ON THE DISTRIBUTION OF GAPS BETWEEN SQUAREFREE NUMBERS

Michael Filaseta*

1. Introduction

Let s_1, $s_2, \ldots$ denote the squarefree numbers in ascending order. In [1], Erdős showed that, if $0 \le \gamma \le 2$, then

$$\sum_{s_{n+1} \le x} (s_{n+1} - s_n)^\gamma \sim B(\gamma)x, \tag{1}$$

where $B(\gamma)$ is a function only of γ. In 1973 Hooley [4] improved the range of validity of this result to $0 \le \gamma \le 3$ and then later gained a further slight improvement by a method he outlined at the International Number Theory Symposium at Stillwater, Oklahoma in 1984. We have, however, independently obtained the better improvement that (1) holds for

$$0 \le \gamma < 29/9 = 3.222\ldots$$

in contrast to the range

$$0 \le \gamma \le 250/79 = 3.16\ldots$$

derived by Hooley. The main purpose of this paper is to substantiate our new result. Professor Hooley has informed me that there are similarities between our methods as well as significant differences.

To obtain our result, we will consider instead the problem of showing that for $0 \le \gamma < 29/9$

$$\sum_{x/2 < s_{n+1} \le x} (s_{n+1} - s_n)^\gamma \sim B'(\gamma)x, \tag{2}$$

where $B'(\gamma)$ is a function of γ. We note that (1) easily follows from (2) by breaking up the sum in (1) into sums of the form given in (2) (and with $B(\gamma) = 2B'(\gamma)$). The reason for considering (2) instead of (1) will become clear in Section 5.

We end the paper (see Section 7) by showing the equivalence of the following 2 conjectures:

Conjecture 1. *For every $\gamma > 0$, there is a $B(\gamma)$ such that the asymptotic formula given by (1) holds.*

Conjecture 2. *For every $\epsilon > 0$, there is an $x_0 = x_0(\epsilon)$ such that if $x \ge x_0$ then there is a squarefree number in the interval $(x, x + x^\epsilon]$.*

A result of Trifonov and the author [3] shows that the latter conjecture holds when one restricts one's attention to $\epsilon > 1/5$. A history of similar results can be found in [2].

---

*Research was supported in part by the NSF under grant number DMS-8903123.

2. Hooley's Result Revisited

In this section, we give an alternative approach to obtaining Hooley's result mentioned above and give some preliminaries to obtaining our improvement. We assume throughout the remainder of the paper that x is sufficiently large. For each positive integer t, we define N_t as the number of positive integers n such that $x/2 < s_{n+1} \le x$ and $s_{n+1} - s_n = t$. Thus, (2) becomes

$$\sum_{t=1}^{\infty} N_t t^{\gamma} \sim B'(\gamma)x. \tag{3}$$

Let $S_2(x,X,T)$ denote the number of 4−tuples (p_1,p_2,k_1,k_2) with p_1 and p_2 primes and k_1 and k_2 positive integers satisfying $k_2p_2^2 > k_1p_1^2$, $X < p_j \le 2X$ for each $j \in \{1,2\}$, and for some $I \subseteq [0,x]$ with $|I| \le 2T$, $k_jp_j^2 \in I$ for each $j \in \{1,2\}$. Note that if $S_2(x,X,T) > 0$, then $X \le \sqrt{x}$. We only consider such X. The bounds given for $S(x,X,T)$ as defined in Hooley [4] are bounds for $S_2(x,X,T)$ as defined here. In particular, it follows from [4] that

$$S_2(x,X,T) \ll \frac{x}{X}. \tag{4}$$

Similarly, we define $S_3(x,X,T)$ as the number of 6−tuples $(p_1,p_2,p_3,k_1,k_2,k_3)$ with p_1, p_2, and p_3 primes and k_1, k_2, and k_3 positive integers satisfying $k_3p_3^2 > k_2p_2^2 > k_1p_1^2$, $X < p_j \le 2X$ for each $j \in \{1,2,3\}$, and for some $I \subseteq [0,x]$ with $|I| \le 2T$, $k_jp_j^2 \in I$ for each $j \in \{1,2,3\}$.

Let r be the greatest integer $\le \log x$. Then $2^{r-1} \ge \sqrt{x}$. For a given T sufficiently large, set $X_j = 2^{j-1}(T/10)\log T$ for $j \in \{1,\dots,r\}$. Thus, if p^2 divides an integer $\le x$, then either $p \in (X_j, 2X_j]$ for some $j \in \{1,\dots,r\}$ or $p \le (T/10)\log T$. Observe that in a given gap of length $t \in (T,2T]$ between squarefree numbers, the number of integers divisible by some p^2 with $p \le (1/10)T\log T$ is

$$\le \sum_{p \le (1/10)T\log T} \left(\frac{t}{p^2}+1\right) < t\left(\sum_{k=2}^{\infty}\frac{1}{k^2}\right) + \pi\left((1/10)T\log T\right) < \frac{3}{4}t \tag{5}$$

provided that $T \ge T_0$ for some sufficiently large constant T_0. Prior to Section 6 in this paper, we shall consider only T sufficiently large so that we can assume the bound in (5) holds. Thus, each gap of size $t \in (T,2T]$ between consecutive squarefree numbers contains at least $\frac{1}{4}t > \frac{1}{4}T$ integers which are divisible by some p^2 with $p > X_1$. Suppose I is such a gap (so $I = (s_n, s_{n+1})$ for some n). Then $|I| \in (T,2T]$. For $j \in \{1,2,\dots,r\}$, define

$$Y_j = Y_j(I) = \left|\{m \in I : \text{ there exists a prime } p \in (X_j,2X_j] \text{ such that } p^2|m\}\right|.$$

Observe that

$$Y_1 + Y_2 + \cdots + Y_r \ge \frac{1}{4}T.$$

Let $\epsilon \in (0, 1/5)$, and let

$$C = C(\epsilon) = \sum_{j=1}^{\infty} (1+\epsilon)^{-j} = \frac{1}{\epsilon}.$$

Then there is a $j = j(I) = j(I, \epsilon) \in \{1, \ldots, r\}$ such that

$$Y_j \geq \frac{1}{4C}(1+\epsilon)^{-j} T. \tag{6}$$

We now turn the situation around. We begin with a fixed $j \in \{1, \ldots, r\}$ and estimate the number of intervals $I = (s_n, s_{n+1})$ with $x/2 < s_{n+1} \leq x$, $|I| \in (T, 2T]$, and $j(I) = j$. First, we consider the number of such I with $1 \leq Y_j \leq 2$. For such I, we get from (6) that

$$T \leq 8C(1+\epsilon)^j.$$

For each $p \in (X_j, 2X_j]$, there are $< x/X_j^2$ positive integral multiples of p^2 which are $\leq x$. Since there are clearly $\leq X_j$ primes $p \in (X_j, 2X_j]$, the number of intervals I containing a multiple of p^2 with $p \in (X_j, 2X_j]$ is

$$\leq \frac{x}{X_j^2} X_j = \frac{x}{X_j}.$$

This is easily an upper bound for the number of I with $1 \leq Y_j \leq 2$. Let N_t' denote the number of intervals $I = (s_n, s_{n+1})$ such that $x/2 < s_{n+1} \leq x$, $|I| = t$, and $1 \leq Y_{j(I)} \leq 2$. Then for any $\gamma < 4$ and for any $\epsilon \in (0, 1/5)$, the above implies that

$$\begin{aligned}\sum_{T<t\leq 2T} N_t' t^\gamma &\leq \sum_{j=1}^{r} \frac{x}{X_j} T^\gamma \ll \sum_{j=1}^{r} \frac{x}{2^j T} \frac{T^3}{T^{3-\gamma}} \\ &\ll_\epsilon \sum_{j=1}^{r} \frac{x}{T^{4-\gamma}} \left(\frac{(1+\epsilon)^3}{2} \right)^j \ll_\epsilon \frac{x}{T^{4-\gamma}}.\end{aligned}$$

Hence, one easily gets that

$$\sum_{t>T_0} N_t' t^\gamma \ll_\epsilon \frac{x}{T_0^{4-\gamma}}. \tag{7}$$

Let $N_t'' = N_t - N_t'$. To obtain an upper bound on $\sum_{t>T_0} N_t t^\gamma$, we will use (7) together with a bound on $\sum_{t>T_0} N_t'' t^\gamma$.

Observe that N_t'' is the number of intervals $I = (s_n, s_{n+1})$ such that $x/2 < s_{n+1} \leq x$, $|I| = t$, and $Y_{j(I)} \geq 3$. By (6), each such I contributes at least $\binom{Y_j}{2} \gg_\epsilon (1+\epsilon)^{-2j} T^2$ of the 4−tuples counted by $S_2(x, X_j, T)$ and at least $\binom{Y_j}{3} \gg_\epsilon (1+\epsilon)^{-3j} T^3$ of the 6−tuples counted by $S_3(x, X_j, T)$. Thus, the total number of intervals I as above is

$$\ll_\epsilon \min \left\{ \frac{S_2(x, X_j, T)}{T^2}(1+\epsilon)^{2j}, \frac{S_3(x, X_j, T)}{T^3}(1+\epsilon)^{3j} \right\}.$$

This now implies that

$$(8) \qquad \sum_{T<t\le 2T} N_t'' \ll_\epsilon \sum_{j=1}^{r} \min\left\{ \frac{S_2(x,X_j,T)}{T^2}(1+\epsilon)^{2j}, \frac{S_3(x,X_j,T)}{T^3}(1+\epsilon)^{3j} \right\}.$$

Observe that (4) implies that

$$\sum_{j=1}^{r} \frac{S_2(x,X_j,T)}{T^2}\left(\frac{11}{10}\right)^{2j} \ll \sum_{j=1}^{r} \frac{x}{2^j T^3 \log T}\left(\frac{11}{10}\right)^{2j} \ll \frac{x}{T^3 \log T}$$

so that by (8) with $\epsilon = 1/10$

$$\sum_{T<t\le 2T} N_t'' t^\gamma \ll \frac{x}{T^{3-\gamma}\log T}.$$

For $\gamma < 3$, this implies that (3) holds (see section 5 of [4]). Hooley [4] had slightly better estimates for $S_2(x,X,T)$ than that given by (4) (see (10) below) which enabled him to also establish (3) when $\gamma = 3$.

3. Estimates for $S_j(x,X,T)$

In this section, we give upper bounds for $S_2(x,X,T)$ and $S_3(x,X,T)$. Our estimate for $S_2(x,X,T)$ is contained in [4]; we include its proof to help explain our approach to estimating $S_3(x,X,T)$.

We begin with $S_2(x,X,T)$. Fix X and T with $X \ge (T/10)\log T$. Let p and $p+a$ be fixed primes (with a possibly ≤ 0) such that $X < p \le 2X$ and $X < p+a \le 2X$. We estimate the number of positive integers k_1 and k_2 satisfying $k_2(p+a)^2 > k_1p^2$ and for some $I \subseteq [0,x]$ with $|I| \le 2T$, k_1p^2 and $k_2(p+a)^2$ are in I. First, we show that $a \ne 0$. If $a = 0$, then since k_1p^2 and k_2p^2 are in I and $X \ge (T/10)\log T$, we get that $|k_2 - k_1| \le 2T/X^2 < 1$ so that $k_2 = k_1$. But this contradicts that $k_2(p+a)^2 > k_1p^2$. Hence, $a \ne 0$. Since p and $p+a$ are primes, we get that p and $p+a$ are relatively prime. Now,

$$(9) \qquad 0 < k_2(p+a)^2 - k_1p^2 \le 2T$$

implies that there are $\le 2T$ choices for k_2 modulo p^2. Since $k_2(p+a)^2 \le x$, we get that $k_2 \le x/X^2$. Thus, the total number of choices for k_2 is

$$\ll \left(\frac{x/X^2}{X^2} + 1\right)T = \left(\frac{x}{X^4}+1\right)T.$$

Also, since $p^2 > X^2 > 2T$, we get from (9) that k_1 is uniquely determined by k_2. The above all holds for p and $p+a$ fixed. There are $\le (\pi(2X) - \pi(X))^2 \ll X^2/\log^2 X$ choices for the pair $(p, p+a)$; hence,

$$(10) \qquad S_2(x,X,T) \ll \left(\frac{x}{X^4}+1\right)T \times \frac{X^2}{\log^2 X} = \frac{xT}{X^2\log^2 X} + \frac{TX^2}{\log^2 X}.$$

We now estimate $S_3(x, X, T)$. Again, we fix X and T with $X \geq (T/10)\log T$. Let p, $p+a$, and $p+a+b$ be fixed primes in $(X, 2X]$. We estimate the number of positive integers k_1, k_2, and k_3 satisfying $k_3p_3^2 > k_2p_2^2 > k_1p_1^2$ and for some $I \subseteq [0, x]$ with $|I| \leq 2T$, $k_jp_j^2 \in I$ for each $j \in \{1, 2, 3\}$. Analogous to the above, we get that $ab(a+b) \neq 0$. Observe that (9) still holds so that there are $\leq 2T$ choices for k_2 modulo p^2. Also, $k_2 \leq x/X^2$. Let $k \in \{0, 1, \ldots, p^2 - 1\}$ be a representative of one of the $\leq 2T$ residue classes modulo p^2 such that $k_2 \equiv k \pmod{p^2}$ can hold. Thus, any $k_2 \equiv k \pmod{p^2}$ can be written in the form $k + up^2$ where u is a non-negative integer. Note that $k_2 \leq x/X^2$ implies that $u \leq x/X^4$. Analogous to (9), we have that

$$0 < k_3(p+a+b)^2 - k_2(p+a)^2 \leq 2T. \tag{11}$$

Thus, with k fixed, k_2 and, hence, u is in one of $\leq 2T$ residue classes modulo $(p+a+b)^2$. Since $u \leq x/X^4$, there are

$$\leq \left(\frac{x/X^4}{X^2} + 1\right) 2T$$

possible values of u. Letting k vary now, we get that there are

$$\leq \left(\frac{x}{X^6} + 1\right) 4T^2$$

possible values of k_2. Given k_2, (9) and (11) imply that k_1 and k_3 are uniquely determined. Allowing p_1, p_2, and p_3 to vary, we now get that

$$S_3(x, X, T) \ll \left(\frac{x}{X^6} + 1\right) T^2 \times \frac{X^3}{\log^3 X} = \frac{xT^2}{X^3 \log^3 X} + \frac{T^2 X^3}{\log^3 X}. \tag{12}$$

These are the estimates that we will use in the sections to follow.

4. An Estimate for Short Gaps

We now proceed to use (12) to obtain results about short gaps between squarefree numbers.

Lemma 1. *Let $\gamma \in [3, 4)$ and $\theta < 2/(5\gamma - 5)$. Then*

$$\sum_{T_0 < t \leq x^\theta} N_t'' t^\gamma \ll_{\gamma,\theta} \frac{x}{T_0^{4-\gamma}} + o(x)$$

as $x \to \infty$.

Proof. Let $T \geq T_0$, and let s and $s' \in \{1, \ldots, r\}$ where r is as in Section 2. It follows from

(8), (12), (10), and (4) that

$$\begin{aligned}\sum_{T<t\le 2T} N_t'' &\ll_\epsilon \sum_{j=1}^{s} \frac{S_3(x,X_j,T)}{T^3}(1+\epsilon)^{3j} + \sum_{j=s+1}^{r} \frac{S_2(x,X_j,T)}{T^2}(1+\epsilon)^{2j} \\ &\ll_\epsilon \sum_{j=1}^{s}\left(\frac{x}{TX_j^3\log^3 X_j} + \frac{X_j^3}{T\log^3 X_j}\right)(1+\epsilon)^{3j} \\ &\quad + \sum_{j=s+1}^{s'}\left(\frac{x}{TX_j^2\log^2 X_j} + \frac{X_j^2}{T\log^2 X_j}\right)(1+\epsilon)^{2j} + \sum_{j=s'+1}^{r}\frac{x}{T^2X_j}(1+\epsilon)^{2j}.\end{aligned}$$

Let $\delta = (2/(5\gamma-5)) - \theta > 0$. Fix $\epsilon \in (0,1/5)$ such that

$$(1+\epsilon)^{3r} \le (1+\epsilon)^{3\log x} = x^{3\log(1+\epsilon)} < x^{\delta/2}.$$

Take $s \in \{0,1,\dots,r-1\}$ as small as possible such that

$$x^{(1-\delta)/3}T^{(1-\gamma)/3} \le X_{s+1}.$$

In particular, if $s \ge 1$, then

$$X_s < x^{(1-\delta)/3}T^{(1-\gamma)/3} \le X_{s+1}. \tag{13}$$

Let $s' \in \{0,1,\dots,r-1\}$ as small as possible such that

$$x^{\delta}T^{\gamma-2} \le X_{s'+1}.$$

By the definition of X_j, we get that

$$\sum_{j=1}^{s}\frac{x}{TX_j^3\log^3 X_j}(1+\epsilon)^{3j} \le \sum_{j=1}^{s}\frac{x}{T^4}\left(\frac{1+\epsilon}{2}\right)^{3j} \le \frac{x}{T^4}\sum_{j=1}^{\infty}\left(\frac{3}{4}\right)^{3j} \ll \frac{x}{T^4}.$$

Also, by (13),

$$\sum_{j=1}^{s}\frac{X_j^3}{T\log^3 X_j}(1+\epsilon)^{3j} \ll \frac{x^{1-\delta}}{T^\gamma}x^{\delta/2} = \frac{x^{1-(\delta/2)}}{T^\gamma}.$$

For $T \le x^\theta$, we also get that

$$\begin{aligned}\sum_{j=s+1}^{s'}\left(\frac{x}{TX_j^2\log^2 X_j} + \frac{X_j^2}{T\log^2 X_j}\right)(1+\epsilon)^{2j} &\ll \frac{x^{1+(\delta/2)}}{TX_{s+1}^2} + \frac{X_{s'+1}^2x^{\delta/2}}{T} \\ &\ll x^{(1/3)+(7\delta/6)}T^{(2\gamma-5)/3} + T^{2\gamma-5}x^{5\delta/2} \ll \frac{x^{1-\delta}}{T^\gamma}.\end{aligned}$$

By the definition of s', we get that

$$\sum_{j=s'+1}^{r} \frac{x}{T^2 X_j}(1+\epsilon)^{2j} \leq \frac{x^{1-(\delta/2)}}{T^\gamma}.$$

Thus,

$$\sum_{T<t\leq 2T} N_t'' t^\gamma \ll_\epsilon \frac{x}{T^{4-\gamma}} + x^{1-(\delta/2)}.$$

Taking T of the form $2^j T_0$ and summing, we get that

$$\sum_{T_0<t\leq x^\theta} N_t'' t^\gamma \ll_{\gamma,\theta} \frac{x}{T_0^{4-\gamma}} + x^{1-(\delta/2)} \log x.$$

The lemma now follows.

Observe that if $3 \leq \gamma < 29/9$ and $\theta \leq 9/50 = 0.18$, then the lemma implies that

$$\sum_{T_0<t\leq x^\theta} N_t'' t^\gamma \ll_{\gamma,\theta} \frac{x}{T_0^{4-\gamma}} + o(x).$$

By a recent result of Trifonov and the author [3], there is a constant c such that $N_t = 0$ (and, hence, $N_t'' = 0$) if $t > cx^{1/5}\log x$. The idea now is to estimate

$$\sum_{x^\theta<t\leq cx^{1/5}\log x} N_t'' t^\gamma.$$

5. Estimating Large Gaps

In this section, we will obtain the final estimates needed to show that one can obtain (2) and, hence, (1) with $\gamma \in [0, 29/9)$. In Section 6, we will combine the estimates we have obtained to establish our main result.

Fix $\epsilon \in (0, 1/5)$ and $\theta > 1/6$. We now seek to estimate $\sum_{T<t\leq 2T} N_t''$ with $T \in [x^\theta, cx^{1/5}\log x]$. Recall that (8) was obtained by fixing $j \in \{1, 2, \ldots, r\}$ and estimating the number of intervals I such that $j(I) = j$. Let j and I satisfy the above with $|I| = t \in (T, 2T]$ and $T \geq x^\theta$. Then

$$Y_j(I) \geq \frac{1}{4C}(1+\epsilon)^{-j} T. \tag{14}$$

We will show that (14) cannot hold for j small. More specifically, we will show that (14) cannot hold when $X_j \leq c_1 T^{10/9}$ where c_1 is a sufficiently small positive constant. Note that $T \leq cx^{1/5}\log x$ implies that $T^{10/9} \leq x^{1/4}$.

Fix $X \in ((1/10)T\log T, c_1T^{10/9}]$, and let

$$Y(I) = |\{m \in I: \text{ there exists a prime } p \in (X, 2X] \text{ such that } p^2|m\}|.$$

We will find an upper bound on $Y(I)$. Observe that $X \geq (1/10)T\log T \geq (1/60)T\log x > t$ for x sufficiently large. Hence, given $p \in (X, 2X]$, there is at most 1 integer $m \in I$ such that $p^2|m$. Thus,

$$Y(I) \leq |\{p \in (X, 2X]: \text{ there exists an integer } m \in I \text{ such that } p^2|m\}|.$$

This easily implies that $Y(I) \leq |S(I)|$ where

$$S(I) = \{u \in (X, 2X] \cap \mathbb{Z}: \text{ there exists an integer } m \in I \text{ such that } u^2|m\}.$$

We will use divided differences to estimate $|S(I)|$. This idea has its origins in the work of Trifonov [6,7] (also see [2]). The conditions $X \leq c_1T^{10/9} \leq x^{1/4}$ and $T \leq cx^{1/5}\log x$ imply that for any constant c_2,

$$\max\left\{X^{5/6}x^{-1/6}, 2Xx^{-1/4}T^{1/4}\right\} \leq c_2X^{3/5}x^{-1/10} \leq X \tag{15}$$

(the latter inequality being obvious). Let J be a subinterval of $(X, 2X]$ with

$$\frac{c_2}{2}X^{6/10}x^{-1/10} \leq |J| \leq c_2X^{6/10}x^{-1/10}, \tag{16}$$

where c_2 is a sufficiently small positive constant. Suppose that u, a_1, a_2, a_3, and a_4 are positive integers such that u, $u+a_1$, $u+a_1+a_2, \dots,$ $u+a_1+\dots+a_4 \in S(I) \cap J$, and let k_0, k_1, k_2, k_3, and k_4 be integers such that $k_j(u+a_1+\dots+a_j)^2 \in I$ for each $j \in \{0, 1, \dots, 4\}$. Let $y \in I$. Recall from Section 2 that we are interested in intervals I of the form $I = (s_n, s_{n+1})$ where $s_{n+1} \in (x/2, x]$ with x sufficiently large. Thus, we may choose $y \geq [(x-1)/2] > x/3$. We have that

$$k_j = \frac{y}{(u+a_1+\dots+a_j)^2} + O\left(tX^{-2}\right) \qquad \text{for } j \in \{0, 1, \dots, 4\}. \tag{17}$$

We use a divided difference for $f(u) = y/u^2$ to approximate $f^{(4)}(u)$. Let

$$D_j = (-1)^{4-j}\left(\prod_{i=0}^{j-1}(a_{i+1}+a_{i+2}+\dots+a_j)\right)\left(\prod_{i=j+1}^{4}(a_{j+1}+a_{j+2}+\dots+a_i)\right)$$

for $j \in \{0, 1, \dots, 4\}$, where the value 1 is assigned to empty products. Viewing a_1, a_2, a_3, and a_4 as variables, $\prod_{j=0}^4 D_j$ is the square of a polynomial, say D, in $\mathbb{Z}[a_1, \dots, a_4]$. Thus,

$$D = \left(\prod_{j=0}^{4} D_j\right)^{1/2}$$

is an integer. Note that D_j divides D for each $j \in \{0, 1, \ldots, 4\}$. Let

$$M_j = D/D_j \quad \text{for } j \in \{0, 1, \ldots, 4\}.$$

Then

$$\sum_{j=0}^{4} D_j^{-1} \frac{y}{(u + a_1 + \cdots + a_j)^2} = \frac{f^{(4)}(u)}{4!} + O\left(f^{(4)}(u) \max_{1 \le j \le 4} \{a_j\}/u \right).$$

We have used here that $\max_{1 \le j \le 4}\{a_j\} \le |J| \le c_2 X^{6/10} x^{-1/10}$, which is small compared to u which is $\ge X$. Multiplying both sides above by D and using (17), we obtain

$$\sum_{j=0}^{4} M_j k_j = \frac{5Dy}{u^6} + O\left(Dy \max_{1 \le j \le 4} \{a_j\}/u^7 \right) + \sum_{j=0}^{4} O\left(M_j t X^{-2} \right). \tag{18}$$

Observe that the left-hand side of (18) is an integer. A formula similar to (18) in which $f^{(3)}(u)$ is approximated by divided differences can be obtained in an analogous fashion (also, see (15) in [2]). Specifically, suppose that u', a_1', a_2', and a_3' are positive integers such that u', $u' + a_1'$, $u' + a_1' + a_2'$, and $u' + a_1' + a_2' + a_3' \in S(I) \cap J$, and let k_0', k_1', k_2', and k_3' be integers such that $k_j'(u' + a_1' + \cdots + a_j')^2 \in I$ for each $j \in \{0, 1, 2, 3\}$. Defining

$$M_0' = -a_2' a_3' (a_2' + a_3'), \;\; M_1' = a_3' (a_1' + a_2')(a_1' + a_2' + a_3'),$$

$$M_2' = -a_1' (a_2' + a_3')(a_1' + a_2' + a_3'), \;\text{ and }\; M_3' = a_1' a_2' (a_1' + a_2'),$$

we get that

$$\sum_{j=0}^{3} M_j' k_j' = \frac{4D'y}{(u')^5} + O\left(D'y \max_{1 \le j \le 3} \{a_j'\}/(u')^6 \right) + \sum_{j=0}^{3} O\left(M_j' t X^{-2} \right), \tag{19}$$

where

$$D' = a_1' a_2' a_3' (a_1' + a_2')(a_2' + a_3')(a_1' + a_2' + a_3').$$

We use (19) immediately and then use (18). The left-hand side of (19) is an integer. Recall that $u' \in (X, 2X]$, $y \in (x/3, x]$ and $T \le cx^{1/5} \log x$. Also, $X \le x^{1/4}$. Thus, if $\max\{a_1', a_2', a_3'\} \le (1/2) X^{5/6} x^{-1/6}$, then the right-hand side of (19) can be shown to be in $(0, 1)$ and, hence, cannot be an integer. This implies that

$$\max\{a_1', a_2', a_3'\} > \frac{1}{2} X^{5/6} x^{-1/6}. \tag{20}$$

Shortly, we will make use of (20), but we note for now that it is really only of value when $X \ge x^{1/5}$.

We now return to (18). Recall that $X \le c_1 T^{10/9}$. It is easy to check that if $T \le x^{9/49}$, then the main term on the right-hand side of (18) dominates the right-hand side of (18) (provided that c_1 is sufficiently small). Recall that $y \in (x/3, x]$ and $u \in (X, 2X]$. Since $a_j \leq |J| \leq c_2 X^{6/10} x^{-1/10}$ for $j \in \{1, \ldots, 4\}$, we get that

$$\frac{5Dy}{u^6} \ll \frac{c_2^{10} X^6 x^{-1} y}{u^6} \ll c_2^{10}.$$

Thus, with c_2 chosen sufficiently small, we get that the right-hand side of (18) is $\in (0, 1)$, contradicting that it must also be an integer. In other words, there cannot be integers $u, a_1, \ldots, a_4$ as above. Hence, we obtain that

$$\text{if } T \leq x^{9/49}, \text{ then } |S(I) \cap J| \leq 4. \tag{21}$$

Now, suppose that $T > x^{9/49}$. We consider two possibilities. First, we consider the case that $X \leq x^{1/2} T^{-3/2}$. One easily checks that

$$X x^{-1/4} T^{1/4} \leq X^{5/6} x^{-1/6}.$$

Recalling (20), we see that if $u', u' + a_1', u' + a_1' + a_2'$, and $u' + a_1' + a_2' + a_3' \in S(I) \cap J$, then $\max\{a_1', a_2', a_3'\} > (1/2) X^{5/6} x^{-1/6} \geq (1/2) X x^{-1/4} T^{1/4}$. We view J as a union of $\left[2|J| x^{1/6} / X^{5/6}\right] + 1$ subintervals of equal length. Since (15) and (16) imply that $|J| \geq (1/2) X^{5/6} x^{-1/6}$, each subinterval has length $\in \left[(1/4) X^{5/6} x^{-1/6}, (1/2) X^{5/6} x^{-1/6}\right]$. Thus, each such subinterval contains ≤ 3 elements from $S(I)$. We consider a sufficiently large positive integer K. We now show that there are $\leq K$ subintervals which contain 1 or more elements from $S(I)$; once this is done, we will have the result that

$$\text{if } T > x^{9/49} \text{ and } X \leq x^{1/2} T^{-3/2}, \text{ then } |S(I) \cap J| \leq 3K. \tag{22}$$

Assume that there are $> K$ subintervals as above, and consider u, $u + a_1, \ldots, u + a_1 + \cdots + a_4$ as being in different subintervals spaced so that

$$a_j > \frac{K}{20} X^{5/6} x^{-1/6} \geq \frac{K}{20} X x^{-1/4} T^{1/4} \qquad \text{for } j \in \{1, 2, 3, 4\}.$$

Recall that M_j divides D. With K sufficiently large, one gets that the main term on the right-hand side of (18) dominates the error terms. The procedure used above to establish (21) gives that the right-hand side of (18) is in $(0, 1)$, which is a contradiction. Thus, (22) holds.

Now, consider the case that $X > x^{1/2} T^{-3/2}$. Observe that since $X \leq c_1 T^{10/9}$, we get that in this case $T > x^{9/47}$. We again view J as a union of subintervals. However, we take $\left[|J| x^{1/4} / (X T^{1/4})\right] + 1$ subintervals of equal length. Here, (15) and (16) imply that $|J| \geq X x^{-1/4} T^{1/4}$ so that each subinterval has length $\in \left[(1/2) X x^{-1/4} T^{1/4}, X x^{-1/4} T^{1/4}\right]$. We then consider each such subinterval as a union of smaller subintervals of length $\leq (1/2) X^{5/6} x^{-1/6}$ (which may be < 1). Each smaller subinterval will by (20) have ≤ 3 elements from $S(I)$. Note that in this case, $X x^{-1/4} T^{1/4} > X^{5/6} x^{-1/6}$. Thus, each of the original subintervals of length $\leq X x^{-1/4} T^{1/4}$ has

$$\leq 3 \left(\frac{X x^{-1/4} T^{1/4}}{(1/2) X^{5/6} x^{-1/6}} + 1 \right) \leq 10 X^{1/6} x^{-1/12} T^{1/4}$$

elements from $S(I)$. As in the previous case, for K sufficiently large, one can now show that $\leq K$ of the subintervals of length $\leq Xx^{-1/4}T^{1/4}$ can contain 1 or more elements from $S(I)$. Thus, we get that

$$\text{if } X > x^{1/2}T^{-3/2}, \text{ then } |S(I) \cap J| \leq 10KX^{1/6}x^{-1/12}T^{1/4}. \tag{23}$$

Recall that (23) only applies when $T > x^{9/47}$. Combining (21), (22), and (23), we obtain that

$$|S(I)| \ll X^{4/10}x^{1/10} \quad \text{if } T \leq x^{9/47}$$

and

$$|S(I)| \ll X^{4/10}x^{1/10} + X^{17/30}x^{1/60}T^{1/4} \quad \text{if } T > x^{9/47}.$$

Fix $\gamma < 29/9$ and $\theta \in (9/50, 2/(5\gamma - 5))$. Let $\omega = (40\theta + 9)/(90\theta)$. Observe that

$$\omega = \frac{4}{9} + \frac{1}{10\theta} < \frac{4}{9} + \frac{5}{9} = 1.$$

Note that if $T \geq x^{\theta}$, then $x^{1/10} \leq T^{1/(10\theta)}$. Since $X \leq c_1T^{10/9}$, we get that

$$|S(I)| \ll T^{(40\theta+9)/(90\theta)} = T^{\omega} \quad \text{for } x^{\theta} \leq T \leq x^{9/47}.$$

Also, if $T > x^{9/47}$, then

$$X^{4/10}x^{1/10} \ll T^{4/9}T^{47/90} = T^{29/30}$$

and

$$X^{17/30}x^{1/60}T^{1/4} \ll T^{17/27}T^{47/540}T^{1/4} = T^{29/30}.$$

Thus,

$$|S(I)| \ll T^{29/30} \quad \text{for } T > x^{9/47}.$$

Hence, $|S(I)| \ll T^{\omega} + T^{29/30}$ whenever $X \leq c_1T^{10/9}$ and $T \geq x^{\theta}$. Observe that for $\theta \in (9/50, 2/(5\gamma - 5))$, for $T \geq x^{\theta}$, for $j \in \{1, \ldots, r\}$, and for $\epsilon = \epsilon(\theta) > 0$ sufficiently small, we now get that

$$\begin{aligned}\frac{\epsilon}{4}(1+\epsilon)^{-j}T \geq \frac{\epsilon}{4}(1+\epsilon)^{-\log x}T &= \frac{\epsilon}{4}x^{-\log(1+\epsilon)}T \\ &\geq \frac{\epsilon}{4}T^{1-(\log(1+\epsilon)/\theta)} > |S(I)|.\end{aligned}$$

Recalling that $|S(I)|$ is an upper bound on $Y(I)$, we get

Lemma 2. *Let $\gamma \in [3, 29/9)$ and $\theta \in (9/50, 2/(5\gamma - 5))$. Let $I = (s_n, s_{n+1})$ where $s_{n+1} \in (x/2, x]$. Suppose that $|I| = t \in (T, 2T]$ where $T \geq x^{\theta}$. Let $\epsilon > 0$ and $C = 1/\epsilon$. There is a positive constant c_1 such that if $\epsilon = \epsilon(\theta) > 0$ is sufficiently small and*

$$Y_j(I) \geq \frac{1}{4C}(1+\epsilon)^{-j}T,$$

then

$$X_j \geq c_1T^{10/9}.$$

6. The Main Result

Let γ and θ be as in Lemma 2, and let $\epsilon = \epsilon(\theta) > 0$ be sufficiently small. Recall that for $I = (s_n, s_{n+1})$ with $s_{n+1} \in (x/2, x]$, there is a $j = j(I)$ such that

$$Y_j(I) \geq \frac{1}{4C}(1+\epsilon)^{-j}T \geq \frac{1}{4C}(1+\epsilon)^{-\log x}T.$$

Observe that with ϵ sufficiently small and $T \geq x^\theta$, the right-hand side above is ≥ 2. Let $|I| = t \in (T, 2T]$ where $T \geq x^\theta$; then Lemma 2 implies that

$$X_j \geq c_1 T^{10/9}.$$

Let r be the greatest integer $\leq \log x$, and let $X'_j = 2^{j-1} c_1 T^{10/9}$. Following the argument in Section 2 and beginning with a fixed $j \in \{1, \ldots, r\}$, we get that the number of intervals $I = (s_n, s_{n+1})$ with $s_{n+1} \in (x/2, x]$, with $|I| = t \in (T, 2T]$ where $T \geq x^\theta$, and with

$$\left|\{m \in I : \exists \text{ a prime } p \in (X'_j, 2X'_j] \text{ such that } p^2 | m\}\right| \geq \frac{1}{4C}(1+\epsilon)^{-\log x}T,$$

is

$$\ll_\epsilon \frac{S_2(x, X'_j, T)}{T^2}(1+\epsilon)^{2\log x}.$$

Thus, if $T \geq x^\theta$, then

$$\sum_{T < t \leq 2T} N''_t \ll_\epsilon \sum_{j=1}^{r} \frac{S_2(x, X'_j, T)}{T^2}(1+\epsilon)^{2\log x}$$

(where we may actually replace N''_t above with N_t). Let $s \in \{1, \ldots, r-1\}$ such that

$$X'_s < x^{1/4} \leq X'_{s+1}.$$

Note that if $X \leq x^{1/4}$, then

$$\frac{xT}{X^2 \log^2 X} \geq \frac{TX^2}{\log^2 X}$$

so that by (10) we obtain that

$$\begin{aligned}\sum_{j=1}^{s} \frac{S_2(x, X'_j, T)}{T^2}(1+\epsilon)^{2\log x} &\ll \sum_{j=1}^{s} \frac{x}{T(X'_j)^2 \log^2 X'_j}(1+\epsilon)^{2\log x} \\ &\ll \sum_{j=1}^{s} \frac{x}{2^{2j} T^{29/9}} (1+\epsilon)^{2\log x} \ll \frac{x}{T^{29/9}}(1+\epsilon)^{2\log x}.\end{aligned}$$

Recall that we can restrict our attention to $T \leq cx^{1/5}\log x$ so that if $X > x^{1/4}$, then $X \geq T^{11/9}$. Hence, by (4),

$$\sum_{j=s+1}^{r} \frac{S_2(x, X_j', T)}{T^2}(1+\epsilon)^{2\log x} \ll \sum_{j=s+1}^{r} \frac{x}{T^2 X_j'}(1+\epsilon)^{2\log x}$$
$$\ll \sum_{j=s+1}^{r} \frac{x}{T^{29/9}} \frac{(1+\epsilon)^{2\log x}}{2^{j-s-1}} \ll \frac{x}{T^{29/9}}(1+\epsilon)^{2\log x}.$$

Thus, we get that

$$\sum_{T<t\leq 2T} N_t'' \ll_\epsilon \frac{x}{T^{29/9}}(1+\epsilon)^{2\log x}.$$

Recall that $\gamma < 29/9$ and $\theta > 9/50$. Hence, for $T \geq x^\theta$,

$$\sum_{T<t\leq 2T} N_t'' t^\gamma \ll_\epsilon \frac{x}{T^{(29/9)-\gamma}}(1+\epsilon)^{2\log x} \ll_\epsilon x^{1-\theta((29/9)-\gamma)}(1+\epsilon)^{2\log x} \ll_\epsilon \frac{x}{\log^2 x}$$

provided ϵ is sufficiently small depending on γ and θ. We now fix such an ϵ. Hence,

$$\sum_{t>x^\theta} N_t'' t^\gamma = o(x).$$

From (7) and Lemma 1, we now get that for $\gamma < 29/9$

$$\sum_{t>T_0} N_t t^\gamma \ll \frac{x}{T_0^{4-\gamma}} + o(x).$$

We now follow Hooley [4]. We use that there is a function $A(t)$ such that

$$N_t = A(t)x + o(x).$$

This follows from a result of Mirsky [5]. Thus, viewing T_0 as a fixed but large number, we get that

$$\sum_{t\leq T_0} N_t t^\gamma = x\sum_{t\leq T_0} A(t) + o(x).$$

Hence,

$$\frac{1}{x}\sum_{t=1}^{\infty} N_t t^\gamma = \sum_{t\leq T_0} A(t) + O\left(\frac{1}{T_0^{4-\gamma}}\right) + o(1).$$

Hence,

$$\limsup_{x\to\infty} \frac{1}{x}\sum_{t=1}^{\infty} N_t t^\gamma = \sum_{t\leq T_0} A(t) + O\left(\frac{1}{T_0^{4-\gamma}}\right)$$

and

$$\liminf_{x\to\infty} \frac{1}{x}\sum_{t=1}^{\infty} N_t t^\gamma = \sum_{t\leq T_0} A(t) + O\left(\frac{1}{T_0^{4-\gamma}}\right).$$

Each of these imply that $\sum_{t=1}^{\infty} A(t)$ converges. The result now follows by setting $B'(\gamma) = \sum_{t=1}^{\infty} A(t)$ and letting T_0 approach infinity above.

7. A Further Remark

We end by commenting briefly on how one can use the ideas from this paper to establish the equivalence of Conjectures 1 and 2 stated in the introduction. First, assume that Conjecture 2 is not true. Then there is an $\epsilon > 0$ such that there exist arbitrarily large x for which the interval $(x, x + x^{\epsilon}]$ contains no squarefree numbers. It is easy to see then that Conjecture 1 is false for $\gamma > 1/\epsilon$.

Now, assume that Conjecture 2 is true. For j a positive integer, we define $S_j(x, X, T)$ as the number of $2j-$tuples $(p_1, \dots, p_j, k_1, \dots, k_j)$ with $p_1, \dots, p_j$ primes and $k_1, \dots, k_j$ positive integers satisfying $k_j p_j^2 > \cdots > k_2 p_2^2 > k_1 p_1^2$, $X < p_i \le 2X$ for each $i \in \{1, \dots, j\}$, and for some $I \subseteq [0, x]$ with $|I| \le 2T$, $k_i p_i^2 \in I$ for each $i \in \{1, \dots, j\}$. The arguments used to estimate $S_2(x, X, T)$ and $S_3(x, X, T)$ in Section 3 generalize to give that

$$S_j(x, X, T) \ll_j \frac{xT^{j-1}}{X^j \log^j X} + \frac{T^{j-1} X^j}{\log^j X}. \tag{24}$$

Furthermore, the argument in Section 2 implies that for $T \ge T_0$ and for each j,

$$\sum_{T < t \le 2T} N_t'' \ll_\delta \sum_{i=1}^{r} \min\left\{ \frac{S_2(x, X_i, T)}{T^2}(1+\delta)^{2i}, \frac{S_j(x, X_i, T)}{T^j}(1+\delta)^{ji} \right\}. \tag{25}$$

Fix $\gamma > 1$, and fix j to be a positive integer with $j > \gamma$. Fix $\epsilon < 1/(4j\gamma)$. Note that $N_t = 0$ if $t > x^{1/(4j\gamma)}$ (with x sufficiently large). Thus, we need only consider $T \le x^{1/(4j\gamma)}$. Let $s \in \{1, \dots, r-1\}$ as small as possible such that

$$x^{1/(2j)} \le X_{s+1}.$$

For $\delta \in (0, 1/5)$ we get from (24) that

$$\begin{aligned} \sum_{i=1}^{s} \frac{S_j(x, X_i, T)}{T^j}(1+\delta)^{ji} &\ll \sum_{i=1}^{s} \frac{x}{TX_i^j \log^j X_i}(1+\delta)^{ji} \\ &\ll \sum_{i=1}^{s} \frac{x}{T^{j+1}} \left(\frac{1+\delta}{2}\right)^{ji} \ll \frac{x}{T^{\gamma+1}}. \end{aligned}$$

Observe that for $T \le x^{1/(4j\gamma)}$ and $X \ge x^{1/(2j)}$, we have that $X \ge T^{\gamma} x^{1/(4j)}$. Thus, by (4),

$$\begin{aligned} \sum_{i=s+1}^{r} \frac{S_2(x, X_i, T)}{T^2}(1+\delta)^{2i} &\ll \sum_{i=s+1}^{r} \frac{x}{T^2 X_i}(1+\delta)^{2i} \ll \sum_{i=s+1}^{r} \frac{x}{T^{\gamma+2}} x^{-1/(4j)}(1+\delta)^{2i} \\ &\ll_\delta \frac{x}{T^{\gamma+2}} x^{-1/(4j)} (1+\delta)^{2\log x} \ll \frac{x}{T^{\gamma+2}} \end{aligned}$$

provided δ is sufficiently small. Fixing δ so that the above all holds, we now get from (25) that

$$\sum_{T < t \le 2T} N_t'' \ll \frac{x}{T^{\gamma+1}}$$

from which it follows that

$$\sum_{t>T_0} N_t'' t^\gamma \ll \frac{x}{T_0}.$$

If δ is sufficiently small, one can obtain a modified version of (7) so that the argument in Section 6 implies now that the asymptotic formula given in (1) holds for γ. Since γ was an arbitrary number > 1, Conjecture 1 now follows as a consequence of Conjecture 2, completing the proof that these conjectures are equivalent.

We note before closing this section that (24) and (25) imply that for a given $\gamma > 0$, there is an $\epsilon > 0$ such that

$$\sum_{\substack{s_{n+1}\le x \\ s_{n+1}-s_n\le x^\epsilon}} (s_{n+1}-s_n)^\gamma \sim B(\gamma)x. \tag{26}$$

In other words, (1) holds for γ when the sum is restricted to gaps of size x^ϵ where ϵ is sufficiently small. This fact implies then that if in (26) x^ϵ is replaced by a function which tends to infinity with x and which is $o(x^\epsilon)$ for every $\epsilon > 0$, then (26) will hold for all $\gamma > 0$.

In conclusion, the author expresses his gratitude to Sid Graham for helpful suggestions during the writing of this paper.

References

1. P. Erdős, *Some problems and results in elementary number theory*, Publ. Math. Debrecen **2** (1951), 103–109.
2. M. Filaseta and O. Trifonov, *On gaps between squarefree numbers*, Analytic Number Theory, Proceedings of a Conference in Honor of Paul T. Bateman (Progress in Mathematics Series, Vol. 85), edited by Berndt, Diamond, Halberstam, and Hildebrand, Birkhäuser, Boston (1990), 235–253.
3. M. Filaseta and O. Trifonov, *On gaps between squarefree numbers II*, J. London Math. Soc. (2) **45** (1992), 215–221.
4. C. Hooley, *On the distribution of square-free numbers*, Can. J. Math. **25** (1973), 1216 –1223.
5. L. Mirsky, *Arithmetical pattern problems related to divisibility by r-th powers*, Proc. London Math. Soc. **50** (1949), 497–508.
6. O. Trifonov, *On the squarefree problem*, C. R. Acad. Bulgare Sci. **41** (1988), 37–40.
7. O. Trifonov, *On the squarefree problem II*, Mathematica Balcanika **3** (1989), 284–295.

Mathematics Department
University of South Carolina
Columbia, SC 29208