.1 applies to sieves of dimension less than 1, and even to certain problems without a well-defined dimension.

In view of (5.6) and (5.7), one obtains the following inequalities.

Corollary 5.1 *Under the condition of Theorem 5.1 we have*

$$S(\mathcal{A},\mathcal{P};z) \le XW(z)\{F(s)+O(e^{-s}(\log y)^{-1/3})\} + \sum_{d<y}|R_d|$$

and

$$S(\mathcal{A},\mathcal{P};z) \ge XW(z)\{f(s)+O(e^{-s}(\log y)^{-1/3})\} - \sum_{d<y}|R_d|.$$

Example 1. Let $\mathcal{A}=\{n\in\mathbb{N}:\ n\le x\}$, $X=x$, $\omega(d)=1$. Then $R_d=O(1)$ since

$$\#\mathcal{A}_d = [\frac{x}{d}] = \frac{X\omega(d)}{d} + O(1)\,.$$

From Corollary 5.1 one obtains

$$S(\mathcal{A},\mathbb{P};z) \ge x\prod_{p<z}\big(1-\frac{1}{p}\big)(f(s)+o(1)) - \sum_{d<y}|R_d|. \tag{5.8}$$

Let $y=\frac{x}{(\log x)^2}$ and suppose that $y^{1/4}<z<y^{1/2}$. Then

$$2 < s = \frac{\log y}{\log z} < 4,$$

whence

$$f(s) = \frac{2e^{\gamma}\log(s-1)}{s}.$$

Thus (5.8), in conjunction with Mertens Theorem (1.4), gives

$$\begin{aligned} S(\mathcal{A},\mathbb{P};z) \ &\ge\ x\{\frac{e^{-\gamma}}{\log z}+O(\frac{1}{(\log z)^2})\}\left(\frac{2e^{\gamma}\log(s-1)}{s}+O((\log y)^{-1/3})\right) \\ &\qquad +O\big(\frac{x}{(\log x)^2}\big) \\ &=\ \frac{2x}{\log y}\{1+o(1)\}\log(s-1)+O\big(\frac{x}{(\log x)^2}\big). \end{aligned}$$

Hence

$$S(\mathcal{A},\mathbb{P};z) \gg \frac{x}{\log x}$$

for any constant value of s strictly greater than 2. To detect primes, one would need to consider the situation with $s=2$. We just fail to find primes, which is not surprising since the sequence $\mathcal{A}^-$ contains no primes (the *parity phenomenon!*).

Example 2. Let $0 < \theta < 1$, and choose

$$\mathcal{A} = \{\, n \,:\, x - x^{\theta} < n \le x \,\},$$

$X = x^{\theta},\ \omega(d) = 1$, and

$$y = \frac{x^{\theta}}{\log x},\ z = x^{\frac{\theta}{2}-\delta}$$

with $\delta > 0$, so that

$$s = \frac{\log y}{\log z} = \frac{2\theta}{\theta - 2\delta} + o(1) > 2$$

and $R_d = O(1)$. As in Example 1, it follows that

$$S(\mathcal{A}, \mathbb{P}; z) \gg \frac{x^{\theta}}{\log x}\,.$$

Hence, if x is large enough, the interval $x - x^{\theta} < n \le x$ contains at least one integer n all of whose prime factors p satisfy $p \ge x^{\theta/2-\delta}$. In particular if $r \in \mathbb{N}$ and $\theta > 2/(r+1)$, then we may choose $\delta > 0$ so that $\theta/2 - \delta > 1/(r+1)$. Thus we will have $p > x^{1/(r+1)}$ so that n can have at most r prime factors.

In general we say that a positive integer n is an *almost prime* of type P_r, if it has at most r prime factors, counted according to multiplicity. We may then conclude that if $\theta > 2/(r+1)$, and if x is sufficiently large, then the interval $x - x^{\theta} < n \le x$ contains at least one P_r number. The necessary size for θ can be reduced (see Example 1 after Theorem 6.2) and it is an interesting problem to know just small it may be taken.

Example 3 (The *twin primes problem*). Let

$$\mathcal{A} = \{\, p + 2 \,:\, p \in \mathbb{P},\ p \le x \,\},$$

and take $X = \pi(x)$ and

$$\omega(p) = \begin{cases} \frac{p}{p-1}, & p > 2, \\ 0, & p = 2. \end{cases}$$

In view of the Bombieri-Vinogradov Theorem, one can take $y = x^{1/2}(\log x)^{-c_2}$ to obtain

$$\sum_{d<y} |R_d| = O\Big(\frac{x}{(\log x)^3}\Big)\,.$$

As above, one then concludes that

$$S(\mathcal{A}, \mathbb{P}; z) \gg \frac{x}{(\log x)^2} \quad \text{if } s > 2\,.$$

Since $s = (\log y)/(\log z)$ we may therefore use $z = x^\theta$ with any constant exponent $\theta < 1/4$. It follows then that the sequence $\mathcal{A}$ contains a growing number of P_4 integers as x tends to infinity. As we shall see in Theorem 6.3, this has been improved by Chen [3] who shows that the same is true for P_2 integers.

Example 4. Let $\mathcal{A} = \{\, n^2 + 1 \,:\, n \le x \,\}$, and take $X = x$ and

$$\omega(p) = \begin{cases} 2, & p \equiv 1 \pmod 4, \\ 0, & p \equiv -1 \pmod 4, \\ 1, & p = 2. \end{cases}$$

Then $R_d = O(\omega(d))$ and

$$\sum_{w \le p < z} \frac{\omega(p) \log p}{p} = \log \frac{z}{w} + O(1).$$

If we take $y = x(\log x)^{-2}$ it follows that

$$\sum_{d<y} |R_d| = O\big(\frac{x}{(\log x)^2}\big).$$

Thus if $z = x^\theta$ with a constant exponent $\theta > 1/2$ we will find that $s > 2$ and hence

$$S(\mathcal{A}, \mathbb{P}; z) \gg \frac{x}{\log x}.$$

It follows, on taking $\theta > 1/5$, that there are infinitely many P_4 numbers of the form $n^2 + 1$.

In a similar way one can prove that if f is an irreducible integer polynomial such that the values $\{\, f(n) \,|\, n \in \mathbb{N} \,\}$ contain no common factor, then $f(n) = P_{2k}$ infinitely often. Here also improvements are possible.

We conclude this section with the following important result.

The "Fundamental Lemma" *Suppose that*

$$\sum_{d<y} |R_d| \ll \frac{x}{(\log x)^2}\,.$$

Then

$$S(\mathcal{A}, \mathcal{P}; z) \sim XW(z),$$

if $\log z = o(\log y)$, *that is, if* $s \to \infty$.

Proof. It suffices to observe that $F(s) = 1 + O(e^{-s})$ and $f(s) = 1 + O(e^{-s})$, as $s \to \infty$.

Remarks.

1) This result should be compared with Corollary 2.1, in which one required $z \leq \log y$.

2) Since one can often choose $y = X^{\delta}$ for some $\delta > 0$ it follows from the Fundamental Lemma that

$$S(\mathcal{A}, \mathcal{P}; z) \sim XW(z),$$

if $z = X^{\varepsilon(X)}$ with $\varepsilon(X) \to 0$ as $X \to \infty$. By considering the sequences $\mathcal{A}^{\pm}$ one sees that one cannot have such a result when $\varepsilon(X) \not\to 0$.

3) The significance of the Fundamental Lemma is that one can sieve out "small primes" (namely those below $X^{\varepsilon(X)}$) as an initial stage in some more complicated argument, and still have an asymptotic formula. In order to make use of this information one usually wants a quantitative form of the Fundamental Lemma, but this is easily established.

4) One can obtain analogous results on the weaker assumption that

$$\sum_{p \leq z} \frac{\omega(p) \log p}{p} \ll \log z$$

as $z \to \infty$.

6 The weighted sieve

If the sequence $\mathcal{A}$ contains only positive integers $n \leq N$, and we can show that $S(\mathcal{A}, \mathbb{P}, N^{\theta}) > 0$ for some $\theta > 1/(r+1)$, then we can conclude that $\mathcal{A}$ contains at least one P_r number. However one can often derive better results by using a *weighted sieve*, in which certain P_r numbers with prime factors below $N^{1/(r+1)}$ are also counted.

In general we let $N = \max_{n \in \mathcal{A}} |n|$ and we choose constants $0 < \alpha < \beta$. We then set

$$W := W(\mathcal{A}, \mathcal{P}; \alpha, \beta) = \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi(\mathcal{P}, N^{\alpha}))=1}} \Big(1 - \sum_{\substack{p|n \\ N^{\alpha} \leq p < N^{\beta}}} w_p\Big),$$

where the weights $w_p \geq 0$ are to be chosen so that

$$\sum_{\substack{p|n \\ N^{\alpha} \leq p < N^{\beta}}} w_p \ \geq \ 1 \quad \text{if} \quad \nu(n) \geq r+1. \tag{6.1}$$

This condition ensures that

$$
\begin{aligned}
W &\leq \sum_{\substack{n\in\mathcal{A},\,\nu(n)\leq r\\(n,\Pi(\mathcal{P},N^\alpha))=1}} \big(1-\sum_{\substack{p|n\\N^\alpha\leq p<N^\beta}} w_p\big)\\
&\leq \sum_{n\in\mathcal{A},\,\nu(n)\leq r} 1.
\end{aligned}
$$

If we know that

$$\#\{n\in\mathcal{A}:\ \exists p^2|n,\ N^\alpha\leq p<N^\beta\}\ll\frac{X}{(\log N)^2},$$

we can deduce that

$$W\leq\#\{n\in\mathcal{A}:\ n=P_r\}+O(\frac{X}{(\log N)^2}).$$

Thus if we can also show that

$$W\gg\frac{X}{\log N}, \tag{6.2}$$

we will be able to deduce that

$$\#\{n\in\mathcal{A}:\ n=P_r\}\gg\frac{X}{\log N}.$$

The optimal choice for the weights w_p is not known, however the choice

$$w_p=\frac{\beta}{(r+1)\beta-1}\big(1-\frac{\log p}{\beta\,\log N}\big)$$

leads to some fairly satisfactory results. These are known as *Richert's logarithmic weights.*

We shall assume henceforth that

$$(r+1)\beta-1>0.$$

Then if $p<N^\beta$ we will have $w_p\geq 0$ as required. Now suppose that $n\in\mathcal{A}$ with h.c.f.$(n,\Pi(\mathcal{P},N^\alpha))=1$, and consider the sum

$$S=\sum_{\substack{p|n\\N^\alpha\leq p<N^\beta}} w_p\,.$$

Then

$$\begin{aligned} S &= \frac{\beta}{(r+1)\beta-1} \sum_{\substack{p|n \\ N^\alpha \le p < N^\beta}} \big(1 - \frac{\log p}{\beta \log N}\big) \\ &\ge \frac{\beta}{(r+1)\beta-1} \sum_{p|n} \big(1 - \frac{\log p}{\beta \log N}\big), \end{aligned}$$

since n has no prime factors $p < N^\alpha$, and

$$1 - \frac{\log p}{\beta \log N} \le 0$$

for any prime factor $p \ge N^\beta$. However

$$\sum_{p|n} \log p \le \log |n| \le \log N,$$

so that

$$\sum_{p|n} \big(1 - \frac{\log p}{\beta \log N}\big) \ge \nu(n) - \frac{1}{\beta}.$$

Hence if $\nu(n) \ge r+1$ we will have

$$S \ge \frac{\beta}{(r+1)\beta-1} \big(\nu(n) - \frac{1}{\beta}\big) \ge 1,$$

as required for (6.1).

We now examine the estimate (6.2). By definition,

$$W = S(\mathcal{A}, \mathcal{P}; N^\alpha) - \sum_{N^\alpha \le p < N^\beta} w_p \, S(\mathcal{A}_p, \mathcal{P}; N^\alpha)\,.$$

We plan to apply Corollary 5.1. We therefore make the assumption that

$$\sum_{d < N^\gamma} |R_d| \ll \frac{X}{(\log N)^2}$$

for some fixed $\gamma > 0$, and we set $y = N^\gamma$. According to Corollary 5.1 we will then have

$$S(\mathcal{A}, \mathcal{P}; N^\alpha) \ge XW(N^\alpha)\,\big(f\big(\frac{\gamma}{\alpha}\big) + o(1)\big).$$

Moreover, since

$$(\mathcal{A}_p)_d = \frac{X\omega(p)}{p}\,\frac{\omega(d)}{d} + R_{pd},$$

it follows that

$$S(\mathcal{A}_p,\mathcal{P};N^\alpha) \le \frac{X\omega(p)}{p}\,W(N^\alpha)\,\bigl(F\bigl(\frac{\log N^\gamma/p}{\log N^\alpha}\bigr)+o(1)\bigr)+\sum_{d<\frac{N^\gamma}{p},\, d|\Pi(\mathcal{P},N^\alpha)}|R_{pd}|\,.$$

Moreover,

$$\sum_{N^\alpha\le p<N^\beta} w_p \sum_{d<N^\gamma/p,\, d|\Pi(\mathcal{P},N^\alpha)} |R_{pd}| \ll \sum_{k<N^\gamma}|R_k| \ll \frac{X}{(\log N)^2},$$

since an integer $k<N^\gamma$ can have at most γ/α prime factors $p\ge N^\alpha$. It follows that

$$\begin{aligned}
&\sum_{N^\alpha\le p<N^\beta} w_p\, S(\mathcal{A}_p,\mathcal{P};N^\alpha)\\
&\le\ XW(N^\alpha)\sum_{N^\alpha\le p<N^\beta} w_p\,\frac{\omega(p)}{p}\,\{F\bigl(\frac{\log N^\gamma/p}{\log N^\alpha}\bigr)+o(1)\}+O(\frac{X}{(\log N)^2})\\
&\le\ XW(N^\alpha)\sum_{N^\alpha\le p<N^\beta} w_p\,\frac{\omega(p)}{p}F\bigl(\frac{\log N^\gamma/p}{\log N^\alpha}\bigr)+o(\frac{X}{\log N}).
\end{aligned}$$

We therefore conclude that

$$W \ge XW(N^\alpha)\{f\bigl(\frac{\gamma}{\alpha}\bigr)-\sum_{N^\alpha\le p<N^\beta} w_p\,\frac{\omega(p)}{p}F\bigl(\frac{\log N^\gamma/p}{\log N^\alpha}\bigr)\}+o(\frac{X}{\log N}).$$

We summarize our conclusions as follows.

Theorem 6.1 *Suppose that*

$$\mathcal{A}\subseteq\ \mathbb{Z}\cap[-N,N]$$

and that

$$\#\mathcal{A}_d = X\,\frac{\omega(d)}{d}+R_d\,,$$

and assume that the following conditions hold.

(i)

$$\sum_{z<p\le w}\frac{\omega(p)\log p}{p} = \log w/z + O(1),\quad (2\le z\le w);$$

(ii)

$$\beta > \alpha > 0, \quad (r+1)\beta - 1 > 0\,;$$

(iii)

$$\#\{n \in \mathcal{A} :\, \exists p^2 | n,\, N^\alpha \le p < N^\beta\} \ll \frac{X}{(\log N)^2};$$

(iv)

$$\sum_{d < N^\gamma} |R_d| \ll \frac{X}{(\log N)^2};$$

(v)

$$f\big(\frac{\gamma}{\alpha}\big) - \sum_{N^\alpha \le p < N^\beta} w_p \frac{\omega(p)}{p}\, F\big(\frac{\log N^\gamma / p}{\log N^\alpha}\big) \gg 1$$

where

$$w_p = \frac{\beta}{(r+1)\beta - 1}\big(1 - \frac{\log p}{\beta \log N}\big)\,.$$

Then the sequence $\mathcal{A}$ *contains* $\gg \frac{X}{\log N}$ *, numbers of type* P_r.

Our task now is to examine condition (v) in the above theorem. Let

$$S(t) = \sum_{N^\alpha \le p < t} \frac{\omega(p) \log p}{p}$$

and

$$h(p) = \frac{w_p}{\log p}\, F\big(\frac{\log N^\gamma / p}{\log N^\alpha}\big)\,,$$

so that

$$\sum_{N^\alpha \le p < N^\beta} w_p \frac{\omega(p)}{p}\, F\big(\frac{\log N^\gamma / p}{\log N^\alpha}\big) = \sum_{N^\alpha \le p < N^\beta} \frac{\omega(p) \log p}{p}\, h(p).$$

By partial summation we find that

$$\begin{aligned} \sum_{N^\alpha \le p < N^\beta} \frac{\omega(p) \log p}{p}\, h(p) &= \big[S(t)h(t)\big]_{N^\alpha}^{N^\beta} - \int_{N^\alpha}^{N^\beta} S(t) h'(t)\, dt \\ &= -\int_{N^\alpha}^{N^\beta} S(t) h'(t)\, dt, \qquad (6.3) \end{aligned}$$

since $S(N^\alpha) = 0$ and $h(N^\beta) = 0$. According to assumption (i) we have $S(t) = \log t - \alpha \log N + O(1)$. The contribution to (6.3) arising from the error term is

$$\ll \int_{N^\alpha}^{N^\beta} |h'(t)|\, dt = \int_\alpha^\beta |\frac{d\,h(N^v)}{d\,v}|\, dv.$$

However

$$h(N^v) = \frac{1}{\log N}\frac{\beta}{(r+1)\beta - 1}\big(\frac{1}{v} - \frac{1}{\beta}\big) F\big(\frac{\gamma - v}{\alpha}\big),$$

whence

$$\frac{d\,h(N^v)}{d\,v} \ll (\log N)^{-1}$$

uniformly for $\alpha \le v \le \beta$. We therefore conclude that (6.3) is

$$\begin{aligned}
&-\int_{N^\alpha}^{N^\beta} \{\log t - \alpha \log N + O(1)\} h'(t)\, dt \\
&= -\big[\{\log t - \log N^\alpha\} h(t)\big]_{N^\alpha}^{N^\beta} + \int_{N^\alpha}^{N^\alpha} \frac{h(t)}{t} dt + O\big(\frac{1}{\log N}\big) \\
&= \frac{\beta}{(r+1)\beta - 1}\int_\alpha^\beta \big(\frac{1}{v} - \frac{1}{\beta}\big) F\big(\frac{\gamma - v}{\alpha}\big)\, dv + O\big(\frac{1}{\log N}\big).
\end{aligned}$$

Hence the inequality

$$f\big(\frac{\gamma}{\alpha}\big) > \frac{\beta}{(r+1)\beta - 1}\int_\alpha^\beta \big(\frac{1}{v} - \frac{1}{\beta}\big) F\big(\frac{\gamma - v}{\alpha}\big)\, dv \tag{6.4}$$

is necessary and sufficient for condition (v) of Theorem 6.1

In order to express f and F in terms of elementary functions we shall impose the condition $\gamma/4 \le \alpha \le \gamma/2$. For this range we will have

$$f\big(\frac{\gamma}{\alpha}\big) = 2e^{\gamma_0}\frac{\log(\gamma/\alpha - 1)}{\gamma/\alpha}$$

and

$$F\big(\frac{\gamma - v}{\alpha}\big) = 2e^{\gamma_0}\frac{\alpha}{\gamma - v},$$

where we have written γ_0 for Euler's constant, to avoid confusion with the parameter γ. Thus (6.4) is equivalent to the condition

$$\log\big(\frac{\gamma}{\alpha} - 1\big) > \frac{\beta}{(r+1)\beta - 1}\int_\alpha^\beta \frac{\gamma}{\gamma - v}\big(\frac{1}{v} - \frac{1}{\beta}\big)\, dv\,.$$

We can now perform the integration on the right hand side, to obtain

$$\log\big(\frac{\gamma}{\alpha}-1\big) > \frac{1}{(r+1)\beta-1}\big(\beta\log\frac{\beta}{\alpha}-(\gamma-\beta)\log\frac{\gamma-\alpha}{\gamma-\beta}\big)\,. \tag{6.5}$$

We shall choose $\alpha=\gamma/4$ and

$$\beta=\frac{\gamma}{1+3^{-r}}.$$

(These are in fact optimal, as a relatively easy calculation shows. However we do not need to know that the choice is optimal to proceed.) The above values are compatible with condition (ii) of Theorem 6.1 providng that

$$\gamma > \frac{1+3^{-r}}{r+1}. \tag{6.6}$$

Moreover (6.5) then reduces to

$$\gamma > \frac{1}{r+1-\frac{\log 4/(1+3^{-r})}{\log 3}},$$

which is a stronger condition than (6.6).

We therefore have the following result.

Theorem 6.2 *Suppose the assumptions of Theorem 6 hold, with*

$$\alpha=\gamma/4, \quad \textit{and} \quad \beta=\frac{\gamma}{1+3^{-r}},$$

and with condition (v) replaced by

$$\gamma > \frac{1}{\Lambda_r}\,,$$

where

$$\Lambda_r := r+1-\frac{\log 4/(1+3^{-r})}{\log 3}\,.$$

Then the sequence $\mathcal{A}$ contains $\gg \frac{X}{\log N}$, numbers of type P_r.

Remarks.

1) For $r\geq 2$ we have $r-\frac{2}{7}<\Lambda_r<r-\frac{1}{7}$. In particular we have $\Lambda_2\geq\frac{11}{6}$.

2) The only parameters which enter into the theorem in a crucial way are N, which measures the size of elements of $\mathcal{A}$, and γ which measures the size of the remainders, in terms of N. The parameter γ is often called the "level

of distribution" (or more precisely, since we may not know the optimal value for γ, an "admissible level of distribution").

Example 1. Let

$$\mathcal{A} = \{n \in \mathbb{N} : \ x - x^{\theta} < n \le x\},$$

and take $\mathcal{P} = \mathbb{P}$, $X = x^{\theta}$ and $\omega(p) = 1$. Then $R_d \ll 1$ so that we may choose any $\gamma < \theta$. Then the assumptions of Theorem 6.2 hold true providing that $\gamma > \Lambda_r^{-1}$. We therefore conclude that $\mathcal{A}$ contains a P_r almost-prime if x is large enough, providing that $\theta > \Lambda_r^{-1}$.

For $r = 2$ much stronger results are known. According to work of Baker, Harman and Pintz [1], the sequence $\mathcal{A}$ actually contains a prime, for the exponent $\theta = 0.525$, which is smaller than Λ_2^{-1}. Moreover Liu [13] has show that there are P_2's as soon as $\theta \ge 0.436$. It would be nice to know that $\theta > 1/r$ sufficed to ensure the existence of P_r's in $\mathcal{A}$, for every r.

Example 2. Let N be an even integer and put

$$\mathcal{A} = \{\, N - p \, : \, p \in \mathbb{P}, \ 3 \le p \le N - 3 \,\},$$

$$\mathcal{P} = \{p \in \mathbb{P} : \ p \nmid N\},$$

$$\omega(p) = \begin{cases} 0, & \text{if } p|N, \\ \frac{p}{p-1}, & \text{if } p \nmid N, \end{cases}$$

and $X = \mathrm{Li}(N)$. As in our discussion of this example in §3, we find, via the Bombieri-Vinogradov Theorem, that any $\gamma < 1/2$ will be admissable for the remainder sum. Since $\Lambda_3 > 2$ this suffices to show that$\mathcal{A}$ contains a P_3 for large enough N, so that every sufficiently large even integer $2n$ may be written as a sum of a prime and a P_3 almost prime.

In this second example we see that Λ_2 is only just less than 2, so we come quite close to handling P_2's this way. However to achieve this requires an ingenious new idea.

Chen's theorem. *Every sufficiently large even integer N is a sum of a prime and a P_2 almost-prime.*

More precisely, for every sufficiently large positive integer N we have

$$\#\{p \in \mathbb{P} : \ N - p \ \in \mathbb{P}_2\} \ge 0.335\, C_2 \,(\prod_{p|N,\, p\ne 2} \frac{p-1}{p-2}) \, \frac{N}{(\log N)^2},$$

where

$$C_2 := 2 \prod_{p|N,\, p\ne 2} (1 - \frac{1}{(p-1)^2})$$

as in § 3.

Sketch proof. Given an even positive integer N, let

$$\mathcal{A} = \{\, N - p \,:\, p \in \mathbb{P},\ 3 \le p \le N-3\,\}.$$

Define $\mathcal{P}$ and $\omega(d)$ as in the previous example, and let $z > 2$. For $n \in \mathbb{N}$, let $a(n) = 1$ if

$$n = p_1 p_2 p_3, \quad p_i \text{ prime}, \quad p_1 < N^{1/3} \le p_2 \le p_3, \tag{6.7}$$

and let $a(n) = 0$ otherwise. We then consider the sum

$$S_0 := S(\mathcal{A}, \mathcal{P}; z) - \frac{1}{2}\sum_{z \le p < N^{1/3}} S(\mathcal{A}_p, \mathcal{P}; z) - \frac{1}{2}\sum_{n\in\mathcal{A},\,(n,\Pi(\mathcal{P},z))=1} a(n).$$

The first two terms of this may be thought of as giving a weighted sieve, with constant weights $w_p = 1/2$.

Write S_0^* for the contribution to S_0 arising from those values of n which are not square-free. If z is a positive power of N then it is easily shown that

$$S_0^* \ll \frac{N}{(\log N)^3},$$

which will be negligible. (In proving this it is useful to note that if $p^2|n$ and $(n, \Pi(\mathcal{P}, z)) = 1$, then $p \ge z$.)

Let $w(n)$ be the weight attached to n in the expression S_0. Clearly we have $w(n) \le 1$ for every n. We claim that $w(n) \le 0$ for any square-free integer $n \in \mathcal{A}$, unless n is a P_2. Subject to this assertion, we will then have

$$S_0 \le \#\{n \in \mathcal{A} :\ n = P_2\} + O(\frac{N}{(\log N)^3}).$$

To verify the claim take a square-free integer $n \in \mathcal{A}$ with $w(n) > 0$. Then we will have $(n, \Pi(\mathcal{P}, z)) = 1$. Moreover there can be at most one prime factor $p|n$ in the range $z \le p < N^{1/3}$, and clearly any integer $n < N$ can have at most two prime factors $p \ge N^{1/3}$. Thus if n is not a P_2 almost-prime it must be of the form (6.7), so that $a(n) = 1$. However it is clear that in this case we have $w(n) = 0$. This establishes the claim.

The terms $S(\mathcal{A}, \mathcal{P}; z)$ and $S(\mathcal{A}_p, \mathcal{P}; z)$ are estimated from below and above respectively, just as in the standard weighted sieve. However it is necessary to choose z somewhat smaller than before, as $z = N^{1/10}$. As a result one has to evaluate $f(5)$, for example, by numerical integration.

However the key new ingredient is the treatment of the sum

$$\sum_{n\in\mathcal{A},\,(n,\Pi(\mathcal{P},z))=1)} a(n).$$

Hitherto the only information about $\mathcal{A}$ that we have used comes from the estimate

$$\#\mathcal{A}_d = X\frac{\omega(d)}{d} + R_d. \tag{6.8}$$

However we now use the precise structure of $\mathcal{A}$ to re-write the sum above as

$$\#\{p \in \mathcal{B}\} = S(\mathcal{B}, \mathbb{P}; N^{1/2}\} + O(N^{1/2}),$$

where

$$\mathcal{B} = \{\, N - p_1p_2p_3 \,:\, p_1p_2p_3 < N,\ z \le p_1 < N^{1/3} \le p_2 < p_3 \,\}.$$

Thus we change our attention to a quite different sequence. This device has been called the *"reversal of rôles"*, or *"Chen's twist"*. (One should note however that although Chen's application of this idea is arguably the most spectacular, the principle was independently discovered by Iwaniec, amongst others.)

Although the set $\mathcal{B}$ looks complicated, it is, in fact, essentially as simple as $\mathcal{A}$. An analogue of the Bombieri-Vinogradov Theorem can be establised, showing that if

$$\#\mathcal{B}_d = X\frac{\omega(d)}{d} + R_d$$

with a suitable value for X, then

$$\sum_{d<N^\gamma} |R_d| \ll \frac{X}{(\log X)^3}$$

for any fixed $\gamma < \frac{1}{2}$.

The standard Selberg upper bound for $S(\mathcal{B}, \mathbb{P}; N^{\frac{1}{2}})$ now allows one to complete the proof of a positive lower bound for S_0.

Remark

While the parity phenomenon gives a limitation to the power of sieve methods which are based purely on the relation (6.8), it is no longer relevant once one uses additional information. Thus the reversal of rôles trick has the potential to circumvent the parity problem.

Other applications of the reversal of rôles trick.

1) One can show via the circle method that are infinitely many triples of distinct primes p_1, p_2, p_3 which form an arithmetic progression, so that $p_2 - p_1 = p_3 - p_2$. On the other hand it is an open problem whether or not there are infinitely many 4-tuples of distinct primes in arithmetic progression.

However, one can combine the circle method with the sieve, and use the reversal of rôles trick to give infinitely many 4-tuples p_1, p_2, p_3, n in arithmetic progression with n a P_2 almost-prime, (see Heath-Brown [8]).

2) In Example 3 of §5 we applied the Rosser-Iwaniec lower bound sieve, together with the Bombieri-Vinogradov Theorem, to the sequence

$$\mathcal{A} = \{p+2 : \, p \in \mathbb{P}, \, p \leq x\}.$$

This was enough to show that if $\theta < 1/4$ then

$$S(\mathcal{A}, \mathbb{P}, x^{\theta}) \gg \frac{x}{(\log x)^2}. \tag{6.9}$$

In particular this shows that $\mathcal{A}$ contains P_4 numbers if we choose $\theta > 1/5$.

However by using the reversal of rôles trick it is possible to show the existence of an admissible constant $\theta > 1/4$ for which (6.9) still holds, thereby showing that $\mathcal{A}$ contains P_3 numbers, without the need for a weighted sieve.

References

[1] R.C. Baker, G. Harman and J. Pintz, The difference between consecutive primes, II, *Proc. London Math. Soc. (3),* 83 (2001), 532-562.

[2] E. Bombieri, J.B. Friedlander and H. Iwaniec, Primes in arithmetic progressions to large moduli, *Acta Math.,* 156 (1986), 203-251.

[3] J.-R. Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes, *Sci. Sinica,* 16 (1973), 157-176.

[4] J.-R. Chen, On the Goldbach's problem and the sieve methods, *Sci. Sinica,* 21 (1978), 701-739.

[5] G.R.H. Greaves, *Sieves in number theory,* Ergebnisse der Mathematik (3), 43. (Springer-Verlag, Berlin, 2001).

[6] H. Halberstam and H.-E. Richert, *Sieve methods* (Academic Press, London-New York, 1974).

[7] D.R. Heath-Brown, Almost-primes in arithmetic progressions and short intervals, *Math. Proc. Cambridge Philos. Soc.,* 83 (1978), 357-375.

[8] D.R. Heath-Brown, Three primes and an almost-prime in arithmetic progression *J. London Math. Soc. (2),* 23 (1981), 396-414.

[9] D.R. Heath-Brown, The number of primes in a short interval, *J. Reine Angew. Math.,* 389 (1988), 22-63.

[10] D. Hensley and I. Richards, On the incompatibility of two conjectures concerning primes, *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, 123-127. (Amer. Math. Soc., Providence, R.I., 1973).

[11] H. Iwaniec, Almost-primes represented by quadratic polynomials, *Invent. Math.,* 47 (1978), 171-188.

[12] H. Iwaniec, Rosser's sieve, *Acta Arith.,* 36 (1980), 171-202.

[13] H.-Q. Liu, Almost primes in short intervals, *J. Number Theory,* 57 (1996), 303-322.

[14] H. Maier, Primes in short intervals, *Michigan Math. J.,* 32 (1985), 221-225.

[15] H.L. Montgomery and R.C, Vaughan, The large sieve, *Mathematika,* 20 (1973), 119-134.

[16] K. Prachar, *Primzahlverteilung,* (Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957).