Asymptotic Semismoothness Probabilities

Eric Bach [*] *René Peralta* [†]

Abstract

We call an integer *semismooth* with respect to y and z if each of its prime factors is $\leq y$, and all but one are $\leq z$. Such numbers are useful in various factoring algorithms, including the quadratic sieve. Let $G(\alpha, \beta)$ be the asymptotic probability that a random integer n is semismooth with respect to n^β and n^α. We present new recurrence relations for G and related functions. We then give numerical methods for computing G, tables of G, and estimates for the error incurred by this asymptotic approximation.

1991 *Mathematics Subject Classification.* Primary 11N25; Secondary 11Y05, 11Y70.

[*]Computer Sciences Department, University of Wisconsin–Madison, 1210 W. Dayton St., Madison, WI 53706. E-mail: bach@cs.wisc.edu. Supported in part by NSF Grants DCR-8552596 and CCR-9208639.

[†]Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, P.O. Box 784, Milwaukee, WI 53201. E-mail: peralta@cs.uwm.edu. Supported in part by NSF Grant CCR-9207204.

1 Introduction

Many number-theoretic algorithms, such as the quadratic sieve factoring method [19], rely on auxiliary numbers whose prime factors lie within prescribed bounds. In practice, one often uses so-called "large prime" versions of these algorithms, in which the auxiliary numbers are composed of one moderately large prime factor and a number of smaller ones. In analyzing these, it is useful to know the asymptotic probability that a random number has this form. In this paper, we show how to compute this probability quickly and accurately and assess the accuracy of our asymptotic approximations.

Following Knuth and Trabb Pardo [8], we factor a positive integer n uniquely as $n = n_1 n_2 \ldots$, where each n_i is prime, and $n_1 \geq n_2 \geq \ldots$. In other words, n_i is the ith largest prime factor of n, counted by multiplicity. If i is greater than the number of prime factors, we define n_i to be 1.

We will say that n is *semismooth* with respect to y and z if $n_1 \leq y$ and $n_2 \leq z$. That is, all the prime factors of n are bounded by z, with the possible exception of a prime factor bounded by y. We let

$$\Psi(x, y, z) = \#\{n \leq x : n_1 \leq y, n_2 \leq z\}.$$

This generalizes de Bruijn's function [3]

$$\Psi(x, y) = \#\{n \leq x : n_1 \leq y\}.$$

We will prove that for every α, β satisfying $0 < \alpha < \beta < 1$,

$$G(\alpha, \beta) = \lim_{x \to \infty} \Psi(x, x^\beta, x^\alpha)/x \tag{1.1}$$

exists. This should be thought of as the asymptotic joint distribution of the relative lengths of n_1 and n_2. Thus, the function $\sigma(u, v) = G(1/u, 1/v)$ can be considered a two-dimensional analog of Dickman's well-known rho function.

The function G satisfies some interesting recurrence relations. In §3 we use these to show the limit in (1.1) exists, and to estimate the rate of convergence. In §4 we discuss methods for computing G numerically and tabulate the results in §5. Finally, §6 discusses the accuracy of our asymptotic approximations.

2 Background

The *Dickman rho function* is defined for real $x \geq 0$ by the relation

$$\rho(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, \\ \frac{1}{x}\int_{x-1}^{x} \rho(t)dt & \text{otherwise.} \end{cases} \tag{2.1}$$

We also let $F(\alpha) = \rho(1/\alpha)$.

Norton [17] surveys some useful properties of the rho function, which we summarize here. First, $0 < \rho(x) \leq 1$, and

$$\rho'(x) = -\rho(x-1)/x \tag{2.2}$$

when $x \geq 1$ (at $x = 1$ we take the right derivative). This implies that ρ is non-increasing, and $|\rho'(x)| \leq 1$. In fact, the rho function decreases very rapidly for large x; we have $\rho(x) \leq 1/x!$.

The differential-delay equation (2.2) implies that ρ is piecewise analytic. More precisely, there is an analytic function ρ_k agreeing with $\rho(x)$ when $k-1 \leq x \leq k$, for $k = 1, 2, 3, \ldots$. We have, for example, $\rho_1 = 1$, and $\rho_2 = 1 - \log x$. It is also easy to see that ρ belongs to the class C^k on the interval $[k, \infty)$.

Let $\pi(x)$ denote the number of primes $\leq x$, and let $\mathrm{li}(x) = \int_0^x dt/\log t$ (the Cauchy principal value is intended here). We will use the prime number theorem, in the form

$$\pi(x) = \mathrm{li}(x) + O(\frac{x}{\log^c x}); \tag{2.3}$$

this relation holds for any $c > 0$. We write $\epsilon(x)$ for the error term, so that $\pi(x) = \mathrm{li}(x) + \epsilon(x)$. Schoenfeld proved, assuming the Riemann hypothesis, that

$$|\epsilon(x)| < (\sqrt{x}\log x)/(8\pi) \tag{2.4}$$

provided $x \geq 2,657$. (See (6.18) of [22].)

The prime number theorem implies

$$\sum_{p<x} \frac{1}{p} = \log\log x + O(1) \tag{2.5}$$

and

$$\sum_{p<x} \frac{1}{p\log p} = O(1). \tag{2.6}$$

Let $0 < \alpha < 1$. Results of de Bruijn imply that if $0 < \alpha \leq \gamma$ and $t^\alpha \geq 2$, we have

$$\Psi(t, t^\gamma) = tF(\gamma) + O(\frac{t}{\alpha \log t}). \tag{2.7}$$

(To prove this, combine (1.4) and (5.3) of [3] with (2.3) above, taking $c = 4$.)

In results such as the above, an unadorned *"O"* symbol indicates an absolute constant.

3 Recurrence Relations for Smoothness Distributions

Many of the useful properties of asymptotic smoothness distributions can be derived from a simple heuristic model, which we call *random bisection.* The idea is that asymptotically, the relative lengths of the prime factors of a random number can be obtained by choosing a random λ uniformly from $(0, 1)$ – this gives the relative length of the first factor – and then proceeding recursively with the smaller interval $(0, 1-\lambda)$. (This was previously applied to prime factorizations in [1].)

To illustrate, we derive a recurrence for $F(\alpha)$, the asymptotic probability that none of n's prime factors exceed n^α. This is the probability that all lengths chosen by random bisection are $\leq \alpha$; conditioning on the first length λ, we should have

$$F(\alpha) = \int_0^\alpha F(\frac{\alpha}{1-\lambda}) d\lambda. \tag{3.1}$$

This is equivalent to (2.1), as the substitutions $t = (1-\lambda)/\alpha$ and $x = 1/\alpha$ show.

Using a similar argument, one can deduce that $F_2(\alpha)$, the asymptotic probability that $n_2 \leq n^\alpha$, should satisfy

$$F_2(\alpha) = \int_0^\alpha F_2(\frac{\alpha}{1-\lambda}) d\lambda + \int_\alpha^1 F(\frac{\alpha}{1-\lambda}) d\lambda. \tag{3.2}$$

(Compare with (3.8) and (3.11) of [8].)

Now, we let $G(\alpha, \beta)$ denote the asymptotic probability that $n_2 \leq n^\alpha$ and $n_1 \leq n^\beta$. Again by conditioning on the first length λ, we conclude that G, if it exists, should satisfy

$$G(\alpha, \beta) = \int_0^\alpha G(\frac{\alpha}{1-\lambda}, \frac{\beta}{1-\lambda}) d\lambda + \int_\alpha^\beta F(\frac{\alpha}{1-\lambda}) d\lambda. \tag{3.3}$$

We will prove this rigorously below, using a different relation for G that is not as easy to motivate:

$$G(\alpha,\beta) = F(\alpha) + \int_\alpha^\beta F(\frac{\alpha}{1-\lambda})\frac{d\lambda}{\lambda}. \tag{3.4}$$

We can, however, give it a probabilistic interpretation. We condition on the *largest* length λ produced by random bisection. Either $\lambda \le \alpha$ (which accounts for the term $F(\alpha)$), or it lies between α and β. The second event contributes a term

$$\int_\alpha^\beta \Pr[\lambda_{(2)} \le \alpha | \lambda_{(1)} = \lambda] dF(\lambda).$$

(Here $\lambda_{(1)} > \lambda_{(2)} > \cdots$ are the lengths produced by random bisection, in sorted order.) The distribution of $\lambda_{(1)}$ is absolutely continuous; from (2.2), we get

$$dF(\lambda) = F(\frac{\lambda}{1-\lambda})\frac{d\lambda}{\lambda}.$$

Because (3.4) holds for arbitrary $\alpha \le \beta$, a standard theorem of analysis (see [23, p. 360]) implies that we can take

$$\Pr[\lambda_{(2)} \le \alpha | \lambda_{(1)} = \lambda] = F(\frac{\alpha}{1-\lambda})/F(\frac{\lambda}{1-\lambda}). \tag{3.5}$$

So far, we have relied on heuristic arguments. We now prove (3.4) and (3.3).

Theorem 3.1 *If $0 < \alpha < \beta < 1$, then*

$$\Psi(x, x^\beta, x^\alpha) = xF(\alpha) + x\int_\alpha^\beta F(\frac{\alpha}{1-\lambda})\frac{d\lambda}{\lambda} + O(\frac{\log(\alpha^{-1})}{\alpha(1-\beta)}\frac{x}{\log x}).$$

Therefore, the limit

$$G(\alpha,\beta) = \lim_{x\to\infty} \Psi(x, x^\beta, x^\alpha)/x$$

exists, and satisfies (3.4).

Proof. The basic idea of the proof is to carefully repeat the conditioning argument for (3.4), employing a uniform estimate for the Ψ function and the prime number theorem.

We have

$$\Psi(x, x^\beta, x^\alpha) = \sum_{p\le x^\alpha} \#\{n \le x : n_1 = p\} + \sum_{x^\alpha < p \le x^\beta} \#\{n \le x : n_1 = p, n_2 \le x^\alpha\}. \tag{3.6}$$

For the first sum, we have

$$\sum_{p\le x^\alpha} \#\{n\le x : n_1 = p\} = \#\{n\le x : n_1\le x^\alpha\} = xF(\alpha) + O(\frac{x}{\alpha\log x}).$$

The second sum requires more work. We first observe that

$$\sum_{x^\alpha<p\le x^\beta} \#\{n\le x : n_1 = p, n_2\le x^\alpha\} = \sum_{x^\alpha<p\le x^\beta} \#\{m\le x/p : m_1\le x^\alpha\}$$

$$= \sum_{x^\alpha<p\le x^\beta} \#\{m\le x/p : m_1\le (x/p)^{\frac{\alpha}{1-\log p/\log x}}\}. \qquad (3.7)$$

When $x^\alpha < p \le x^\beta$,

$$0\le\alpha<\frac{\alpha}{1-\alpha}<\frac{\alpha}{1-\log p/\log x}\le\frac{\alpha}{1-\beta}.$$

The estimate (2.7) applies, so

$$\sum_{x^\alpha<p\le x^\beta} \#\{m\le x/p : m_1\le (x/p)^{\frac{\alpha}{1-\log p/\log x}}\}$$

$$= \sum_{x^\alpha<p\le x^\beta} \frac{x}{p}F(\frac{\alpha}{1-\log p/\log x}) + O(\frac{1}{\alpha}\sum_{x^\alpha<p\le x^\beta}\frac{x/p}{\log(x/p)}).$$

Applying (2.5) and (2.6), we get

$$\Psi(x,x^\beta,x^\alpha) = xF(\alpha) + \sum_{x^\alpha<p\le x^\beta}\frac{x}{p}F(\frac{\alpha}{1-\log p/\log x}) + O(\frac{\log(\alpha^{-1})}{\alpha(1-\beta)}\frac{x}{\log x}). \qquad (3.8)$$

Using Stieltjes integration, we have

$$\sum_{x^\alpha<p\le x^\beta}\frac{1}{p}F(\frac{\alpha}{1-\log p/\log x}) = \int_{x^\alpha}^{x^\beta}\rho(\frac{1-\log t/\log x}{\alpha})\frac{d\pi(t)}{t}. \qquad (3.9)$$

If we integrate by parts, substitute $\pi(t) = \mathrm{li}(t) + \epsilon(t)$, and recombine the terms involving $\mathrm{li}(t)$, we obtain

$$\int_{x^\alpha}^{x^\beta}\rho(\frac{1-\log t/\log x}{\alpha})\frac{d\pi(t)}{t} = \int_{x^\alpha}^{x^\beta}\rho(\frac{1-\log t/\log x}{\alpha})\frac{dt}{t\log t}$$

$$+\Big[\rho(\frac{1-\log t/\log x}{\alpha})\frac{\epsilon(t)}{t}\Big]_{x^\alpha}^{x^\beta} - \int_{x^\alpha}^{x^\beta}\frac{d}{dt}(\frac{1}{t}\rho(\frac{1-\log t/\log x}{\alpha}))\epsilon(t)dt. \qquad (3.10)$$

We now show the error terms in (3.10) are small. Using $|\rho|\le 1$ and (2.3) (with $c=1$), we obtain

$$\Big[\rho(\frac{1-\log t/\log x}{\alpha})\frac{\epsilon(t)}{t}\Big]_{x^\alpha}^{x^\beta} = O(\frac{1}{\alpha\log^2 x}).$$

After differentiating the quotient and estimating each resulting term separately, we get

$$\int_{x^\alpha}^{x^\beta} \frac{d}{dt}(\frac{1}{t}\rho(\frac{1-\log t/\log x}{\alpha}))\epsilon(t)dt = O(\int_{x^\alpha}^{x^\beta} \frac{dt}{t\log^2 t}) = O(\frac{1}{\alpha \log x}).$$

This shows that

$$\Psi(x, x^\beta, x^\alpha) = xF(\alpha) + \int_{x^\alpha}^{x^\beta} \rho(\frac{1-\log t/\log x}{\alpha})\frac{dt}{t\log t} + O(\frac{\log(\alpha^{-1})x}{\alpha(1-\beta)\log x}).$$

Making the substitution $\lambda = \log t/\log x$, we obtain the first statement of the theorem. The second follows from dividing by x and letting $x \to \infty$. $\square$

The novelty in the above theorem is a careful estimate of the error term. Knuth and Trabb Pardo gave (3.4) for the special case $\beta = 2\alpha$. Weaker statements of Theorem 3.1 (that is, without error estimates) appear in [9] and [15]. We now prove (3.3), which we believe to be new.

Theorem 3.2 *We have*

$$G(\alpha,\beta) = \int_0^\alpha G(\frac{\alpha}{1-\lambda}, \frac{\beta}{1-\lambda})d\lambda + \int_\alpha^\beta F(\frac{\alpha}{1-\lambda})d\lambda.$$

Proof. If $0 < \gamma < 1$, we have

$$F(\gamma) = \int_0^\gamma F(\frac{\gamma}{1-\zeta})d\zeta.$$

Now substitute $\zeta = \lambda/(1-\nu)$ and $\gamma = \alpha/(1-\nu)$, and rearrange terms to obtain

$$F(\frac{\alpha}{1-\nu}) = \int_0^\alpha F(\frac{\alpha}{1-\nu-\lambda})d\lambda + \nu F(\frac{\alpha}{1-\nu}).$$

If we divide this by ν, integrate over $\alpha \le \nu \le \beta$, reverse the order of integration, and substitute $\nu = (1-\lambda)\mu$, we get

$$\int_\alpha^\beta F(\frac{\alpha}{1-\nu})\frac{d\nu}{\nu} = \int_0^\alpha d\lambda \int_{\frac{\alpha}{1-\lambda}}^{\frac{\beta}{1-\lambda}} F(\frac{\alpha}{(1-\lambda)(1-\mu)})\frac{d\mu}{\mu} + \int_\alpha^\beta F(\frac{\alpha}{1-\nu})d\nu.$$

If we add $F(\alpha) = \int_0^\alpha F(\alpha/(1-\lambda))d\lambda$ to both sides and apply (3.4), we get

$$G(\alpha,\beta) = F(\alpha) + \int_\alpha^\beta F(\frac{\alpha}{1-\nu})\frac{d\nu}{\nu}$$

$$= \int_0^\alpha d\lambda \Big[\int_{\frac{\alpha}{1-\lambda}}^{\frac{\beta}{1-\lambda}} F(\frac{\alpha}{(1-\lambda)(1-\mu)})\frac{d\mu}{\mu} + F(\frac{\alpha}{(1-\lambda)})\Big] + \int_\alpha^\beta F(\frac{\alpha}{1-\nu})d\nu$$

$$= \int_0^\alpha G(\frac{\alpha}{1-\lambda}, \frac{\beta}{1-\lambda})d\lambda + \int_\alpha^\beta F(\frac{\alpha}{1-\lambda})d\lambda. \qquad \square$$

4 Numerical Methods

Several authors have discussed computing smoothness distributions such as the Dickman rho function. We briefly discuss this work and then present our numerical methods for the semismoothness distribution G.

Implicit in the random bisection idea is the notion that smoothness distributions can be computed by Monte Carlo methods. This was done for the rho function by Chamayou [5], albeit with a different probabilistic model than ours. Although one could also approximate G by simulation, we have not done this because the probabilities of current interest are so small.

It is also possible to combine a recurrence relation with numerical integration. This was done by van de Lune and Wattel [13] and Knuth and Trabb Pardo [8]. For example, replacing the integral in (2.1) with an appropriate quadrature rule gives a linear equation that can be solved to obtain an approximation to $\rho(x)$. Either of the relations (3.3) and (3.4) can be used in this way to compute G. In practice, however, we were dissatisfied with the performance of the resulting methods. Use of the recurrence relation (3.3) involves computing values of G in a two-dimensional region and interpolating the values on a line of integration. The relation (3.4) is more useful, as it only relies on values of F (i.e., ρ); however, one needs an accurate table of this function before numerical integration is feasible.

The best methods for calculating ρ are based on the following idea. Recall that there is an analytic function ρ_k that agrees with ρ on the interval $[k-1, k]$. Knowing the Taylor series for ρ_k, one can use (2.2) to get the Taylor series for ρ_{k+1} up to a constant term, which can be then determined from (2.1). This was used by Cheer and Goldston [6], Marsaglia, Zaman, and Marsaglia [14], and Patterson and Rumsey [18] to evaluate ρ and similar functions.

To compute G, we used Patterson and Rumsey's method for ρ, which we summarize as follows. (Its derivation is similar to §3 of [6].) Let $0 \leq \xi \leq 1$. Define coefficients $c_i^{(k)}$ by

$$\rho_k(k-\xi) = \sum_{i=0}^{\infty} c_i^{(k)} \xi^k, \qquad k = 1, 2, \ldots .$$

Then we have

$$c_0^{(1)} = 1, \quad c_i^{(0)} = 0 \text{ for } i \geq 1,$$

$$c_0^{(2)} = 1 - \log 2, \quad c_i^{(2)} = 1/(i2^i) \text{ for } i \geq 1, \tag{4.1}$$

and for $k > 2$

$$c_i^{(k)} = \sum_{j=0}^{i-1} \frac{c_j^{(k-1)}}{ik^{i-j}}, \tag{4.2}$$

with

$$c_0^{(k)} = \frac{1}{k-1} \sum_{j=1}^{\infty} \frac{c_j^{(k)}}{j+1}. \tag{4.3}$$

It can be shown that $0 \leq c_i^{(k)} \leq 1/2^i$, so that $m+1$ terms of the series will approximate ρ_k within an (absolute) error of 2^{-m}. Empirically, we found that 55 coefficients were enough to compute ρ to IEEE standard double precision (relative error about 10^{-17}) in the range $0 \leq x \leq 20$.

Although this suffices for our purposes, we remark that the method of [14] is superior when one wishes to compute $\rho(x)$ to high precision. It expands ρ_k in circles of radius 1/2 about $k - 1/2$, using simpler recurrences than (4.1)–(4.3). (We do call attention to one oversight in [14]: the authors state that "$\rho(x)$ behaves asymptotically like $x^{\alpha x}$," and provide data suggesting that $\alpha \doteq -1.18$. However, de Bruijn [4] proved that $\log \rho(x) \sim -x \log x$ as $x \to \infty$, so $\alpha = -1$.)

Our method for computing G uses (3.4), together with term-by-term integration of the Taylor series determined by (4.1)–(4.3). Rather than use (3.4) directly, it is more convenient to work with $\sigma(u, v) = G(1/u, 1/v)$, which satisfies

$$\sigma(u, v) = \rho(u) + \int_v^u \rho(u - u/t) \frac{dt}{t}.$$

(To prove this, make the substitutions $\alpha = 1/u$, $\beta = 1/v$, and $\lambda = 1/t$ in (3.4).) We define

$$J(u, v, w) = \int_v^u \rho(w - w/t) \frac{dt}{t},$$

so that

$$\sigma(u, v) = \rho(u) + J(u, v, u).$$

We now show how to compute $J(u, v, w)$. Let $k = \lceil w - w/u \rceil$, and define $\xi(t)$ by $w - w/t = k - \xi(t)$.

If $\xi(t) \in [0,1]$ for $v \le t \le u$, we can proceed as follows:

$$J(u,v,w) = \int_v^u \rho(k-\xi(t))\frac{dt}{t} = \sum_{i=0}^{\infty} c_i^{(k)} \int_v^u \xi(t)^i \frac{dt}{t}$$

$$= \sum_{i=0}^{\infty} c_i^{(k)} \int_{w/u}^{w/v} \frac{(\eta+k-w)^i}{\eta} d\eta.$$

(Here we have substituted $\eta = w/t$.) If $H_i(u,v,w) = \int_{w/u}^{w/v} \frac{(\eta+k-w)^i}{\eta} d\eta$, then writing $(\eta+k-w)^i/\eta$ as $(\eta+k-w)^{i-1} + (\eta+k-w)^{i-1}(k-w)/\eta$ gives

$$H_i(u,v,w) = \begin{cases} \log(u/v) & \text{if } i=0, \\ \frac{(w/v+k-w)^i-(w/u+k-w)^i}{i} + (k-w)H_{i-1}(u,v,w) & \text{otherwise.} \end{cases}$$

Solving the recurrence yields

$$H_i(u,v,w) = C^i \left[\log(u/v) + \sum_{j=1}^{i} \frac{(A/C)^j}{j} - \sum_{j=1}^{i} \frac{(B/C)^j}{j} \right], \tag{4.4}$$

where $A = w/v + k - w$, $B = w/u + k - w$, and $C = k - w$.

Thus, in the case $\xi(t) \in [0,1]$, we have

$$J(u,v,w) = \sum_{i=0}^{\infty} c_i^{(k)} H_i(u,v,w), \tag{4.5}$$

where the H_i's are defined by (4.4). If $\xi(t) \notin [0,1]$, we must split the integral. Note that when $t = w/(w-k+1)$, we have $\xi(t) = 1$, that is, $w - w/t = k - 1$. In this case, we have

$$\begin{aligned} J(u,v,w) &= \int_v^{w/(w-k+1)} \rho(w-w/t)\frac{dt}{t} + \int_{w/(w-k+1)}^u \rho(w-w/t)\frac{dt}{t} \\ &= J(w/(w-k+1),v,w) + J(u,w/(w-k+1),w). \end{aligned}$$

The second integral can be computed via (4.5) and the first integral is computed recursively. We note that the integral is split if and only if $v < w/(w-k+1)$.

We can bound the recursion depth for computing $J(u,v,u)$ by observing that at the ith recursive step, u is replaced by $u/(1+i-r)$, where $r = \lceil u \rceil - u$. (This can be verified by induction on i.) From this it can be seen that the integral is split no more than u/v times.

If we approximate $J(u,v,w)$ by n terms of the series (4.5), the tail is bounded by

$$\begin{aligned}
\sum_{i=n+1}^{\infty} c_i^{(k)} \int_v^u \xi(t)^i \frac{dt}{t} &\le \sum_{i=n+1}^{\infty} c_i^{(k)} \int_v^u \frac{dt}{t} \\
&= \log(u/v) \sum_{i=n+1}^{\infty} c_i^{(k)} \\
&\le \log(u/v) \sum_{i=n+1}^{\infty} (1/2)^i \\
&= \frac{\log(u/v)}{2^n}
\end{aligned}$$

when $\xi(t) \in [0,1]$. When computing $J(u,v,u)$, the integral is split into at most u/v pieces, so the total error is at most $\frac{u \log(u/v)}{v 2^n}$.

Some care is required in the computation of $H_i(u,v,w)$, because massive cancellation occurs in (4.4) when i is large. We deal with this in the following way. In any recursive call (i.e., not the top level), it can be shown that $0 \le A \le 1$, $B = 0$, and $C < -1$. For these cases, we replace (4.4) by the convergent series

$$H_i(u,v,w) = \sum_{j=i+1}^{\infty} \frac{B^j}{C^{j-i} j} - \sum_{j=i+1}^{\infty} \frac{A^j}{C^{j-i} j}. \tag{4.6}$$

At the top level, A/C and B/C are unbounded, and naive use of (4.4) can lead to overflow whenever C is close to 0. Here, in our calculations we replaced (4.4) by the equivalent form

$$H_i(u,v,w) = C^i \log(u/v) + \sum_{j=1}^{i} \frac{A^j}{C^{j-i} j} - \sum_{j=1}^{i} \frac{B^j}{C^{j-i} j} \tag{4.7}$$

whenever $|C| < 0.2$.

5 Tables

In this section we give tables of the asymptotic semismoothness distribution, computed with the methods of §4. Our calculations used 22 terms of the Taylor series for $\rho_k(x)$ and 22 terms in the expansion given by (4.5).

As a check on our computations we used an independent computation of $G(\alpha, 2\alpha)$ (using (3.2) and numerical integration), as well as Table 1 in [8]. This table includes values of $G(\alpha, \alpha) = \rho(\alpha^{-1})$, as well as values of $G(\alpha, 2\alpha)$. Our results agree with [8] to seven significant figures.

Table 1 shows $\sigma(u, v) = G(\frac{1}{u}, \frac{1}{v})$ for u, v in the range $2 \leq u \leq 20$ and $2 \leq v \leq 10$.

u	v								
	2.0	**3.0**	**4.0**	**5.0**	**6.0**	**7.0**	**8.0**	**9.0**	**10.0**
2.0	3.068528e-01	—–	—–	—–	—–	—–	—–	—–	—–
3.0	2.246518e-01	4.860839e-02	—–	—–	—–	—–	—–	—–	—–
4.0	9.639901e-02	2.465561e-02	4.910926e-03	—–	—–	—–	—–	—–	—–
5.0	3.079212e-02	6.144568e-03	1.849280e-03	3.547247e-04	—–	—–	—–	—–	—–
6.0	8.511187e-03	1.092267e-03	3.127192e-04	1.051674e-04	1.964970e-05	—–	—–	—–	—–
7.0	2.184024e-03	1.596965e-04	3.754974e-05	1.317947e-05	4.778139e-06	8.745670e-07	—–	—–	—–
8.0	5.297043e-04	2.058327e-05	3.662652e-06	1.179057e-06	4.696284e-07	1.796314e-07	3.232069e-08	—–	—–
9.0	1.221795e-04	2.418992e-06	3.097157e-07	8.573660e-08	3.319264e-08	1.439810e-08	5.732620e-09	1.016248e-09	—–
10.0	2.684198e-05	2.633999e-07	2.352672e-08	5.382861e-09	1.915717e-09	8.358903e-10	3.855071e-10	1.583472e-10	2.770172e-11
11.0	5.627512e-06	2.679280e-08	1.637888e-09	3.019323e-10	9.556196e-11	3.985501e-11	1.890085e-11	9.128686e-12	3.844123e-12
12.0	1.128672e-06	2.559021e-09	1.057229e-10	1.544758e-11	4.254912e-12	1.647475e-12	7.654740e-13	3.859215e-13	1.932249e-13
13.0	2.170148e-07	2.304231e-10	6.373298e-12	7.303377e-13	1.725544e-13	6.086492e-14	2.698020e-14	1.355061e-14	7.158754e-15
14.0	4.009759e-08	1.962928e-11	3.606489e-13	3.218087e-14	6.459110e-15	2.048898e-15	8.518898e-16	4.157376e-16	2.213880e-16
15.0	7.134198e-09	1.587081e-12	1.923443e-14	1.329397e-15	2.252007e-16	6.367265e-17	2.454356e-17	1.145796e-17	6.005714e-18
16.0	1.224713e-09	1.221450e-13	9.701521e-16	5.171612e-17	7.360498e-18	1.843517e-18	6.534072e-19	2.887008e-19	1.467795e-19
17.0	2.032238e-10	8.971949e-15	4.641835e-17	1.901399e-18	2.266007e-19	5.005962e-20	1.621962e-20	6.731238e-21	3.286696e-21
18.0	3.265002e-11	6.304945e-16	2.112631e-18	6.627200e-20	6.595740e-21	1.281259e-21	3.779199e-22	1.465048e-22	6.821566e-23
19.0	5.086645e-12	4.248307e-17	9.169125e-20	2.195713e-21	1.820808e-22	3.103008e-23	8.307233e-24	2.996162e-24	1.323455e-24
20.0	7.695287e-13	2.750199e-18	3.803595e-21	6.932224e-23	4.779992e-24	7.133404e-25	1.729532e-25	5.786581e-26	2.415504e-26

Table 1. Values of $\sigma(u, v) = G(1/u, 1/v)$ for $2 \leq u \leq 20$; $2 \leq v \leq 10$.

Of particular interest nowadays are values of $G(\alpha, \beta)$ for (α, β) near $(1/12, 1/7.5)$. This is so because recent implementations of the multiple polynomial quadratic sieve are designed to factor 100-digit cryptographic integers (i.e., products of two large primes), using auxiliary 60-digit numbers which are semismooth with respect to bounds near 10^8 and 10^5. It is believed that these auxiliary numbers are semismooth with the same probability as random numbers, so that the bulk of the algorithm's work can be viewed as a search for semismooth numbers among what are essentially random 60-digit numbers. Thus the probability of a "hit" is given by $\Psi(10^{60}, 10^8, 10^5)$, which is approximately $G(1/12, 1/7.5)$ (for details, see [10]). Most other factoring algorithms also allow for a "large prime" variation (see [15, 16]). Semismoothness tables should be of aid in choosing optimal parameters for these algorithms as well.

Table 2 gives values of $G(\alpha, \beta)$ for α and β in the current range of interest for factorization algorithms.

u	v						
	6.0	**6.5**	**7.0**	**7.5**	**8.0**	**8.5**	**9.0**
10.0	1.915717e-09	1.246194e-09	8.358903e-10	5.685742e-10	3.855071e-10	2.548590e-10	1.583472e-10
10.5	4.347011e-10	2.790060e-10	1.862869e-10	1.273550e-10	8.785696e-11	6.019803e-11	4.010645e-11
11.0	9.556196e-11	6.027060e-11	3.985501e-11	2.719972e-11	1.890085e-11	1.319996e-11	9.128686e-12
11.5	2.042300e-11	1.261597e-11	8.230457e-12	5.579976e-12	3.879851e-12	2.734419e-12	1.930373e-12
12.0	4.254912e-12	2.567571e-12	1.647475e-12	1.105464e-12	7.654740e-13	5.408660e-13	3.859215e-13
12.5	8.660935e-13	5.094235e-13	3.206906e-13	2.123547e-13	1.459152e-13	1.028856e-13	7.371953e-14
13.0	1.725544e-13	9.875063e-14	6.086492e-14	3.967965e-14	2.698020e-14	1.891864e-14	1.355061e-14
13.5	3.369911e-14	1.873623e-14	1.128729e-14	7.230573e-15	4.854135e-15	3.375809e-15	2.408918e-15
14.0	6.459110e-15	3.484569e-15	2.048898e-15	1.287611e-15	8.518898e-16	5.863181e-16	4.157376e-16
14.5	1.216276e-15	6.360275e-16	3.645830e-16	2.244685e-16	1.461324e-16	9.936033e-17	6.986140e-17
15.0	2.252007e-16	1.140545e-16	6.367265e-17	3.836313e-17	2.454356e-17	1.646200e-17	1.145796e-17

Table 2. Values of $\sigma(u, v) = G(1/u, 1/v)$ for $10 \le u \le 15$; $6 \le v \le 9$.

We observe that, in the range of Table 2, $\log(\sigma(u, v))$ is almost linear in u, v. By analogy with known approximations to Dickman's rho function we performed a least squares fit of a linear function of $u \log u, v \log v$. The resulting approximation is

$$\sigma(u, v) \approx e^{4.55219 - 0.933064 u \log u - 0.280283 v \log v}. \tag{5.1}$$

In the range of Table 2, this approximation has a relative error of no more than 30% (the error increases rapidly outside this range). The approximation also shows that $\sigma(u, v)$ is much more dependent on u than on v.

6 Error Analysis

In this section, we consider the question of how closely asymptotic distributions such as ρ and σ approximate actual smoothness probabilities. We will show, in a certain sense, that if ρ is a good approximation to the smoothness distribution, then σ is a good approximation to the semismoothness distribution.

We first consider the question of whether $x\rho(u)$ is a good approximation to

$\Psi(x, x^{1/u})$. This is of practical importance since asymptotic relations such as $x\rho(u) \sim \Psi(x, x^{1/u})$ do not guarantee that $\rho(u)$ is a good approximation to the probability of $1/u$-smoothness for any numbers of practical interest.

For example, from results of Ramaswami [20] (cited as equations 3.7 and 3.8 of [17]) and Knuth and Trabb Pardo [8], we know that

$$\Psi(x, x^{1/u}) = x\rho(u) + \frac{x(1-\gamma)\rho(u-1)}{\log x} + O(\frac{x}{\log^2 x}),$$

and therefore

$$\frac{\Psi(x, x^{1/u})}{x\rho(u)} = 1 + \frac{\rho(u-1)}{\rho(u)}\frac{(1-\gamma)}{\log x} + O(\frac{1}{\rho(u)\log^2 x}).$$

(Here, $\gamma = 0.5772...$ is Euler's constant.) The unknown part of the relative error is

$$O(\frac{1}{\rho(u)\log^2 x}). \tag{6.1}$$

Taking the crude approximation $\rho(u) \approx u^{-u}$, we note that for (6.1) to be small, we need $\log x \gg u^{u/2}$. This is not likely to be attained in practical situations; for example, if $u = 7.5$ and $x = 10^{60}$, we have $(u^{-u}\log^2 x)^{-1} = 191.5$.

On the other hand, accurate tables of $\rho(u)$ have been available for at least two decades, going well beyond the values of u needed to evaluate current factoring methods. As far as we know, no discrepancy has been observed between values of the rho function and smoothness probabilities, in the range of interest to algorithm designers. For example, Table 3 below exhibits smooth number counts found by Odlyzko (from [21]); as soon as the predicted count of smooth numbers is moderately large, one finds reasonable agreement with the rho function. (We note that Odlyzko only counted numbers whose prime power factors are small, a definition more stringent than ours.)

k	count	$u = \frac{\log(10^{15})}{\log(2^k)}$	$10^5\rho(u)$	ratio
6	0	8.305	0.001144	0.000
7	0	7.118	0.05981	0.000
8	1	6.229	0.9810	1.019
9	6	5.537	7.727	0.777
10	27	4.983	37.18	0.726
11	110	4.530	126.6	0.869
12	326	4.152	336.0	0.970
13	691	3.833	739.3	0.935
14	1425	3.559	1416	1.006
15	2416	3.322	2425	0.996
16	3852	3.114	3816	1.009
17	5691	2.931	5616	1.013
18	7979	2.768	7823	1.020

Table 3. Counts of even 2^k-smooth numbers in $[10^{15}, 10^{15} + 2 \times 10^5]$.

Therefore, we will simply take as given that ρ is a good approximation to smoothness probabilities; investigating this question further is beyond the scope of this paper. We proceed from this assumption to study the question of when $\sigma(u, v)$ is a good approximation to $\Psi(x, x^{1/u}, x^{1/v})/x$.

The following theorem states that if F is a good approximation to the smoothness distribution, then G is a good approximation to the semismoothness distribution. In this result, α and β satisfy $0 < \alpha < \beta < 1$, and $\rho(x)$ is extended to be 1 for negative numbers.

Theorem 6.1 *Assume the Riemann hypothesis. Choose c_1 and c_2 so that*

$$c_1 \leq \frac{\Psi(t, t^\gamma)}{tF(\gamma)} \leq c_2$$

whenever $\frac{\alpha}{1-\alpha} \leq \gamma \leq \frac{\alpha}{1-\beta}$ and $t \geq x^{1-\beta}$. Then, if $x^\alpha \geq 2,657$, we have

$$c_1(1 - \Delta) \leq \frac{\Psi(x, x^\beta, x^\alpha)}{xG(\alpha, \beta)} \leq c_2(1 + \Delta),$$

where

$$|\Delta| \leq \frac{\beta}{4\pi G(\alpha, \beta)} \left[2\rho(\frac{1-\beta}{\alpha}) + \frac{\rho(\frac{1-\alpha-\beta}{\alpha})}{(1-\beta)\log x} \right] \frac{\log x}{x^{\alpha/2}}. \quad (6.2)$$

Proof. From (3.6) and (3.7) we obtain

$$\Psi(x, x^\beta, x^\alpha) = \Psi(x, x^\alpha) + \sum_{x^\alpha < p \le x^\beta} \Psi(x/p, (x/p)^{\frac{\alpha}{1-\log p/\log x}}). \tag{6.3}$$

From the definition of c_2, plus (3.9) and (6.3), we have

$$\begin{aligned} c_2^{-1}\Psi(x, x^\beta, x^\alpha) &\le xF(\alpha) + x \sum_{x^\alpha < p \le x^\beta} \frac{1}{p} F(\frac{\alpha}{1 - \log p/\log x}) \\ &= xF(\alpha) + x \int_{x^\alpha}^{x^\beta} \rho(\frac{1 - \log t/\log x}{\alpha}) \frac{d\pi(t)}{t}. \end{aligned}$$

Using (3.4) and (3.10), and writing $\lambda(t) = \log t/\log x$, we have

$$c_2^{-1}\Psi(x, x^\beta, x^\alpha) \le xF(\alpha) + x\int_{x^\alpha}^{x^\beta} \rho(\frac{1-\lambda(t)}{\alpha}) \frac{dt}{t \log t} + xE_1(\alpha, \beta, x) + xE_2(\alpha, \beta, x)$$

$$= xG(\alpha, \beta) + xE_1(\alpha, \beta, x) + xE_2(\alpha, \beta, x), \tag{6.4}$$

where

$$E_1(\alpha, \beta, x) = \Big[\rho(\frac{1-\lambda(t)}{\alpha})\frac{\epsilon(t)}{t}\Big]_{x^\alpha}^{x^\beta} = \rho(\frac{1-\beta}{\alpha})\frac{\epsilon(x^\beta)}{x^\beta} - \rho(\frac{1-\alpha}{\alpha})\frac{\epsilon(x^\alpha)}{x^\alpha} \tag{6.5}$$

and

$$E_2(\alpha, \beta, x) = -\int_{x^\alpha}^{x^\beta} \frac{d}{dt}(\frac{1}{t}\rho(\frac{1-\lambda(t)}{\alpha}))\epsilon(t)dt$$

$$= \int_{x^\alpha}^{x^\beta} \epsilon(t)t^{-2}[\rho(\frac{1-\lambda(t)}{\alpha}) + \rho'(\frac{1-\lambda(t)}{\alpha})/(\alpha \log x)]dt. \tag{6.6}$$

Schoenfeld's bound (2.4) (which assumes the Riemann hypothesis) and (6.5) imply

$$|E_1| \le \frac{\log x}{8\pi}\left[\rho(\frac{1-\beta}{\alpha})\beta x^{-\beta/2} + \rho(\frac{1-\alpha}{\alpha})\alpha x^{-\alpha/2}\right] \le \frac{\log x}{4\pi}\left[\rho(\frac{1-\beta}{\alpha})\beta x^{-\alpha/2}\right]. \tag{6.7}$$

Similarly, but using (2.2) and (6.6), we have

$$|E_2| \le \int_{x^\alpha}^{x^\beta} \frac{\log t}{8\pi t^{3/2}}\left[\rho(\frac{1-\lambda(t)}{\alpha}) + \rho(\frac{1-\lambda(t)}{\alpha} - 1)/(\log x - \log t)\right] dt.$$

So far, we have assumed that $\rho(x) = 0$ when $x < 0$. If we redefine $\rho(x)$ to be 1 when $x < 0$, the inequality above still holds, and we have made ρ monotonic. Using this new extension of ρ, we find

$$|E_2| \le \frac{A}{8\pi}\int_{x^\alpha}^{x^\beta} t^{-3/2}\log t\ dt \le \frac{A\beta \log x}{8\pi}\int_{x^\alpha}^{\infty} t^{-3/2}dt = \frac{A\beta\log x}{4\pi}x^{-\alpha/2}, \tag{6.8}$$

where A denotes the expression

$$\rho(\frac{1-\beta}{\alpha}) + \rho(\frac{1-\alpha-\beta}{\alpha})\frac{1}{(1-\beta)\log x}.$$

Let

$$\Delta(\alpha,\beta,x) = \frac{(E_1+E_2)}{G(\alpha,\beta)};$$

then (6.4), the inequalities (6.7) and (6.8), and a little algebra give the upper bound in the theorem. The lower bound is proved by an entirely analogous argument, starting with the estimate $\Psi(t,t^\gamma) \geq c_1 t F(\gamma)$. $\square$

With the help of Theorem 6.1, the extra relative error incurred by using the asymptotic two-dimensional smoothness distribution can be explicitly estimated. For example, if $x \approx 10^{60}$, $\alpha = \frac{1}{12}$, and $\beta = \frac{1}{7.5}$, then (6.2) gives $|\Delta| \leq 0.062$.

We also remark that Theorem 6.1 can be improved slightly at some cost in readability. For example, the first inequalities in (6.7) and (6.8) could be used directly (note that the first integral in (6.8) can be expressed in closed form).

In a certain sense, Theorem 6.1 ascribes most of the error in the approximation

$$\Psi(x,x^\alpha,x^\beta) \approx xG(\alpha,\beta)$$

to the use of the rho function. Hildebrand [11] proved that if the Riemann hypothesis holds, then

$$\Psi(x,x^\alpha) = xF(\alpha)\left(1 + O(\frac{\log(\alpha^{-1})}{\alpha\log x})\right), \tag{6.9}$$

as $x \to \infty$. However, Theorem 6.1 and equation (3.4) imply that

$$\Delta = O(\frac{\log x}{x^{\alpha/2}}),$$

which is asymptotically much smaller than the relative error in (6.9).

7 Addendum

Following the ideas in this paper, Robert Lambert [12] has computed the asymptotic probability that a random integer $\leq x$ has exactly two prime factors between x^α and x^β, with all other prime factors $\leq x^\alpha$.

Simon Tavaré has kindly informed us that the random bisection model also plays a role in theoretical population biology. We briefly note the connection to our work. We have studied the asymptotic joint distribution of the normalized lengths of the prime factors (ordered by size) of a random integer $x \leq n$. Letting $n_1 \geq n_2 \geq n_3 \geq \cdots$ be these prime factors, the asymptotic joint distribution of

$$\frac{\log n_1}{\log n}, \frac{\log n_2}{\log n}, \frac{\log n_3}{\log n}, \ldots$$

is identical to the distribution of allele frequencies (in a model with infinitely many alleles), ranked by size. In theoretical biology this is known as the *Poisson-Dirichlet* distribution. Algorithms for computing its marginal distributions (in our terms, the distribution of the length of the kth-largest factor of a random integer), have been given by Griffiths [7]. As far as we know, however, we are the first to publish an algorithm for computing the joint distribution. For more on applications to biology, we refer the reader to [7] and references therein.

Finally, we remark that there is a very efficient algorithm for sampling from the factor length (Poisson-Dirichlet) distribution, which has been analyzed in [2].

References

[1] E. Bach. How to generate factored random numbers. *SIAM J. Comput.*, 17:179–193, 1988.

[2] E. Bach. Exact analysis of a priority queue algorithm for random variate generation. *Proc. 5th Ann. ACM-SIAM Symposium on Discrete Algorithms*, pp. 48–52, 1994.

[3] N. G. de Bruijn. On the number of positive integers $\leq x$ and free of prime factors $> y$. *Indag. Math.*, 13:50–60, 1951.

[4] N. G. de Bruijn. The asymptotic behavior of a function occuring in the theory of primes. *J. Indian Math. Soc. (N.S.)*, 15:25–32, 1951.

[5] J.-M.-F. Chamayou. A probabilistic approach to a differential-difference equation arising in analytic number theory. *Math. Comp.*, 27:197–203, 1973.

[6] A. Y. Cheer and D. A. Goldston. A differential delay equation arising from the sieve of Eratosthenes. *Math. Comp.*, 55:129–141, 1990.

[7] R. C. Griffiths. On the distribution of points in a Poisson Dirichlet process. *J. Appl. Probab.*, 25:336–345, 1988.

[8] D. Knuth and L. Trabb Pardo. Analysis of a simple factorization algorithm. *Theoret. Comput. Sci.*, 3:321–348, 1976.

[9] G. Kolesnik and E. G. Strauss. On the first occurrence of values of a character. *Trans. Amer. Math. Soc.*, 246:385–394, 1978.

[10] A.K. Lenstra and M.S. Manasse. Factoring by electronic mail. In *EUROCRYPT 89*, volume 434 of Lecture Notes in Computer Science, pages 355–371. Springer-Verlag, 1990.

[11] A. Hildebrand. Integers free of large prime factors and the Riemann hypothesis. *Mathematika*, 31:258-271, 1984.

[12] R. Lambert. Asymptotic doubly-semi-smoothness. Manuscript, February 1995.

[13] J. van de Lune and E. Wattel. On the numerical solution of a differential-difference equation arising in analytic number theory. *Math. Comp.*, 23:417–421, 1969.

[14] G. Marsaglia, A. Zaman, and J. C. W. Marsaglia. Numerical solution of some classical differential-difference equations. *Math. Comp.*, 53:191–201, 1989.

[15] P. L. Montgomery. *An FFT extension of the elliptic curve method of factorization.* Ph.D. thesis, University of California - Los Angeles, 1992.

[16] P. L. Montgomery and R. D. Silverman. An FFT extension to the $p-1$ factoring algorithm. *Math. Comp.*, 54:839–854, 1990.

[17] K. K. Norton. Numbers with small prime factors, and the least kth power non-residue. *Mem. Amer. Math. Soc.*, 106, 1971.

[18] N. Patterson. Letter to Eric Bach, November 1988.

[19] C. Pomerance. The quadratic sieve factoring algorithm. In *EUROCRYPT '84*, volume 209 of Lecture Notes in Computer Science, pages 169-182. Springer-Verlag, 1985.

[20] V. Ramaswami. The number of positive integers $< x$ and free of prime divisors $> x^c$, and a problem of S.S. Pillai. *Duke Math. J.*, 16:99–109, 1949.

[21] C.-P. Schnorr and H. W. Lenstra, Jr. A Monte Carlo factoring algorithm with linear storage. *Math. Comp.*, 43:289–311, 1984.

[22] L. Schoenfeld. Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II. *Math. Comp.*, 30:337–360, 1976.

[23] E. C. Titchmarsh, *The Theory of Functions.* Second Edition, Oxford Univ. Press, 1939.