1

NUMBERS WITH A LARGE PRIME FACTOR II

ROGER BAKER AND GLYN HARMAN

To Klaus Roth with warmest good wishes on the occasion of his 80th birthday.

1. Introduction

Let x be a large positive number and $y = x^{1/2}$. Even if we assume the Riemann hypothesis, it appears to be hopelessly difficult to show that there is a prime number p in the interval $\mathcal{I} = (x, x+y]$. One approach is to assume that there are Siegel zeros. By making a precise hypothesis of this nature, Friedlander and Iwaniec [3] show that there are primes in intervals $[x, x + x^{39/79}]$ for long ranges of x.

Ramachandra [13] suggested the problem of showing that there is a number n in $\mathcal{I}$ having a large prime factor p with $p > x^{\phi}$. This is an "approximation" to the original question. Here of course ϕ is to be made as large as possible. Increasing values of ϕ for which such a p can be shown to exist have been provided by Ramachandra [13, 14], Graham [4], Baker [1], Jia [7, 8, 9, 10], Liu [11], Baker and Harman [2], Liu and Wu [12] and Harman [6].

In Harman's book, the value of ϕ is 0.74, and it is noted that recent work on exponential sums due to Wu [17] and Robert and Sargos [16] give room for further progress. In this chapter, we pursue this programme, and prove the following result. We write $P(n)$ for the largest prime factor of a natural number n, and $Q(n)$ for the smallest prime factor of n, with $Q(1) = 1$.

Theorem. *For all sufficiently large x, there are integers n in the interval $\mathcal{I}$ with*

$$P(n) > x^{0.7428}.$$

We shall quote liberally from earlier works on the subject, especially [2, 12, 6]. Let ϵ be a positive number, which we suppose is sufficiently small. Let

$$N(d) = \sum_{\substack{n \in \mathcal{I} \\ d | n}} 1,$$

$L = \log x$, $U = x^{3/5-\epsilon}$ and $\phi = 0.7428$. Then (see [6, Section 6.2])

$$\sum_{d<x} \Lambda(d) N(d) = \sum_{n \in \mathcal{I}} (\log n - \Lambda(n)) = yL + O(y),$$

$$\sum_{d \leqslant U} \Lambda(d) N(d) = \left(\frac{3}{5} - \epsilon \right) yL + O(y),$$

and

$$\sum_{\substack{U<d<x \\ d \text{ not prime}}} \Lambda(d) N(d) = O(y).$$

It suffices for the proof of our theorem to show that

$$\sum_{U<p\leqslant x^{\phi}} (\log p)N(p) < \left(\frac{2}{5}+\frac{\epsilon}{2}\right) yL. \tag{1.1}$$

For then the above inequalities yield the existence of $p > x^{\phi}$ with $N(p) = 1$. Obviously $N(n) = 0$ or 1 for $n > y$.

Thus we have reduced the question to an upper bound sieve problem. Let $v \in [U, x^{3/4}]$. Define θ by $v = x^{\theta}$ and let $\mathcal{K} = (v, ev]$, $\mathcal{A} = \{n: n \in \mathcal{K},\ N(n) = 1\}$ and $\mathcal{B} = \{n: n \in \mathcal{K}\}$. Thus $\mathcal{A}$ is our set to be sieved, and $\mathcal{B}$ is a "comparison set".

For a finite set $\mathcal{E}$ of natural numbers, we write $\mathcal{E}_d = \{n: dn \in \mathcal{E}\}$ and let $|\mathcal{E}|$ denote the cardinality of $\mathcal{E}$. We shall be concerned with the quantity

$$S(\mathcal{A}_d, z) = |\{n \in \mathcal{A}_d: Q(n) \geqslant z\}|$$

and its averages over d. In particular, $S(\theta) = S(\mathcal{A}, (\mathrm{e}v)^{1/2})$ is the number of primes in $\mathcal{A}$. It is not hard to see (compare, e.g. [1]) that (1.1) follows from the bound

$$\int_{0.6-\epsilon}^{\phi} \theta S(\theta)\,\mathrm{d}\theta < \frac{2yL}{5}, \tag{1.2}$$

which we shall establish in the following sections.

We close this section with a few remarks on notation. Throughout the chapter, we suppose that $x > C(\epsilon)$. We write $\eta = \exp(-3/\epsilon)$ and $J = [vy^{-1}x^{4\eta}]$. The quantity δ denotes $C\eta$, where C is an absolute constant, not necessarily the same at each occurrence. Constants implied by $\ll$, $\gg$ and $O_\epsilon(\ )$ depend at most on ϵ. Constants implied by $O(\ )$ are absolute. The notation $Y \asymp Z$ means $Y \ll Z \ll Y$, and $m \sim M$ stands for $M < m \leqslant 2M$. We reserve ℓ, m and n for natural number variables and p, q, r, s, t and u, possibly with suffices, for prime variables. Finally, let $\psi(\alpha) = \alpha - [\alpha] - 1/2$.

2. The Arithmetical Information

Our first lemma concerns the "type I" sums S_I associated to the problem,

$$S_I = \sum_{h\sim H} \sum_{n\sim N} \sum_{\substack{m\sim M \\ v<mn\leqslant ev}} b_n\, e\left(\frac{hx}{mn}\right).$$

Lemma 1. *Suppose that* $3/5 \leqslant \theta < 3/4 - \epsilon$, $1/2 \leqslant H \leqslant J$ *and* $|b_n| \leqslant 1$. *Then*

$$S_I \ll vx^{-6\eta} \tag{2.1}$$

provided that either

$$N \ll x^{2/5-\epsilon}, \tag{2.2}$$

or

$$v^6 x^{-13/4+\epsilon} \ll N \ll x^{1/2-\epsilon}. \tag{2.3}$$

Proof. For the case (2.2), see [12, Corollary 2 of Theorem 2]. The condition $v < mn \leqslant ev$ can be removed at the cost of a logarithmic factor; for more details see [6, Section 3.2], for example.

For the case (2.3), we apply Wu [17, Theorem 2], which is essentially an abstraction of a result of Rivat and Sargos [15].

Again, the condition $v < mn \leqslant ev$ can be removed at the cost of a logarithmic factor; this is done in [15], and the extra details can readily be incorporated into [17]. In the notation of [17], take $k = 4$, $\alpha = \gamma = -1$, $\beta = 1$, and replace (H, M, N, X) by (N, H, M, Hxv^{-1}). We then have the bound

$$S_I x^{-\eta} \ll ((Hxv^{-1})^{16} N^{52} H^{68} M^{60})^{1/80} + ((Hxv^{-1}) N^2 H^2 M^4)^{1/4} + NH + N(HM)^{1/2} + N^{1/2}HM + X^{-1/2}HMN.$$

Thus we have to verify that

$$(Hxv^{-1})^{16} N^{52} H^{68} M^{60} \ll v^{80} x^{-\delta}, \tag{2.4}$$

$$H(xv^{-1}) N^2 H^2 M^4 \ll v^4 x^{-\delta}, \tag{2.5}$$

$$NH \ll vx^{-\delta}, \tag{2.6}$$

$$N(HM)^{1/2} \ll vx^{-\delta}, \tag{2.7}$$

and

$$X^{-1/2}HMN \ll xv^{-\delta}. \tag{2.8}$$

The left-hand side of (2.4) is $\ll x^{\delta+8} v^{68} x^{-34} v^{60} N^{-8} \ll v^{80} x^{-\delta}$ from (2.3). The left-hand side of (2.5) is $\ll x^{\delta+1/2} v^6 x^{-1} N^{-2} \ll v^4 x^{-\delta}$ likewise. The left-hand side of (2.6) is $\ll vx^{-\delta}$ since $N < x^{1/2-\epsilon}$. The left-hand side of (2.7) is $\ll N^{1/2}(vx^{-1/2})^{1/2} v^{1/2} x^{\delta} \ll vx^{-\delta}$ likewise. Finally, the left-hand side of (2.8) is $\ll H^{1/2} x^{-1/2+\delta} v^{3/2} \ll x^{\delta-3/4} v^2 \ll vx^{-\delta}$ since $\theta < 3/4 - \epsilon$. This completes the proof. □

In order to state our results for type II sums

$$S_{II} = \sum_{h\sim H} \sum_{\substack{n\sim N \\ v<mn\leqslant ev}} \sum_{m\sim M} a_m b_n \, e\left(\frac{hx}{mn}\right),$$

we introduce some notation that is adapted from [12, 6]. We define ϕ_j by the following table:

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9
$\frac{a}{b}$	$\frac{3}{5}$	$\frac{11}{18}$	$\frac{35}{54}$	$\frac{2}{3}$	$\frac{90}{131}$	$\frac{226}{323}$	$\frac{547}{771}$	$\frac{23}{32}$	$\frac{1857}{2500}$
$\approx$	0.6	0.6111	0.6481	0.6667	0.687	0.6997	0.7095	0.7188	0.7428

In the above $\approx$ gives the decimal to four significant figures. Put $J_j = [\phi_j, \phi_{j+1})$. We then write $\mathcal{J}(\theta) = [\theta - 1/2 + \epsilon, \tau(\theta) - \epsilon]$, where $\tau(\theta)$ is given by the next table:

Interval	J_1	J_2	J_3	J_4	J_5	J_6	J_7
$\tau(\theta)$	$2 - 3\theta$	$\frac{1}{6}$	$\frac{9\theta - 3}{17}$	$\frac{12\theta - 5}{17}$	$\frac{55\theta - 25}{17}$	$\frac{59\theta - 28}{66}$	$\frac{245\theta - 119}{261}$

It is convenient to write $\mathcal{K}(\theta) = [2\theta - 1 + \epsilon, 3/2 - 2\theta - \epsilon]$ for $\theta < 5/8 - \epsilon$.

Lemma 2. *Suppose that $\phi_1 \leqslant \theta \leqslant \phi_8$. Then for $|a_m| \leqslant 1$, $|b_n| \leqslant 1$, $1/2 \leqslant H \leqslant J$, we have $S_{II} \ll vx^{-6\eta}$, provided that either*

$$x^{\theta-1/2+\epsilon} \ll M \ll x^{\tau(\theta)-\epsilon}, \tag{2.9}$$

or

$$\theta < 5/8 - \epsilon \quad \textit{and} \quad x^{2\theta-1+\epsilon} \ll M \ll x^{3/2-2\theta-\epsilon}. \tag{2.10}$$

Proof. Again, we may remove the condition $v < mn \leqslant ev$ at the cost of a logarithmic factor. The case (2.9) is covered in [1], with $\theta \in J_1$, and [12]. For the case (2.10), we appeal to Robert and Sargos [16, Theorem 1], with $X = Hxv^{-1}$. We obtain

$$S_{II}x^{-\eta} \ll HNM\left(\left(\frac{X}{NHM^2}\right)^{1/4} + \frac{1}{(HN)^{1/4}} + \frac{1}{M^{1/2}} + \frac{1}{X^{1/2}}\right).$$

We already dealt with the last term. Next,

$$HNM\left(\frac{X}{HNM^2}\right)^{1/4} \ll v^2x^{-1/2+\delta}\left(\frac{x}{v^2M}\right)^{1/4} = v^{3/2}x^{-1/4+\delta}M^{-1/4} \ll vx^{-\delta}$$

from (2.10). And, $HNM(HN)^{-1/4} \ll H^{3/4}v^{3/4}M^{1/4} \ll v^{3/2}x^{-3/8+\delta}M^{1/4} \ll vx^{-\delta}$ from (2.10). Finally, $HNM^{1/2} \ll v^2x^{-1/2+\delta}M^{-1/2} \ll vx^{-\delta}$ from (2.10). □

The key consequence of Lemma 2 is that (2.9) or (2.10) implies

$$\sum_{\substack{m\sim M\\ v<mn\leqslant ev}}\sum_{n\sim N} a_m b_n \left(\psi\left(\frac{x+y}{mn}\right) - \psi\left(\frac{x}{mn}\right)\right) \ll yx^{-5\eta}$$

and consequently

$$\sum_{\substack{m\sim M\\ mn\in\mathcal{A}}}\sum_{n\sim N} a_m b_n = y \sum_{\substack{m\sim M\\ mn\in\mathcal{B}}}\sum_{n\sim N} \frac{a_m b_n}{mn} + O(yx^{-5\eta});$$

compare [2, 6]. Similarly for Lemma 1.

3. The Alternative Sieve: Initial Stage

The sieve introduced in [5], and discussed at length in [6], was designated the "alternative sieve" in [2] – it is an alternative to the Rosser–Iwaniec sieve. In the present context, we write

$$S(\mathcal{B}_m, \lambda) = y \sum_{\substack{mn\in\mathcal{B}\\ Q(n)\geqslant\lambda}} \frac{1}{mn},$$

and compare this quantity with $S(\mathcal{A}_m, u)$. We can regard $S(\mathcal{B}_m, u)$ as "known".

Lemma 3. *We have*

$$S(\mathcal{B}_m, \lambda) = \omega\left(\frac{\log v/m}{\log\lambda}\right)\frac{y}{m\log\lambda}(1 + O_\epsilon(L^{-1}))$$

for $m \leqslant v^{1-\eta}$ and $x^\epsilon \leqslant \lambda \leqslant v/m$. Here $\omega(u)$ is Buchstab's function.

Proof. This is a slight variant of [2, Lemma 8]. □

Under the conditions of Lemma 1,

$$\sum_{n\sim N} b_n|\mathcal{A}_n| = y\sum_{n\sim N}\frac{b_n}{n} + O_\epsilon(yx^{-5\eta}) \tag{3.1}$$

(compare [2, Lemma 9]), and one can easily deduce that

$$\sum_{n\sim N}\sum_{\substack{d<x^\epsilon\\ p|d\Rightarrow p<x^\eta}} b_n\left||\mathcal{A}_{nd}| - \frac{y}{nd}\right| = O_\epsilon(yx^{-4\eta}),$$

provided that either

$$N \ll x^{2/5-2\epsilon} \tag{3.2}$$

or

$$v^6x^{-13/4+\epsilon} \ll N \ll x^{1/2-2\epsilon}. \tag{3.3}$$

Of course, if $v < x^{73/120-\epsilon}$, then $v^6x^{-13/4+\epsilon} < x^{2/5-2\epsilon}$, so that (3.1) holds whenever

$$\theta < 73/120 - \epsilon \quad \text{and} \quad N \ll x^{1/2-2\epsilon}. \tag{3.4}$$

Arguing just as in [2, Lemma 10], we now obtain

Lemma 4. *Suppose that one of* (3.2), (3.3) *or* (3.4) *holds. For every* $n\sim N$*, let* $0\leqslant a_n\leqslant 1$*, and* $a_n = 0$ *unless* $Q(n)\geqslant x^\eta$*. Then*

$$\sum_{n\sim N} a_nS(\mathcal{A}_n, x^\eta) = \sum_{n\sim N} a_nS(\mathcal{B}_n, x^\eta)\left(1 + O\left(\exp\left(-\frac{\epsilon}{\eta}\log\frac{\epsilon}{\eta}\right)\right)\right) + O_\epsilon(yx^{-4\eta}).$$

Here, and in succeeding lemmas, it is possible to attain a sharper error term on the right-hand side by following the arguments in [6, Chapter 3]. Since this would not improve our final result, we keep the exposition close to that of [2].

The other structural component of the alternative sieve, in the present application, is

Lemma 5. *Let* $\phi_1\leqslant\theta\leqslant\phi_8$*. Let* $h\geqslant 1$ *be given, and let* $\mathcal{D}\subset\{1,\dots,h\}$*. Let* $1\leqslant M\leqslant M_1$*,* $M_1\leqslant 2M$*, and suppose that one of the two conditions below holds:*

(i) *Either* $x^{\theta-1/2+\epsilon}\ll M\ll x^{\tau(\theta)-\epsilon}$ *or* $x^{\theta-\tau(\theta)+\epsilon}\ll M\ll x^{1/2-\epsilon}$.

(ii) $\theta < 5/8-\epsilon$*; and either* $x^{2\theta-1+\epsilon}\ll M\ll x^{3/2-2\theta-\epsilon}$ *or* $x^{3\theta-3/2+\epsilon}\ll M\ll x^{1-\theta-\epsilon}$.

Then

$$\sum_{p_1}\cdots\sum_{p_h}{}^{*} S(\mathcal{A}_{p_1\dots p_h}, p_1) = \sum_{p_1}\cdots\sum_{p_h}{}^{*} S(\mathcal{B}_{p_1\dots p_h}) + O(yx^{-4\eta}).$$

Here $*$ *in the summation indicates that* $p_1,\dots,p_h$ *satisfy* $x^\eta\leqslant p_1<\cdots<p_h$ *and* $M\leqslant\prod_{j\in\mathcal{D}}p_j<M_1$*, together with no more than* ϵ^{-1} *further conditions of the form*

$$R\leqslant\prod_{j\in\mathcal{F}}p_j\leqslant S. \tag{3.5}$$

Proof. This is proved in exactly the same way as [2, Lemma 12], using Lemma 2 of this chapter in place of [2, Lemma 11]. □

It is convenient to write, for $\phi_1 \leqslant \theta \leqslant \phi_8$, $a = vx^{-1/2-\epsilon}$ and $b = x^{\tau(\theta)-\epsilon}$, and for $\phi_1 \leqslant \theta < 5/8-\epsilon$, $c = v^2x^{-1+\epsilon}$ and $d = x^{3/2-2\theta-\epsilon}$. We also write $w = b/a$ and $w' = d/c$. Note that $w' = w$ for $\theta \leqslant 11/18$, while $w' < w$ for $11/18 < \theta \leqslant 5/8-\epsilon$.

Let $E = E(x,\epsilon)$ be some function of x and ϵ, with $0 < E(x,\epsilon) \leqslant \epsilon$. We can now write down some general conditions under which we have

$$\sum_m a_m S(\mathcal{A}_m, z) = \sum_m a_m S(\mathcal{B}_m, z)(1+O(E)) \tag{3.6}$$

for $x^\eta \leqslant z \leqslant w$; and some further conditions under which we have (3.6) for the range $x^\eta \leqslant z \leqslant w'$. When (3.6) holds, we say for brevity that $\sum_m a_m S(\mathcal{A}_m, z)$ *has an asymptotic formula with error* E.

Let us write

$$\gamma = \begin{cases} 1/2 - 2\epsilon & \text{if } \phi_1 \leqslant \theta < 73/120-\epsilon, \\ 2/5 - 2\epsilon & \text{if } 73/120-\epsilon \leqslant \theta \leqslant \phi_8. \end{cases}$$

Lemma 6. *Let $\phi_1 \leqslant \theta \leqslant \phi_8$. Let $1/2 \leqslant M \ll a$, $1/2 \leqslant N \ll x^\gamma a^{-1}$, $M \leqslant M_1 \leqslant 2M$, $N \leqslant N_1 \ll 2N$ and $x^\eta < z \leqslant w$. Suppose that $\{1,\dots,h\}$ partitions into two sets $\mathcal{C}$ and $\mathcal{D}$. Then*

$$\sum_{p_1} \cdots \sum_{p_h}{}^{*} S(\mathcal{A}_{p_1\dots p_h}, w)$$

has an asymptotic formula with error ϵ. Here $*$ in the summation indicates that $p_1,\cdots,p_h$ satisfy $z \leqslant p_1 < \cdots < p_h$, $M \leqslant \prod_{j\in\mathcal{C}} p_j < M_1$ and $N \leqslant \prod_{j\in\mathcal{D}} p_j < N_1$, together with no more than ϵ^{-1} further conditions of the form (3.5).

Proof. This is proved in exactly the same way as [2, Lemma 13], using Lemma 4 of this chapter in place of [2, Lemma 10]. □

Let $g = v^6x^{-13/4+\epsilon}$ and $\phi_1 \leqslant \theta < 37/60$. The significance of the number 37/60 is that $g < v/b$ holds for $\theta < 37/60$.

Lemma 7. *Let $\phi_1 \leqslant \theta \leqslant 5/8-\epsilon$. Let $P_1 \geqslant x^\eta,\dots,P_h \geqslant x^\eta$, and suppose either that*

$$P_1\dots P_h \ll \frac{v}{d}, \quad x^\eta \leqslant z \leqslant w', \tag{3.7}$$

or that $\theta < 37/60$ and

$$g \ll P_1\cdots P_h \ll \frac{v}{b}, \quad x^\eta \leqslant z \leqslant w, \tag{3.8}$$

where the condition $P_1\dots P_h \gg g$ can be deleted in (3.8) *if $\theta < 73/120-\epsilon$. Then*

$$\sum_{p_1\sim P_1} \cdots \sum_{p_h\sim P_h} S(\mathcal{A}_{p_1\dots p_h}, z) \tag{3.9}$$

has an asymptotic formula with error ϵ.

Proof. This is similar to that of [2, Lemma 13], so we shall be brief. We write

$$\sum_{\mathbf{p}} \quad \text{for} \quad \sum_{p_1\sim P_1} \cdots \sum_{p_h\sim P_h},$$

and $m = p_1\cdots p_h$. Suppose first that (3.7) holds. By Buchstab's identity,

$$\sum_{\mathbf{p}} S(\mathcal{A}_m, z) = \sum_{\mathbf{p}} S(\mathcal{A}_m, x^\eta) - \sum_{\mathbf{p}} \sum_{x^\eta \leqslant q_1 < z} S(\mathcal{A}_{mq_1}, q_1).$$

The first term on the right has an asymptotic formula with error

$$\exp\left(-\frac{\epsilon}{\eta}\log\frac{\epsilon}{\eta}\right), \tag{3.10}$$

by Lemma 4. The subsum of the second term on the right for which $mq_1 \geqslant v/d$ has an asymptotic formula with error $x^{-\eta}$ by Lemma 5, since $mq_1 \leqslant (v/d)z \leqslant v/c$. To the residual sum in which $mq_1 < v/d$, we apply Buchstab again. If we continue in this fashion, the j-th step is the identity

$$\begin{aligned}\sum\nolimits_j &= \sum_{\mathbf{p}}\sum_{(3.11)} S(\mathcal{A}_{mq_1\dots q_j}, q)\\ &= \sum_{\mathbf{p}}\sum_{(3.11)} S(\mathcal{A}_{mq_1\dots q_j}, x^\eta) - \sum_{\mathbf{p}}\sum_{(3.12)} S(\mathcal{A}_{mq_1\dots q_{j+1}}, q_{j+1})\end{aligned}$$

with summation conditions

$$x^\eta \leqslant q_j < \cdots < q_1 < z, \quad mq_1\cdots q_j < \frac{v}{d}, \tag{3.11}$$

$$x^\eta \leqslant q_{j+1} < q_j \cdots q_1 < z, \quad mq_1\cdots q_j < \frac{v}{d}. \tag{3.12}$$

The first of the subtracted pair of sums has an asymptotic formula with error (3.10) by Lemma 4, and the subsum of the second of the pair complementary to $\sum_{j+1}$ has an asymptotic formula with error $x^{-\eta}$, since

$$\frac{v}{d} \leqslant mq_1\cdots q_{j+1} < \left(\frac{v}{d}\right) q_{j+1} < \left(\frac{v}{d}\right) w = \frac{v}{c}. \tag{3.13}$$

The residual sum is $\sum_{j+1}$. After $O_\epsilon(1)$ steps the residual sum is empty, giving a decomposition of $\sum_{\mathbf{p}} S(\mathcal{A}_m, z)$ into a main term and an error term. A corresponding decomposition applies to $\sum_{\mathbf{p}} S(\mathcal{B}_m, z)$, and just as in the proof of [2, Lemma 13], (3.9) has an asymptotic formula with error

$$\eta^{-1}2^{1/\eta}\exp\left(-\frac{\epsilon}{\eta}\log\frac{\epsilon}{\eta}\right) < \epsilon.$$

This completes the proof of the case (3.7). The case (3.8) is very similar, with v/b and w in the roles of v/d and w': thus (3.13) is replaced by

$$\frac{v}{b} \leqslant mq_1\cdots q_{j+1} < \left(\frac{v}{b}\right) q_{j+1} < \left(\frac{v}{b}\right) w = \frac{v}{a}. \qquad \square$$

4. Assembling the Components of the Final Decomposition

For each θ, we shall in Section 5 make a "final decomposition" of $S(\mathcal{A}, (\mathrm{e}v)^{1/2})$ and a corresponding decomposition of $S(\mathcal{B}, (\mathrm{e}v)^{1/2})$, using Buchstab's identity and, in some cases, role reversals. Let us say this takes the form

$$S(\mathcal{A}, (\mathrm{e}v)^{1/2}) = \sum_{j=1}^{k} S_j - \sum_{j=k+1}^{\ell} S_j, \quad S(\mathcal{B}, (\mathrm{e}v)^{1/2}) = \sum_{j=1}^{k} S_j^* - \sum_{j=k+1}^{\ell} S_j^*.$$

Here $S_j \geqslant 0$, $S_j^* \geqslant 0$ and for $j \leqslant k$ and, say, $k+1 \leqslant j < t$, where $t \leqslant \ell$, we have $S_j = S_j^*(1+O(\epsilon))$. Thus we get the upper bound

$$S(\mathcal{A}, (\mathrm{e}v)^{1/2}) \leqslant \left(S(\mathcal{B}, (\mathrm{e}v)^{1/2}) + \sum_{j=t}^{\ell} S_j^* \right) (1+O(\epsilon)).$$

We strive to make the "discarded sums" S_j, with $t \leqslant k \leqslant \ell$, as small as possible, thinking of them as regions in euclidean spaces.

The first step is

$$S(\mathcal{A}, (\mathrm{e}v)^{1/2}) = S(\mathcal{A}, w) - \sum_{w \leqslant p < (\mathrm{e}v)^{1/2}} S(\mathcal{A}_p, p).$$

To continue the process for $p \in I$, an interval where no asymptotic formula is available for

$$\sum_{p \in I} S(\mathcal{A}_p, p), \tag{4.1}$$

we need to give asymptotic formulae for

$$\sum_{p \in I} S(\mathcal{A}_p, w) \quad \text{and} \quad \sum_{p \in I} \sum_{w \leqslant q < p} S(\mathcal{A}_{pq}, w^*),$$

where $w^* = w$ or w', depending on p, q. If this cannot be done, we discard the sum (4.1). These remarks should give context to the lemmas in the present section.

Lemma 8. *Let $\theta \leqslant 0.65 - \epsilon$ and $P < b^2$. Then*

$$\sum_{p \sim P} S(\mathcal{A}_p, w)$$

has an asymptotic formula.

Proof. See [6, Lemma 6.7]. □

Lemma 9. *Let $\phi_1 \leqslant \theta \leqslant 5/8 - \epsilon$, $w \leqslant Q \leqslant P \leqslant (\mathrm{e}v)^{1/2}$, and suppose that $PQ^2 \ll v$, $P, Q \notin [a,b] \cup [c,d]$ and $PQ \notin [vd^{-1}, vc^{-1}] \cup [vb^{-1}, va^{-1}]$. Suppose further that either*

(i) $\theta < 73/120 - \epsilon$; *or*
(ii) $P \leqslant x^{1/2-\epsilon} v^{-1/2}$.

Then

$$\sum_{p \sim P} \sum_{q \sim Q} S(\mathcal{A}_{pq}, w')$$

has an asymptotic formula.

Proof. (i) If $Q < a$, we can apply Lemma 6, since $P \leqslant (\mathrm{e}v)^{1/2} < x^{1/2}/a$. Thus we may suppose that $Q > b$. We cannot have $P > d$, since $b^2 d = x^{11/2 - 8\theta - 3\epsilon} > vx^{\epsilon}$. Thus we have $Q \leqslant P < c$. Accordingly, $PQ < c^2 < x^{1/2-\epsilon} = va^{-1}$. Hence we have $PQ < vb^{-1}$, and the result follows from Lemma 7.

(ii) We have $PQ \leqslant x^{1-2\epsilon} v^{-1} < v/c$. Hence $PQ < v/d$, and we may apply Lemma 7. □

Lemma 10. *Suppose that* $\theta \in [73/120 - \epsilon, 47/75 - \epsilon]$, $evb^{-2} < P < (ev)^{1/2}$, *and either* $P \leqslant x^{2/5-2\epsilon}a^{-1}$, *or* $\theta < 11/18 - \epsilon$ *and* $P > x^{2/5-2\epsilon}a^{-1}$. *Then*

$$\sum_{p \sim P} S(\mathcal{A}_p, p) \geqslant \sum_{p \sim P} S(\mathcal{B}_p, p)(1 + O(\epsilon)) - S_{\nabla}, \tag{4.2}$$

where S_{∇} *and* ∇ *are defined as follows:*

(i) *For* $P \leqslant x^{2/5-2\epsilon}a^{-1}$, *we have*

$$S_{\nabla} = \sum_{\nabla} S(\mathcal{B}_{pqr}, r), \tag{4.3}$$

where ∇ *is the set of conditions*

$$p \sim P, \quad w \leqslant r < q < a, \quad q < \left(\frac{ev}{p}\right)^{1/2}, \quad r < \left(\frac{ev}{pq}\right)^{1/2},$$

it being understood that no combination of the variables p, q, r *satisfies the requirements of Lemma 5.*

(ii) *For* $\theta < 11/18 - \epsilon$ *and* $P > x^{2/5-2\epsilon}a^{-1}$, *we have*

$$S_{\nabla} = \sum_{\nabla} S(\mathcal{B}_{mqu}, u), \tag{4.4}$$

where ∇ *is the set of conditions*

$$mq \sim P, \quad w \leqslant q < a, \quad Q(m) \geqslant q, \quad w \leqslant u < \left(\frac{ev}{P}\right)^{1/2},$$

it being understood that no combination of the variables m, q, u *satisfies the requirements of Lemma 5.*

Proof. This is essentially [6, Lemma 6.8], using Lemma 5 in place of the corresponding result in [6]. □

The role reversal used in the second part of Lemma 10 does not yield useful results if we extend it beyond $\theta = \phi_2$. We now treat a role reversal for

$$\sum_{p \sim P} S(\mathcal{A}_p, p), \tag{4.5}$$

where we assume that $73/120 - \epsilon \leqslant \theta < \phi_2$ and $b < P < vg^{-1}$. Besides primes q with $pq \in \mathcal{A}$, the above sum counts $pq_1q_2 \in \mathcal{A}$ with $p \leqslant q_1 \leqslant q_2$. The dependence of q_1 on p suggests we first show that

$$\sum_{p \sim P} S(\mathcal{A}_p, p) = \sum_{p \sim P} S(\mathcal{A}_p, P) + O\left(L^{-1} \sum_{p \sim P} S(\mathcal{B}_p, p)\right).$$

Clearly, it suffices to show that

$$\sum_{\substack{pq_1q_2 \in \mathcal{A} \\ p \sim P \\ q_1 \leqslant q_2 \\ P \leqslant q_1 < p}} 1 = O\left(L^{-1} \sum_{p \sim P} S(\mathcal{B}_p, p)\right). \tag{4.6}$$

The left-hand side of (4.6) is

$$\leqslant \sum_{\substack{p\sim P\\ q_2 \asymp vP^{-2}}} S(\mathcal{A}_{pq_2}, P) = O\left(\sum_{\substack{p\sim P\\ q_2 \asymp vP^{-2}}} \frac{y}{pq_2 L}\right) = O(yL^{-3})$$

from Lemma 4, since $vP^{-1} > g$. This implies (4.6).

We now proceed as in [6, Section 6.6]. The sum (4.5) is, with acceptable error,

$$\{p\ell : p\ell \in \mathcal{A},\ p\sim P,\ Q(\ell) \geqslant P\} = \sum_{\substack{\ell \asymp v/P\\ Q(\ell)\geqslant P}} S(\mathcal{A}(\ell), (2P)^{1/2}),$$

where $\mathcal{A}(\ell) = \{m\colon m\sim P,\ m\ell \in \mathcal{A}\}$. We rewrite the sum over ℓ as

$$\sum_{\substack{\ell \asymp v/P\\ Q(\ell)\geqslant P}} S(\mathcal{A}(\ell), w) - \sum_{\substack{\ell\\ w\leqslant q<(2P)^{1/2}}} S(\mathcal{A}(\ell)_q, q).$$

The first of the subtracted pair of sums has an asymptotic formula by Lemma 7, since $v/P > g$. For the second sum, we note that $q < x^{13/8+\epsilon}v^{-5/2} < a$, since $\theta > 17/28$. We reverse roles again, so that

$$\sum_{\substack{\ell\\ w\leqslant q<(2P)^{1/2}}} S(\mathcal{A}(\ell)_q, q) = \sum_{\substack{mq\sim P\\ w\leqslant q<(2P)^{1/2}\\ Q(m)\geqslant q}} S\left(\mathcal{A}_{mq}, \left(\frac{\mathrm{e}v}{P}\right)^{1/2}\right).$$

Since $m \ll x^{3/4+3\epsilon}v^{-1} < b$, we can restrict attention to $m < a$ in the last expression. Now m is prime, since $w^2 > a$; write $m = r$. Apply Buchstab once more, so that

$$\sum_{\substack{rq\sim P\\ w\leqslant q<r<a}} S\left(\mathcal{A}_{rq}, \left(\frac{\mathrm{e}v}{P}\right)^{1/2}\right) = \sum_{\substack{rq\sim P\\ w\leqslant q<r<a}} S(\mathcal{A}_{rq}, w) - \sum_{\substack{rq\sim P\\ w\leqslant q<r<a\\ w\leqslant u<(\mathrm{e}v/P)^{1/2}}} S(\mathcal{A}_{rqu}, u).$$

The first sum on the right-hand side satisfies the requirements of Lemma 6. We discard those parts of the second sum for which we cannot give an asymptotic formula by Lemma 5. This establishes the first part of the following result.

Lemma 11. *Suppose that* $73/120 - \epsilon \leqslant \theta < 13/21$ *and either*

(i) $\theta < 11/18$ *and* $b < P < vg^{-1}$*; or*

(ii) $P \leqslant x^{1/2-\epsilon}v^{-1/2}$.

Then (4.2) *holds, and corresponding to the two cases above,* S_∇ *and* ∇ *are defined as follows:*

(i) *We have*

$$S_\nabla = \sum_{\nabla} S(\mathcal{B}_{rqu}, u),$$

where ∇ *is the set of conditions*

$$rq\sim P, \quad w\leqslant q<r<a, \quad w\leqslant u<\left(\frac{\mathrm{e}v}{P}\right)^{1/2},$$

it being understood that no combination of variables r, q, u *satisfies the conditions of Lemma 5.*

(ii) *We have*

$$S_\nabla = \sum_{\nabla} S(\mathcal{B}_{pqr}, r),$$

where ∇ *is the set of conditions*

$$p \sim P, \quad w'' \leqslant r < q < p, \quad r < \left(\frac{\mathrm{e}v}{pq}\right)^{1/2}, \tag{4.7}$$

it being understood that no combination of variables p, q, r *satisfies the conditions of Lemma 5, and* $w'' = w$ *or* w' *depending on whether* $q < a$ *or* $q > b$.

For the second part of the lemma, we apply Buchstab twice to the sum (4.5), taking into account Lemma 9(ii). We then discard the part of $\sum_{(4.7)} S(\mathcal{A}_{pqr}, r)$ to which Lemma 5 does not apply.

5. Completion of the Proof of the Theorem

As noted in the introduction, our treatment is just as in [6, Section 6.7] for $\theta \geqslant 5/8 - \epsilon$. For the moment, suppose that $\phi_1 \leqslant \theta < 5/8 - \epsilon$. We begin our final decomposition with

$$\begin{aligned} S(\mathcal{A}, (\mathrm{e}v)^{1/2}) &= S(\mathcal{A}, w) - \sum_{w \leqslant p < a} S(\mathcal{A}_p, p) - \sum_{p \in [a,b] \cup [c,d]} S(\mathcal{A}_p, p) \\ &\quad - \sum_{\substack{b < p < (\mathrm{e}v)^{1/2} \\ p \notin [c,d]}} S(\mathcal{A}_p, p) \\ &= S_1 - S_2 - S_3 - S_4, \end{aligned} \tag{5.1}$$

say. We have asymptotic formulae for S_1 and S_3. The treatment of S_2 and S_4 raises several questions, the answers depending on θ.

(i) Is there an interval of p within S_2 for which an asymptotic formula holds?

(ii) For the rest of S_2 and S_4, which intervals I of p permit two further decompositions, in the sense

$$\begin{aligned} \sum_{p \in I} S(\mathcal{A}_p, p) &= \sum_{p \in I} S(\mathcal{A}_p, w) - \sum_{\substack{p \in I \\ q \in J_p \\ w \leqslant q < p}} S(\mathcal{A}_{pq}, w) + \sum_{\substack{p \in I \\ q \in J_p \\ w \leqslant r < q < p}} S(\mathcal{A}_{pqr}, r) \\ &\quad - \sum_{\substack{p \in I \\ q \notin J_p \\ w' \leqslant q < p}} S(\mathcal{A}_{pq}, w') + \sum_{\substack{p \in I \\ q \notin J_p \\ w' \leqslant r < q < p}} S(\mathcal{A}_{pqr}, r) \\ &= S_5 - S_6 + S_7 - S_8 + S_9, \end{aligned} \tag{5.2}$$

say?

(iii) In S_7 and S_9, which portions permit two more decompositions to obtain sums

$$\sum S(\mathcal{A}_{pqr}, w_1) - \sum S(\mathcal{A}_{pqrs}, w_2) + \sum S(\mathcal{A}_{pqrst}, t) = S_{10} - S_{11} + S_{12}, \tag{5.3}$$

say? How do we choose w_1 and w_2 according to the region in which (p, q, r), (p, q, r, s) lie?

(iv) Are there further intervals of p in which a role reversal in

$$\sum_{p \in I} S(\mathcal{A}_p, p) \tag{5.4}$$

is to be preferred to discarding the sum in (5.4)?

(v) Can a small part of (5.4) be recovered, rather than discarding all of it?

Of course, a decomposition terminates if some combination of variables allows us to apply Lemma 5. For example, we do not decompose further the portion of S_7 with $pqr \in [c, d]$. There are seven Buchstab decompositions in some cases; these will be noted below.

We write $T(\theta) = (L/y)S(\theta)$. For simplicity of writing, we ignore any terms in the construction of an upper bound for $T(\theta)$ which are $O(\epsilon)$.

We now provide answers to (i)–(v) above.

(i) For $\theta < 11/18$, we have $w^2 > a$ and, as in [6], there is an asymptotic formula for the part S_2' of S_2 with $p < b^{1/2}$. For $\theta \geqslant 11/18$ there is nothing corresponding to S_2'.

(ii)(a) For $\theta < 73/120$, (5.2) is applied for $I = (b^{1/2}, a)$, (b, c), $(d, v^{1/2})$, and J_p consists of (w, a). For $p < v^{1/2}$ implies $p < x^{1/2}a^{-1}$. If $q > b$, we place (p, q) in S_8, since $pq^2 < v$ implies $pq < vb^{-1}$.

(b) For $\theta \in (73/120, 11/18)$, (5.2) is applied for $p < (x/v)^{1/2}$ and $d < p < x^{9/10-\theta}$. Note that if $p > d$, then $q < b$ since $db^2 > v$; in this case, $J_p = (w, a)$. In fact, since $(x/v)^{1/2} < x^{9/10-\theta}$, $J_p = (w, a)$ for $p < (x/v)^{1/2}$; while for $p < (x/v)^{1/2}$, $q > a$, we may place (p, q) in S_8 since $pq < x/v$ implies $pq < v/c$.

(c) For $\theta \in (11/18, 13/21)$, apply (5.2) for $p < (x/v)^{1/2}$, with $J_p = (w, a)$, arguing as in the last paragraph.

(d) For $\theta \in (13/21, 5/8)$, apply (5.2) for $p < a$, with $J_p = (w, a)$. The point here is that w' is too small for numerical results arising from Lemma 7 to be helpful.

(iii) We carry out two more decompositions if either

(a) p, q, r, r can be combined into two products m, n with $m < x^{\gamma}a^{-1}$ and $n < a$; or

(b) $pqr^2 < v/c$; or

(c) $\theta < 73/120$ and $pqr^2 < x^{1/2}$; or

(d) $73/120 < \theta < 37/60$, and (†) p, q, r can be combined into two products m, n with $m < x^{9/10-\theta}$ and $n < a$, or else $pqr < v/d$; and (‡) $pqrw > g$ and $pqr^2 < x^{1/2}$.

If (a) is satisfied, we apply Lemma 6 to S_{10} and S_{11}. If (b) is satisfied, then we apply Lemma 7 to S_{11}, and to S_{10} we can definitely apply Lemma 7 and may be able to apply Lemma 6. If (c) is satisfied, we can apply Lemma 6 or 7 to S_{10} and S_{11}. Since for $w < r < s$, (‡) implies $g < pqrs < x^{1/2}$, we can apply Lemma 7 in case (d) to S_{11}, and either Lemma 6 or 7 applies to S_{10}. It is clear that, for $\theta > 11/18$, we always apply Lemma 6 in preference to Lemma 7 if the necessary hypotheses are fulfilled, and this determines w_1 and w_2.

(iv)(a) A role reversal based on Lemma 10(ii) is used for $73/120 < \theta < 11/18$ and $p > x^{9/10}a^{-1}$.

(b) A role reversal based on Lemma 11 is used for $73/120 < \theta < 11/18$ and $p \in ((x/v)^{1/2}, vg^{-1})$. The latter interval disappears for $\theta > 11/18$.

(v) According to (i)–(iv), $S(\mathcal{A}_p, p)$ is discarded for $vg^{-1} < p < c$ if $\theta \in (73/120, 11/18)$, for $p \in (x^{1/2}/v, c)$ and $p \in (d, v^{1/2})$ if $11/18 < \theta < 13/21$, and for $p \in (b, c)$ and $p \in (d, v^{1/2})$ if $\theta > 13/21$. If $\theta \in (13/21, 5/8)$, then $b^2c < v$ by a generous margin. From the discarded terms $S(\mathcal{A}_p, p)$, we can recover those pqr in $\mathcal{A}$ with $b < p < q < r$ and $c < r < d$.

Seven-dimensional integrals arise for $\theta \in (17/28, 5/8)$, since then $w^6 < v/b$, and it can happen that in S_{12}, p, q, r, s, t, u permit a treatment similar to that in (iii) whenever $w < u < t$: for θ close to $17/28$, this would depend on the inequalities $pqrstu < pqrst^2$ and $pqrstu > pqrstw$.

The above discussion will enable the reader to write down the multidimensional integrals $I_1, \ldots, I_h$ such that

$$T(\theta) \leqslant \frac{1}{\theta} + I_1 + \cdots + I_h. \tag{5.5}$$

In the case $\theta \in (13/21, 5/8)$, there is a further integral arising from (v), namely

$$I_{h+1} = \int_{1/6}^{(1-\theta)/2} \int_{\max\{3\theta-3/2-\alpha,\alpha\}}^{1-\theta-\alpha} \frac{\mathrm{d}\beta}{\beta^2} \frac{\mathrm{d}\alpha}{\alpha^2}$$

such that

$$T(\theta) \leqslant \frac{1}{\theta} + I_1 + \cdots + I_h - I_{h+1}. \tag{5.6}$$

For integrals similar to $I_1, \ldots, I_h$, see the discussion in [6, Section 6.7].

The conclusions that we obtain from (5.5) and (5.6) are as follows. We have

$$\int_{\phi_1}^{\phi_2} \theta T(\theta)\,\mathrm{d}\theta < 0.01153; \tag{5.7}$$

note how close this is to the conjectural value $0.01111\ldots$. Further,

$$\int_{\phi_2}^{\phi_4} \theta T(\theta)\,\mathrm{d}\theta < 0.12455. \tag{5.8}$$

Of course, the saving in (5.8) compared with [6] comes only from $\phi_2 < \theta < 5/8$.

Just as in [6],

$$\int_{\phi_4}^{\phi_8} \theta T(\theta)\,\mathrm{d}\theta < 0.17597, \tag{5.9}$$

and

$$\int_{\phi_8}^{\phi_9} \theta T(\theta)\,\mathrm{d}\theta < \frac{5}{2}(\phi_9^2 - \phi_8^2) < 0.088. \tag{5.10}$$

We may combine (5.7)–(5.10) to give (1.2). This completes the proof of the theorem.

References

[1] R.C. Baker. The greatest prime factor of the integers in an interval. *Acta Arith.*, 47 (1986), 193–231.

[2] R.C. Baker, G. Harman. Numbers with a large prime factor. *Acta Arith.*, 73 (1995), 119–145.

[3] J.B. Friedlander, H. Iwaniec. Exceptional characters and prime numbers in short intervals. *Selecta Math. (N.S.)*, 10 (2004), 61–69.

[4] S.W. Graham. The greatest prime factor of the integers in an interval. *J. London Math. Soc.*, 24 (1981), 427–440.

[5] G. Harman. On the distribution of αp modulo one. *J. London Math. Soc.*, 27 (1983), 9–13.

[6] G. Harman. *Prime-detecting sieves* (London Mathematical Society Monographs 33, Princeton University Press, 2007).

[7] C.H. Jia. The greatest prime factor of integers in short intervals I. *Acta Math. Sinica*, 29 (1986), 815–825.

[8] C.H. Jia. The greatest prime factor of integers in short intervals II. *Acta Math. Sinica*, 32 (1989), 188–199.
[9] C.H. Jia. The greatest prime factor of integers in short intervals III. *Acta Math. Sinica (N.S.)*, 9 (1993), 321–336.
[10] C.H. Jia. The greatest prime factor of integers in short intervals IV. *Acta Math. Sinica (N.S.)*, 12 (1996), 433–445.
[11] H-Q. Liu. The greatest prime factor of the integers in an interval. *Acta Arith.*, 65 (1993), 301–328.
[12] H-Q. Liu, J. Wu. Numbers with a large prime factor. *Acta Arith.*, 89 (1999), 163–187.
[13] K. Ramachandra. A note on numbers with a large prime factor. *J. London Math. Soc.*, 1 (1969), 303–306.
[14] K. Ramachandra. A note on numbers with a large prime factor II. *J. Indian Math. Soc.*, 34 (1970), 39–48.
[15] J. Rivat, P. Sargos. Nombres premiers de la forme $[n^c]$. *Canad. J. Math.*, 53 (2001), 414–433.
[16] O. Robert, P. Sargos. Three-dimensional exponential sums with monomials. *J. reine angew. Math.*, 591 (2006), 1–20.
[17] J. Wu. On the primitive circle problem. *Monatsh. Math.*, 135 (2002), 69–81.