LARGE DEVIATIONS IN THE RANDOM SIEVE

Geoffrey Grimmett

Abstract. The proportion ρ_k of gaps with length k between square-free numbers is shown to satisfy $\log\rho_k = -\big(1+\mathrm{o}(1)\big)(6/\pi^2)k\log k$ as $k\to\infty$. Such asymptotics are consistent with Erdős's challenge to prove that the gap following the square-free number t is smaller than $c\log t/\log\log t$, for all t and some constant c satisfying $c > \pi^2/12$. The results of this paper are achieved by studying the probabilities of large deviations in a certain 'random sieve', for which the proportions ρ_k have representations as probabilities. The asymptotic form of ρ_k may be obtained in situations of greater generality, when the squared primes are replaced by an arbitrary sequence (s_r) of relatively prime integers satisfying $\sum_r 1/s_r < \infty$, subject to two further conditions of regularity on this sequence.

1. Introduction

A positive integer is called *square-free* if it is divisible by no squared prime. The sequence of square-free numbers has density $6/\pi^2$, but the gaps between consecutive square-free numbers can be large. Several authors have studied the lengths of these gaps, in order to try to understand how large they may be. Results obtained to date appear, however, to be far from the best possible.

Write $t_1, t_2, \dots$ for the (increasing) square-free numbers. Erdős [2] observed that

$$t_{i+1} - t_i > \big(1+\mathrm{o}(1)\big)\frac{\pi^2}{12}\,\frac{\log t_i}{\log\log t_i} \quad \text{for infinitely many } i, \tag{1.1}$$

and he asked whether it could possibly be the case that

$$t_{i+1} - t_i < (1+\epsilon)\frac{\pi^2}{12}\,\frac{\log t_i}{\log\log t_i} \quad \text{for all large } i. \tag{1.2}$$

Also, he initiated a study of the moments of the gaps, in proving that the sequence

$$L_n(g) = \frac{1}{n}\sum_{i:\,t_i\le n} g(t_{i+1}-t_i) \tag{1.3}$$

1991 *Mathematics Subject Classification.* 11N36, 11K31, 60F10.

Key words and phrases. Random sieve, large deviations, square-free numbers.

Address of author. Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, United Kingdom.

has a finite limit as $n \to \infty$, when $g(x) = x^\alpha$ and $\alpha = 2$. (A proof of (1.1) is included in Appendix 1, together with a discussion of the constant $\pi^2/12$, which actually appeared as $\pi^2/6$ in [2].)

Further results in these directions have been since obtained, and we mention two of these. Filaseta and Trifonov [5] have shown that

$$t_{i+1} - t_i = \mathrm{O}(n^{1/5} \log n) \quad \text{if} \quad t_i \le n,$$

and Huxley [15] that $L_n(g)$ has a finite limit if $g(x) = x^\alpha$ and $\alpha < \frac{11}{3}$. See also [4, 6, 7, 11, 14].

The (forthcoming) result of the present paper causes us to pose the following provocative extensions of Erdős's question. Does $L_n(g)$ have a finite limit as $n \to \infty$, when

(a) $g(x) = e^{\theta x}$ and $\theta > 0$, or
(b) $g(x) = x^{\theta x}$ ($= \exp\{\theta x \log x\}$) and θ is positive but not too large?

Inequality (1.1) implies that $L_n(g) \to \infty$ as $n \to \infty$ when $g(x) = x^{\theta x}$ and $\theta > 12/\pi^2$.

The emphasis of the current work is probabilistic. Let $S_k(m)$ be the number of square-free numbers in the interval $\{m, m+1, \dots, m+k-1\}$. It follows by a calculation of Mirsky [16] that the limit

$$p_k(j) = \lim_{n\to\infty} \frac{1}{n} \Big| \Big\{ m : S_k(m) = j,\ 1 \le m \le n \Big\} \Big| \tag{1.4}$$

exists for all k and j. The function $p_k(\cdot)$ is a probability mass function on the set $\{0, 1, \dots, k\}$, and the quantity $\pi_k = p_k(0)$ is the density of numbers m such that $\{m, m+1, \dots m+k-1\}$ contains no square-free number. We shall prove the following:

$$\log \pi_k = -\big(1 + \mathrm{o}(1)\big) \frac{6}{\pi^2} k \log k \quad \text{as } k \to \infty. \tag{1.5}$$

(All logarithms in this paper are natural.) Thus the correct order for π_k is $k^{-(6/\pi^2)k}$. This is an improvement over upper bounds of larger order obtained by Hall [12] and Huxley [15], and is relevant to the question (1.2) posed by Erdős. We amplify this statement next.

Let $g(x) = x^{\theta x}$ ($= \exp\{\theta x \log x\}$) where $\theta > 0$, and consider the quantity $L_n(g)$ defined in (1.3). An argument of weak convergence (see (2.5)) suggests that, *if $L_n(g)$ has a finite limit $\lambda(\theta)$, then it is reasonable to expect that*

$$\lambda(\theta) = \sum_k g(k)\, \rho_k$$

where ρ_k denotes the density of sequences $\{m, m+1, \dots, m+k\}$ such that m and $m+k$ are square-free, but $m+i$ is not square-free when $1 \le i < k$. Now $\rho_k \le \pi_{k-1}$

(see (3.18)), whence, by (1.5), $\lambda(\theta) < \infty$ if $\theta < 6/\pi^2$. This calculation supports the possibility that $L_n(g)$ indeed has a finite limit $\lambda(\theta)$ when $g(x) = x^{\theta x}$ and $\theta < 6/\pi^2$. If this is so, then

$$\frac{1}{t_i}\, g(t_{i+1} - t_i) \le L_{t_i}(g) \to \lambda(\theta) \quad \text{as } i \to \infty,$$

yielding (1.2) when $\epsilon > 1$.

Our main result for the square-free numbers is the following theorem, which includes the claim (1.5).

Theorem 1. *We have that, as $k \to \infty$,*

$$\begin{aligned}\log \pi_k &= -\big(1 + \mathrm{o}(1)\big)\frac{6}{\pi^2}\, k \log k,\\ \log \rho_k &= -\big(1 + \mathrm{o}(1)\big)\frac{6}{\pi^2}\, k \log k.\end{aligned}$$

We shall prove Theorem 1 in the context of the *random sieve* introduced in [8] and pursued in [9]. The random sieve will provide the correct setting for a further discussion of some of the points above. It is described in Section 2.

So far, we have concentrated on the square-free numbers. However, similar results are valid for the a-free numbers with $a \ge 2$, and more generally for the set of all integers which remain after sieving by a family $\mathcal{S} = (s_r)$ of relatively prime numbers satisfying $\sum_r 1/s_r < \infty$, subject to two conditions on $\mathcal{S}$. Further details are provided in the next section.

2. The random sieve

Rather than working with the squared primes, we shall work more generally with an increasing sequence $\mathcal{S} = (s_1, s_2, \dots)$ of relatively prime integers (so that $1 < s_1 < s_2 < \dots$, and $(s_i, s_j) = 1$ if $i \ne j$). The 'sieve generated by $\mathcal{S}$' is constructed as follows. For $m \ge 1$, we write G_m for the set of integers r such that $s_r \mid m$, and we write $G = (G_1, G_2, \dots)$. An integer m is called $\mathcal{S}$*-free* if $G_m = \varnothing$, and we denote the $\mathcal{S}$-free numbers by the increasing sequence $t_1, t_2, \dots$.

If $s_r = p_r^a$, the ath power of the rth prime, then the $\mathcal{S}$-free numbers are more commonly called 'a-free', or 'square-free' in the case $a = 2$.

The 'random sieve' is a variant of the above process, in which the action of each s_r is subject to a random translation (see [8]). Let $X_1, X_2, \dots$ be independent random variables, with $P(X_r = k) = 1/s_r$ for $1 \le k \le s_r$. For $m \ge 1$, we write Γ_m for the set of integers r such that $s_r \mid m - X_r$. The outcome is a random vector $\Gamma = (\Gamma_1, \Gamma_2, \dots)$, taking values in the state space $\Omega = (2^{\mathbb{N}})^{\mathbb{N}}$ of sequences of subsets of the natural numbers $\mathbb{N}$. An integer m is called 'randomly $\mathcal{S}$-free' if $\Gamma_m = \varnothing$.

Next we review material taken from [8]. This material is directed at understanding the way in which the *averaging properties* of the sieve generated by $\mathcal{S}$ may be represented as *probabilities* associated with the random sieve. We use the language of weak convergence (see [1]).

The sample space Ω is a product space, and we endow Ω with the product of discrete topologies. We write $\mathcal{F}$ for the σ-field of subsets of Ω generated by the open sets, and sets in $\mathcal{F}$ we call *events*. Next we introduce some probability measures on the measurable pair $(\Omega, \mathcal{F})$. For $n \geq 1$, let N_n be chosen randomly and uniformly from $\{1, 2, \ldots, n\}$, so that $P(N_n = k) = 1/n$ for $1 \leq k \leq n$. Now, let $\gamma_n = (G_{N_n}, G_{N_n+1}, \ldots)$, the sequence obtained from G by deleting an initial segment of random length. Let μ_n denote the probability measure associated with γ_n, in that $\mu_n(A) = P(\gamma_n \in A)$ for $A \in \mathcal{F}$; let μ be the probability measure associated similarly with the random sieve Γ.

Theorem 2 ([8]). *We have that μ_n converges weakly to μ (as $n \to \infty$) if and only if $\sum_r 1/s_r < \infty$.*

The sufficiency of the summability condition is proved in [8]. Its necessity was given in [8] also, and follows from the following observation. Let A be the set of $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$ such that $|\omega_1| < \infty$. If $\sum_r 1/s_r = \infty$, then $\mu_n(A) = 1$ for all n whereas $\mu(A) = 0$. The picture is of course generally different with a different topology on Ω. Note that the harmonic summation of the primes diverges, and therefore the theorem does not apply in this case. Sieves satisfying this summability condition have been studied in [3, 17] and elsewhere.

It is a consequence of this weak convergence that

$$\int f \, d\mu_n \to \int f \, d\mu \quad \text{as } n \to \infty$$

for all bounded continuous functions $f : \Omega \to \mathbb{R}$.

In this paper we are concerned with the incidence of $\mathcal{S}$-free numbers, and especially the lengths of intervening gaps. To this end, we define functions f_{jk} on Ω as follows. For non-negative integers j, k, and a sequence $\omega = (\omega_1, \omega_2, \ldots) \in \Omega$, we define

$$T_k(\omega) = \sum_{i=1}^{k} 1_{\{\omega_i = \varnothing\}}, \qquad f_{jk}(\omega) = 1_{\{T_k(\omega) = j\}},$$

where 1_A denotes the indicator function of an event A. Thus, for example, $f_{jk}(G)$ equals 1 if and only if exactly j of the first k integers are $\mathcal{S}$-free. Furthermore,

$$\int f_{jk} \, d\mu_n = \frac{1}{n} \Big| \Big\{ m : S_k(m) = j, \ 1 \leq m \leq n \Big\} \Big|$$

where $S_k(m)$ is the number of $\mathcal{S}$-free integers in $\{m, m+1, \ldots, m+k-1\}$; cf. (1.4). Clearly f_{jk} is bounded and continuous, and therefore (by Theorem 2)

$$\int f_{jk} \, d\mu_n \to p_k(j) \quad \text{as } n \to \infty,$$

whenever $\sum_r 1/s_r < \infty$, where

$$p_k(j) = \mu\big(f_{jk}(\Gamma) = 1\big) = \mu\big(T_k(\Gamma) = j\big).$$

The mean of the probability mass function $p_k(\cdot)$ is

$$\sum_{j=0}^{k} j\, p_k(j) = E\big(T_k(\Gamma)\big) = \sum_{i=1}^{k} \mu(\Gamma_i = \varnothing) = k\zeta$$

where E denotes expectation, and $\zeta = \zeta(\mathcal{S})$ is given by

$$\zeta = \prod_r \left(1 - \frac{1}{s_r}\right).$$

Our result in this paper is a large-deviation theorem for the mass function p_k, in the limit as $k \to \infty$. Such a theorem is valid subject to additional conditions on the sequence $\mathcal{S}$, and we state these next. Write $\Sigma(n) = |\mathcal{S} \cap [1, n]|$, the *growth function* of $\mathcal{S}$. We shall require that there exists σ satisfying $0 < \sigma < 1$ such that

$$\frac{\log \Sigma(n)}{\log n} \to \sigma \quad \text{as } n \to \infty \tag{2.1}$$

and that

$$\limsup_{n\to\infty} \left\{ \frac{\log \sum_{r:\, s_r > n} s_r^{-1}}{\log n} \right\} \le \sigma - 1. \tag{2.2}$$

At first sight these conditions may appear somewhat artificial. However, they are valid in a variety of instances. It is worthwhile to contrast them with the *stronger* condition of regular variation (of Σ) utilised in [9]. Recall that Σ is said to be *regularly varying* if the limit

$$\ell(c) = \lim_{n\to\infty} \frac{\Sigma(cn)}{\Sigma(n)}$$

exists for all $c > 0$. If Σ is regularly varying then it is easily seen that $\ell(c) = c^\sigma$ for some σ, called the *index* of the function. Furthermore, the condition of regular variation is valid with index $\sigma = 1/a$, if $\mathcal{S}$ is the sequence of ath powers of the primes. Also, for any σ satisfying $0 < \sigma < 1$, there exists a sequence $\mathcal{S}$ of relatively prime numbers for which $\Sigma(n) = |\mathcal{S} \cap [1, n]|$ is regularly varying with index σ. Such a sequence may be obtained by a suitable thinning of the prime numbers.

If Σ is regularly varying with index σ (where $0 < \sigma < 1$), then (2.1) and (2.2) hold (see [9]).

The distribution function associated with the mass function p_k is denoted by F_k, so that

$$F_k(x) = \mu\big(T_k(\Gamma) \le x\big) = \sum_{j=0}^{\lfloor x \rfloor} p_k(j) \quad \text{for } x \ge 0,$$

where $\lfloor x \rfloor$ denotes the integer part of x.

Theorem 3. *Suppose that $\mathcal{S}$ satisfies (2.1) and (2.2) for some σ with $0 < \sigma < 1$. If $0 \le \nu < \zeta$ then*

$$\log F_k(\nu k) = -(\zeta - \nu)\left(\frac{1}{\sigma} - 1\right)\big(1 + \mathrm{o}(1)\big)k \log k \quad \text{as } k \to \infty. \tag{2.3}$$

This theorem asserts that the probability of a deviation having size $(\zeta - \nu)k$ *beneath* the mean (i.e., $\nu < \zeta$) is of order $k^{-f(\nu,\sigma)k}$ for an appropriate positive quantity $f(\nu, \sigma)$. Deviations *above* the mean have even smaller probability. More specifically, it may be seen by the Chinese Remainder Theorem that, if $\nu > \zeta$, then $p_k(j) = 0$ if $j \ge \nu k$ and k is sufficiently large. Other properties of the distribution associated with F_k were established in [8, 9], particularly results concerning its height.

Applying Theorem 3 with $\nu = 0$, we find that

$$\log p_k(0) = -\big(1 + \mathrm{o}(1)\big)\left(\frac{1}{\sigma} - 1\right)\left\{\prod_r \left(1 - \frac{1}{s_r}\right)\right\} k \log k. \tag{2.4}$$

If $\mathcal{S}$ is the sequence of squared primes, then $\sigma = \frac{1}{2}$ and

$$\log p_k(0) = -\big(1 + \mathrm{o}(1)\big)\frac{6}{\pi^2}\, k \log k$$

in agreement with the first claim of Theorem 1. In this case (and more generally for the a-free numbers) one may in principle obtain further information about the o(1) term in (2.4), by using the more detailed asymptotics associated with the distribution of the primes.

The second claim of Theorem 1 is similar to the first, as is its proof. We have by Mirsky's theorem (and Theorem 2) that

$$\rho_k = \mu\big(\Gamma_1 = \varnothing,\ \Gamma_i \ne \varnothing \text{ for } 1 < i \le k,\ \Gamma_{k+1} = \varnothing\big).$$

Theorem 4. *Suppose that $\mathcal{S}$ satisfies (2.1) and (2.2) for some σ with $0 < \sigma < 1$.*
(a) *If $s_1 > 2$ then*

$$\log \rho_k = -\big(1 + \mathrm{o}(1)\big)\left(\frac{1}{\sigma} - 1\right)\left\{\prod_r \left(1 - \frac{1}{s_r}\right)\right\} k \log k \qquad \text{as } k \to \infty.$$

(b) *If $s_1 = 2$, the same asymptotic relation is valid so long as k tends to infinity through the even numbers. We have that $\rho_k = 0$ if k is odd.*

Finally in this section, we discuss the matter of the lengths of gaps between $\mathcal{S}$-free numbers. The appropriate function $h : \Omega \to \mathbb{R}$ is given by

$$h(\omega) = 1_{\{\omega_1 = \varnothing\}} \times \Big(\inf\{k \ge 2 : \omega_k = \varnothing\} - 1\Big)$$

for $\omega = (\omega_1, \omega_2, \dots) \in \Omega$. For $g : \mathbb{R} \to \mathbb{R}$, we have that

$$\int g\big(h(\cdot)\big)\, d\mu_n = \frac{1}{n} \sum_{i:\, t_i \le n} g(t_{i+1} - t_i),$$

where $t_1, t_2, \dots$ is the sequence of $\mathcal{S}$-free numbers as before. The function $g\big(h(\cdot)\big)$ maps Ω into $\mathbb{R}$, and is continuous but not generally bounded. It follows by weak convergence that

$$\int g\big(h(\cdot)\big)\, d\mu_n \to E\left(g\big(h(\Gamma)\big)\right) \quad \text{as } n \to \infty \tag{2.5}$$

subject to the condition of uniform integrability, namely that

$$\sup_n \int_{A_M} g\big(h(\cdot)\big)\, d\mu_n \to 0 \quad \text{as } M \to \infty$$

where $A_M = \big\{\omega \in \Omega : |g(h(\omega))| \ge M\big\}$; see [1, p. 32]. In the interesting case, when g is an unbounded function of the positive integers, uniform integrability amounts to an upper bound on the gaps $t_{i+1} - t_i$.
Note. There is a minor error in the proof of Theorem 2 appearing in [8], and this is an appropriate place to acknowledge this. See Appendix 2.

3. Proof of Theorems 3 and 4

Proof of Theorem 3. As in [9], the strategy of the proof is to divide $\mathcal{S}$ into classes (small, medium, and large, in this case), and to show that deviations from the mean can occur only if the effect of the 'large' members of $\mathcal{S}$ is aberrant in a special way. Owing to a convenient definition of the 'large' members, the probability of such aberrant behaviour may be estimated.

Assume that $\mathcal{S}$ and Σ satisfy (2.1) and (2.2) where $0 < \sigma < 1$. In particular, $\sum_r 1/s_r < \infty$, so that

$$\zeta = \prod_r \left(1 - \frac{1}{s_r}\right) > 0. \tag{3.1}$$

Let $0 \le \nu < \zeta$ and $0 < \epsilon < \min\{\sigma, 1-\sigma, \zeta - \nu\}$, noting that $\epsilon < \sigma^{-1} - 1$. We call s_r $(\in \mathcal{S})$ *small* if $r \le R$ where $R = R(\epsilon)$ is chosen in such a way that

$$\sum_{r > R} \frac{1}{s_r} < \tfrac{1}{3}\epsilon, \tag{3.2}$$

and

$$\prod_{r=1}^{R}\left(1-\frac{1}{s_r}\right) \le \zeta + \tfrac{2}{3}\epsilon. \tag{3.3}$$

(These two inequalities are of course related.) With $U = s_1 s_2 \dots s_R$, let

$$H_k = \{j : 1 \le j \le k,\ r \in \Gamma_j \text{ for some } 1 \le r \le R\}, \tag{3.4}$$

the set of j (satisfying $1 \le j \le k$) such that Γ_j contains at least one member of $\{1, 2, \dots, R\}$. Using the Chinese Remainder Theorem and the coprimality of $s_1, s_2, \dots, s_R$, we may bound $|H_k|$ by

$$|H_k| \le \left\{1 - \prod_{r=1}^{R}\left(1-\frac{1}{s_r}\right)\right\} U\lfloor k/U \rfloor + U. \tag{3.5}$$

Therefore

$$\frac{1}{k}|H_k| \le (1-\zeta) + \frac{U}{k}\,. \tag{3.6}$$

We choose N_1 such that $U/N_1 < \frac{1}{3}\epsilon$, and obtain

$$\frac{1}{k}|H_k| \le 1 - \zeta + \tfrac{1}{3}\epsilon \quad \text{if } k \ge N_1. \tag{3.7}$$

Let $0 < \eta < \epsilon$, and let $V = \lfloor k^{\sigma^{-1}-\eta} \rfloor$ where k is sufficiently large that $s_R \le V$ (say $k \ge N_2 \ge N_1$). We partition $\mathcal{S}$ as $\mathcal{S} = \mathcal{M} \cup \mathcal{N}$ where $\mathcal{M} = \{s_r : s_r \le V\}$, $\mathcal{N} = \{s_r : s_r > V\}$. Members of $\mathcal{M} \setminus \{s_1, s_2, \dots, s_R\}$ are called 'medium', and members of $\mathcal{N}$ are called 'large'.

We may bound the cardinality of the set

$$J_k = \{j : 1 \le j \le k,\ r \in \Gamma_j \text{ for some } s_r \in \mathcal{M}\} \tag{3.8}$$

by

$$\begin{aligned} |J_k| &\le |H_k| + \sum_{\substack{r:\, s_r \in \mathcal{M} \\ r > R}} \left(1 + \frac{k}{s_r}\right) \\ &\le k(1 - \zeta + \tfrac{1}{3}\epsilon) + |\mathcal{M}| + k \cdot \tfrac{1}{3}\epsilon \qquad \text{if } k \ge N_2 \end{aligned} \tag{3.9}$$

by (3.2) and (3.7). By (2.1), there exists ρ (> 0) such that, for all large k,

$$|\mathcal{M}| = \Sigma(V) \le k^{1-\rho}.$$

Therefore there exists N_3 $(\geq N_2)$ such that $|\mathcal{M}| \leq \frac{1}{3}\epsilon k$ if $k \geq N_3$, giving from (3.9) that

$$|J_k| \leq k(1-\zeta+\epsilon) \quad \text{if } k \geq N_3. \tag{3.10}$$

Next we introduce the random set whose cardinality is to be estimated:

$$K_k = \{j : 1 \leq j \leq k,\ j \notin J_k,\ r \in \Gamma_j \text{ for some } s_r \in \mathcal{N}\}. \tag{3.11}$$

We have that

$$\begin{aligned} F_k(\nu k) &= \mu\big(|J_k| + |K_k| \geq (1-\nu)k\big) \\ &\leq \mu\big(|K_k| \geq (\zeta-\nu-\epsilon)k\big) \quad \text{by (3.10)}. \end{aligned} \tag{3.12}$$

We shall bound the last probability using Markov's inequality (see [10, p. 278]). Note that $|K_k| \leq K$ where

$$K = \sum_{r:\, s_r \in \mathcal{N}} 1_{\{X_r \leq k\}},$$

the sum of independent Bernoulli random variables with respective means

$$E\big(1_{\{X_r \leq k\}}\big) = \mu(X_r \leq k) = \frac{k}{s_r};$$

the X_r were given towards the beginning of Section 2.

Lemma. *Let $Y_1, Y_2, \dots$ be independent Bernoulli random variables with*

$$P(Y_r = 1) = 1 - P(Y_r = 0) = q_r,$$

where $\mu = \sum_r q_r < \infty$. Then

$$P\left(\sum_{r=1}^{\infty} Y_r \geq \gamma\mu\right) \leq \exp\big\{-\mu(\gamma\log\gamma + 1 - \gamma)\big\} \quad \text{for } \gamma > 1. \tag{3.13}$$

Proof. By Markov's inequality, if $\theta > 0$,

$$\begin{aligned} P\left(\sum_{r=1}^{\infty} Y_r \geq \gamma\mu\right) &\leq e^{-\gamma\mu\theta} \prod_{r=1}^{\infty} E\big(e^{\theta Y_r}\big) \\ &= e^{-\gamma\mu\theta} \prod_{r=1}^{\infty} \big(1 + q_r(e^\theta - 1)\big) \\ &\leq \exp\big\{-\gamma\mu\theta + \mu(e^\theta - 1)\big\}. \end{aligned}$$

We set $\theta = \log\gamma$ to obtain (3.13). $\square$

Returning to the proof of the theorem, the mean of $|K_k|$ satisfies

$$E|K_k| \le E(K) = \sum_{r:\, s_r\in\mathcal{N}} \frac{k}{s_r} \le k^{2-\sigma^{-1}+\epsilon} \tag{3.14}$$

for all large k (say for $k \ge N_4 \ge N_3$) by (2.2).

By (3.12), (3.14), and the lemma,

$$\log F_k(\nu k) \le \log \mu\big(K \ge (\zeta-\nu-\epsilon)k\big) \le -E(K)(\gamma\log\gamma + 1 - \gamma),$$

where

$$\gamma = \frac{(\zeta-\nu-\epsilon)k}{E(K)} \ge (\zeta-\nu-\epsilon)k^{\sigma^{-1}-1-\epsilon} \quad \text{if } k \ge N_4$$

by (3.14). This implies that

$$\log F_k(\nu k) \le -\big(1+\mathrm{o}(1)\big)(\zeta-\nu-\epsilon)k\log k^{\sigma^{-1}-1-\epsilon} \quad \text{as } k\to\infty$$

and hence

$$\limsup_{k\to\infty}\left\{\frac{\log F_k(\nu k)}{k\log k}\right\} \le -(\zeta-\nu-\epsilon)(\sigma^{-1}-1-\epsilon) \quad \text{for all } \epsilon > 0.$$

We let $\epsilon \downarrow 0$ to obtain the upper bound necessary for the theorem.

Finally we establish a lower bound for $F_k(\nu k)$. First note a lower bound for $|H_k|$, similar to the upper bound (3.5),

$$|H_k| \ge \left\{1 - \prod_{r=1}^{R}\left(1-\frac{1}{s_r}\right)\right\} U\lfloor k/U\rfloor$$

and therefore

$$1-\zeta-\epsilon \le \frac{1}{k}|H_k| \le 1-\zeta+\epsilon \quad \text{if } k \ge N_2$$

by (3.3) and (3.7). Now, for $k \ge N_2$,

$$F_k(\nu k) = \mu\big(T_k(\Gamma)\le \nu k\big) = \sum_H \mu\big(T_k(\Gamma)\le\nu k \mid H_k = H\big)\mu(H_k = H) \tag{3.15}$$

where the summation is over all subsets H of $\{1,2,\dots,k\}$ satisfying

$$(1-\zeta-\epsilon)k \le |H| \le (1-\zeta+\epsilon)k. \tag{3.16}$$

We say that s_r 'strikes' the integer i if $r \in \Gamma_i$. Conditional on the event $\{H_k = H\}$, we have that $T_k(\Gamma) \le \nu k$ if and only if the medium and large s_r (i.e., the s_r satisfying $r > R$) strike at least $|\overline{H}| - \nu k$ elements of $\overline{H}$ (the complement of a set H is denoted as $\overline{H}$). This is certainly achieved if the earliest $|\overline{H}| - \nu k$ such s_r each strikes a new integer of $\overline{H}$, i.e., an integer struck by no smaller s_r (here and later we sometimes use real numbers where integers are required, but it may easily be checked that this notational convenience has no influence on the outcome). That is

$$\mu\big(T_k(\Gamma) \le \nu k \mid H_k = H\big) \ge \mu(A_H)$$

where

$$A_H = \Big\{X_{R+1} \in \overline{H},\, X_{R+2} \in \overline{H} \setminus \{X_{R+1}\}, \dots ,\\ X_{R+W} \in \overline{H} \setminus \{X_{R+1}, X_{R+2}, \dots , X_{R+W-1}\}\Big\},$$

with $W = |\overline{H}| - \nu k$. Now

$$\begin{aligned}\mu(A_H) &= \frac{|\overline{H}|}{s_{R+1}} \cdot \frac{|\overline{H}|-1}{s_{R+2}} \cdots \frac{|\overline{H}|-W+1}{s_{R+W}} \\ &\ge \frac{|\overline{H}|!}{(|\overline{H}|-W)!} \left(\frac{1}{s_{R+W}}\right)^{W} \\ &\ge \frac{\{(\zeta-\epsilon)k\}!}{(\nu k)!} \left(\frac{1}{s_{R+W}}\right)^{(\zeta+\epsilon-\nu)k}\end{aligned}$$

by (3.16). Furthermore, by (2.1),

$$s_{R+W} \le s_{(\zeta+2\epsilon-\nu)k} \le \{(\zeta+2\epsilon-\nu)k\}^{\sigma^{-1}+\epsilon}$$

for large k, say $k \ge N_5$ ($\ge N_2$). Substituting into (3.15), we obtain

$$\liminf_{k\to\infty} \left\{\frac{\log F_k(\nu k)}{k \log k}\right\} \ge (\zeta-\epsilon-\nu) - (\zeta+\epsilon-\nu)(\sigma^{-1}+\epsilon) \tag{3.17}$$

for all $\epsilon > 0$. Now let $\epsilon \downarrow 0$ to obtain the required lower bound. The proof is complete. □

Proof of Theorem 4. Using the stationarity of the sequence Γ, we have that

$$\rho_k \le \mu\big(\Gamma_i \ne \varnothing \text{ for } 1 < i \le k\big) = \pi_{k-1}. \tag{3.18}$$

This provides an upper bound for $\log \rho_k$ of the required order. We note that

$$\rho_k = \pi_{k-1} - 2\pi_k + \pi_{k+1}, \tag{3.19}$$

by using the stationarity of Γ; however, we shall not make use of this fact. (Equation (3.19) is shown as follows. Denote by $u\,0^{k-1}\,v$, for $u, v = 0, 1$, the event that $1_{\{\Gamma_1=\varnothing\}} = u$, $\Gamma_2 \neq \varnothing, \ldots, \Gamma_k \neq \varnothing$, $1_{\{\Gamma_{k+1}=\varnothing\}} = v$. Now,

$$\rho_k = \mu(1\,0^{k-1}\,1) = \mu(\cdot\,0^{k-1}\,\cdot) - \mu(0\,0^{k-1}\,\cdot) - \mu(\cdot\,0^{k-1}\,0) + \mu(0\,0^{k-1}\,0)$$

where a dot in position j indicates no constraint on Γ_j. Equation (3.19) follows using stationarity.)

Turning to lower bounds for ρ_k, we remark first that, when $s_1 = 2$, we have that $2 \in \Gamma_1 \cup \Gamma_{k+1}$ if k is odd; therefore $\rho_k = 0$ in this case. Suppose now that $s_1 > 2$ (a similar argument holds if $s_1 = 2$ and k is even). Basically we follow the relevant part of the proof of Theorem 3. First,

$$\rho_k = \mu\big(\Gamma_i \neq \varnothing \text{ for } 1 < i \leq k \,\big|\, A\big)\mu(A) \tag{3.20}$$

where $A = \{\Gamma_1 = \Gamma_{k+1} = \varnothing\}$. Now,

$$\mu(A) = \bigg\{\prod_{r:\,s_r|k}\Big(1 - \frac{1}{s_r}\Big)\bigg\}\bigg\{\prod_{r:\,s_r\nmid k}\Big(1 - \frac{2}{s_r}\Big)\bigg\} \geq \prod_r \Big(1 - \frac{2}{s_r}\Big). \tag{3.21}$$

The last quantity is strictly positive since $s_1 > 2$.

We now compute a lower bound for the conditional probability in (3.20). Conditioning on A amounts to conditioning on the event that $X_r \neq 1, k+1 \pmod{s_r}$ for all r. Under A, the X_r are conditionally independent with distributions which are similar to their unconditional distributions. The argument of the previous proof now goes through, with $\nu = 0$, and with H_k replaced by

$$H_k = \big\{j : 2 \leq j \leq k,\ r \in \Gamma_j \text{ for some } 1 \leq r \leq R\big\},$$

and other minor changes arising from the altered distributions of the X_r. The conclusion is, as in (3.17), that

$$\liminf_{k\to\infty}\left\{\frac{\log \mu\big(\Gamma_i \neq \varnothing \text{ for } 1 < i \leq k \,\big|\, A\big)}{k \log k}\right\} \geq -\zeta(\sigma^{-1} - 1).$$

This may be combined with (3.21), as required. □

Appendix 1

We present here a sketch proof of (1.1), which was given in [2] but with the constant $\pi^2/12$ unfortunately replaced by $\pi^2/6$. The proof is elementary, but there is some value in giving brief details.

Let $s_i = p_i^2$, the square of the ith prime, let t_i be the ith square-free number, and write

$$\zeta = \prod_{i=1}^{\infty}\left(1 - \frac{1}{s_i}\right) = \frac{6}{\pi^2}.$$

Let $\epsilon > 0$, and pick R such that

$$\prod_{i=1}^{R}\left(1 - \frac{1}{s_i}\right) \le \zeta + \epsilon,$$

implying that the numbers $s_1, s_2, \ldots, s_R$ divide at least a proportion $1 - \zeta - \epsilon$ of integers. Let k be an integer which is larger than $P = \prod_{i=1}^{R} s_i$. For an integer m, write $x_m(1), x_m(2), \ldots, x_m(M)$ for the increasing subsequence of $m+1, m+2, \ldots, m+k$ containing all integers not divisible by any s_i (for $1 \le i \le R$), noting that

$$0 \le M - \zeta P\lfloor k/P\rfloor \le P + \epsilon P\lfloor k/P\rfloor. \tag{A.1}$$

For fixed large k, what is the smallest value of m such that

$$x_m(i) \equiv 0 \mod s_{R+i}, \quad \text{for } 1 \le i \le M? \tag{A.2}$$

By (A.1) and the Chinese Remainder Theorem, these congruences have a solution for some m satisfying $1 \le m \le n$, if $n = \prod_{i=1}^{N} s_i$ and

$$N \ge R + (\zeta + \epsilon)P\lfloor k/P\rfloor + P. \tag{A.3}$$

Now

$$\log\left\{\prod_{i=1}^{N} s_i\right\} = 2\sum_{i=1}^{N}\log p_i \sim 2N\log N \quad \text{as } N \to \infty,$$

by [13, Thm 420]. Therefore, by (A.3), we may take n such that

$$\log n = 2(1 + \mathrm{o}(1))(\zeta + \epsilon)k\log k \quad \text{as } k \to \infty.$$

With $n = n(k)$ given thus, there exists by (A.2) a square-free number t_i satisfying $t_i \le n$ and $t_{i+1} - t_i > k$. This implies that

$$t_{i+1} - t_i > (1 + \mathrm{o}(1))\frac{1}{2(\zeta + \epsilon)}\frac{\log t_i}{\log\log t_i} \quad \text{for infinitely many } i \tag{A.4}$$

as required for (1.1).

We have a further note, relevant to the appearances in (1.5) and (A.4) of the respective constants $\pi^2/6$ and $\pi^2/12$. The calculation above implies that

$$t_{i+1} - t_i > k \quad \text{for some} \quad t_i \le \exp\bigl\{2(1+\epsilon)\zeta k \log k\bigr\}. \tag{A.5}$$

In fact there are many such t_i, for the following reason. The congruences (A.2) may be replaced by

$$x_m(i) \equiv 0 \mod s_{R+j(i)}, \quad \text{for } 1 \le i \le M \tag{A.6}$$

where $j = (j(1), j(2), \ldots, j(M))$ is any given permutation of $(1, 2, \ldots, M)$. By counting the number of such permutations, one finds after a little work that there are at least $\exp\{(1-\epsilon)\zeta k \log k\}$ integers m satisfying $m \le \exp\{2(1+\epsilon)\zeta k \log k\}$ such that the sets $\{m+1, m+2, \ldots, m+k\}$ are disjoint and contain no square-free number. If these numbers m were 'uniformly spread' over the integers between 1 and $\exp\bigl\{2(1+\epsilon)\zeta k \log k\bigr\}$, then the first would be smaller than $\exp\bigl\{(1+\epsilon)\zeta k \log k\bigr\}$. In fact, we do not know how these numbers m are distributed. In the averaging process of (1.4) however, it is the 'average' gap which manifests itself, and this accounts for the value $\pi^2/6$ appearing in (1.5) and Theorem 1.

Appendix 2

We take this opportunity to correct two minor omissions in [8]. Firstly, on page 2, each set $\mathcal{G}_i$ is the set of all subsets of the label set $\mathcal{G}$. Secondly, in (2.13) on page 6, we assume further that $t_j < \frac{1}{2} s_j$. The final display on that page becomes

$$\frac{1}{n} |I \cap \{1, 2, \ldots, n\}| \le \frac{1}{n} \sum_{\substack{j:\ j \ge R \\ s_j \le 2n}} \left(1 + \frac{n}{s_j}\right) t_j \le 3 \sum_{j \ge R} \frac{t_j}{s_j} \le 3\epsilon.$$

Acknowledgements

The author is grateful to Maury Bramson for taking an interest in the problem, and for making a suggestion which led to the current note. In addition, Richard Hall has continued to give invaluable advice on number-theoretic matters. The author benefitted from a discussion with Pál Erdős of the results of [2]. The work was aided by partial support from the European Union under contract CHRX–CT93–0411. It was completed during a visit by the author to the Department of Mathematics at the University of Utah.

References

1. Billingsley, P. (1968). *Weak Convergence of Probability Measures.* John Wiley & Sons, New York.
2. Erdős, P. (1951). Some problems and results in elementary number theory. *Publicationes Mathematicae Universitatis Debreceniensis* **2**, 103–109.
3. Erdős, P. (1966). On the difference of consecutive terms of sequences defined by divisibility properties. *Acta Arithmetica* **12**, 175–182.
4. Filaseta, M. (1993). On the distribution of gaps between squarefree numbers. *Mathematika* **40**, 88–101.
5. Filaseta, M. and Trifonov, O. (1992). On the gaps between squarefree numbers II. *Journal of the London Mathematical Society* **45**, 215–221.
6. Filaseta, M. and Trifonov, O. (1994). The distribution of fractional parts with applications to gap results in number theory (to appear).
7. Graham, S. W. (1993). Moments of gaps between k-free numbers. *Journal of Number Theory* **44**, 105–117.
8. Grimmett, G. R. (1991). Statistics of sieves and square-free numbers. *Journal of the London Mathematical Society* **43**, 1–11.
9. Grimmett, G. R. and Hall, R. R. (1991). The asymptotics of random sieves. *Mathematika* **38**, 285–302.
10. Grimmett, G. R. and Stirzaker, D. R. (1992). *Probability and Random Processes.* Second edition. Oxford University Press, Oxford.
11. Hall, R. R. (1982). Squarefree numbers on short intervals. *Mathematika* **29**, 7–17.
12. Hall, R. R. (1989). The distribution of squarefree numbers. *Journal für die Reine und Angewandte Mathematik* **394**, 107–111.
13. Hardy, G. H. and Wright, E. M. (1979). *An Introduction to the Theory of Numbers.* Fifth edition. Oxford University Press, Oxford.
14. Hooley, C. (1973). On the distribution of square-free numbers. *Canadian Journal of Mathematics* **25**, 1216–1223.
15. Huxley, M. N. (1995). Moments of differences between square-free numbers (to appear).
16. Mirsky, L. (1948). Arithmetical pattern problems connected with r-free numbers. *Proceedings of the London Mathematical Society* **50**, 497–508.
17. Szemerédi, E. (1973). On the difference of consecutive terms of sequences defined by divisibility properties, II. *Acta Arithmetica* **23**, 359–361.