The Number of Prime Factors on Average in Certain Integer Sequences

Rafael Jakimczuk
División Matemática
Universidad Nacional de Luján
Buenos Aires, Argentina
jakimczu@mail.unlu.edu.ar

Matilde Lalín
Département de mathématiques et de statistique
Université de Montréal
CP 6128, succ. Centre-ville
Montreal, QC H3C 3J7, Canada
mlalin@dms.umontreal.ca

Abstract

Let $\Omega(n)$ denote the total number of primes in the factorization of the integer n. We obtain asymptotic formulae for the sums $\sum_{n\leq x}\Omega(n)$, where n runs over the square-free numbers, the square-full numbers, the perfect powers, or generalizations of these families of numbers.

1 Introduction

Let us consider the prime factorization of a positive integral number n, namely

$$n = q_1^{s_1}\cdots q_r^{s_r}, \tag{1}$$

where the q_i $(i = 1,\dots,r)$ are its distinct prime factors and the s_i $(i = 1,\dots,r)$ are their respective multiplicities. Let $\Omega(n)$ be the total number of prime factors in the factorization of n, that is, $\Omega(n) = s_1 + \cdots + s_r$. The following formula is well-known (see, for example, [2, Thm. 430, p. 355] and [1, Sect. 1.4.4]):

$$\sum_{n\leq x}\Omega(n) = x\log\log x + B_2 x + O\left(\frac{x}{\log x}\right) \qquad (x \geq 2), \tag{2}$$

where

$$B_2 = B_1 + \sum_p \frac{1}{p(p-1)},$$

B_1 is the Mertens constant given by

$$B_1 = \gamma + \sum_p \left(\log\left(1 - \frac{1}{p}\right) + \frac{1}{p} \right),$$

and

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right) \approx 0.57721566\ldots$$

is the Euler–Mascheroni constant.

A positive integer n is *square-free* if and only if its prime factorization has no factors with an exponent larger than one, that is, $n = q_1 \cdots q_r$, where the q_i $(i = 1, \ldots, r)$ are distinct. Let $\mathcal{S}_2$ denote the set of square-free numbers, and let $Q_2(x)$ denote the cardinality of the set of square-free numbers not exceeding x. It is well-known [2, Thm. 333, p. 269] that these numbers have density $\frac{1}{\zeta(2)} = \frac{6}{\pi^2}$, where $\zeta(s)$ denotes the Riemann zeta function, that is,

$$Q_2(x) = \frac{x}{\zeta(2)} + O\left(x^{1/2}\right) = \frac{6}{\pi^2}x + O\left(x^{1/2}\right). \tag{3}$$

More generally, given a positive integer n with prime factorization $n = q_1^{s_1} \cdots q_r^{s_r}$ as in (1), we shall say that n is *h-free* if $s_i \le h - 1$ $(i = 1, \ldots, r)$. In particular, if $h = 2$ we obtain the square-free numbers. Let $\mathcal{S}_h$ denote the set of h-free numbers, and let $Q_h(x)$ be the cardinality of the set of h-free numbers not exceeding x. It is well-known that these numbers have density $\frac{1}{\zeta(h)}$, that is,

$$Q_h(x) = \frac{x}{\zeta(h)} + O\left(x^{1/h}\right), \qquad (h \ge 2). \tag{4}$$

In this note we show a proof for the following statement.

Theorem 1. *Let $h \ge 2$ be an arbitrary fixed integer, and let $\mathcal{S}_h$ be the set of h-free numbers. The following asymptotic formula holds*

$$\sum_{\substack{n \le x \\ n \in \mathcal{S}_h}} \Omega(n) = \frac{1}{\zeta(h)} x \log\log x + O(x).$$

Let $h \ge 2$ be an arbitrary fixed integer. A positive integer is said to be *h-full* if all the factors in its prime factorization have exponent greater than or equal to h. That is, the number $n = q_1^{s_1} \cdots q_r^{s_r}$ is h-full if $s_i \ge h$ $(i = 1, \ldots, r)$. If $h = 2$ these numbers are called *square-full* or *powerful*. Let $\mathcal{N}_h$ denote the set of h-full numbers and let $A_h(x)$ denote the cardinality of the set of h-full numbers not exceeding x. The number $A_h(x)$ has been estimated by Ivić and Shiu [3] with the following asymptotic formula:

$$A_h(x) = \gamma_{0,h}x^{1/h} + \gamma_{1,h}x^{1/(h+1)} + \cdots + \gamma_{h-1,h}x^{1/(2h-1)} + \Delta_h(x),$$

where $\gamma_{0,h}, \gamma_{1,h} \ldots, \gamma_{h-1,h}$ are certain computable constants and $\Delta_h(x) \ll x^\rho$ with $\rho \leq \frac{263}{2052}$. (In fact, ρ can be much smaller, depending on h.)

The constant for the main term in the above asymptotics can be precisely described by the following [5]:

$$\gamma_{0,h} = \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{\sigma(u(n))n^{1/h}} = \prod_p\left(1 + \frac{p - p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right),$$

where $\sigma(n)$ denotes the sum of divisors of n and $u(n)$ is the greatest square-free integer that divides n. Thus, if $n = q_1^{s_1}\cdots q_r^{s_r}$, then $\sigma(u(n)) = (q_1+1)\cdots(q_r+1)$.

In this note we prove the following result.

Theorem 2. *Let $h \geq 2$ be an arbitrary fixed integer, and let $\mathcal{N}_h$ be the set of h-full numbers. The following asymptotic formula holds*

$$\begin{aligned}\sum_{\substack{n\leq x\\ n\in\mathcal{N}_h}}\Omega(n) &= h\gamma_{0,h}x^{1/h}\log\log x\\ &+ \left(h(B_2-\log h) + \sum_p \frac{(h+1)p^{1+1/h} - hp - 2hp^{2/h} + (2h-1)p^{1/h}}{(p-1)\left(p^{1/h}-1\right)\left(p^{1+1/h}+p^{1/h}-p\right)}\right)\gamma_{0,h}x^{1/h}\\ &+ O_h\left(\frac{x^{1/h}}{\sqrt{\log x}}\right).\end{aligned}$$

Setting $h = 2$ in the previous result, we immediately obtain the following corollary.

Corollary 3. *The following asymptotic formula holds*

$$\begin{aligned}\sum_{\substack{n\leq x\\ n\in\mathcal{N}_2}}\Omega(n) &= 2\frac{\zeta(3/2)}{\zeta(3)}x^{1/2}\log\log x\\ &+ \left(2(B_2-\log 2) + \sum_p\frac{3}{p^{3/2}+1}\right)\frac{\zeta(3/2)}{\zeta(3)}x^{1/2} + O\left(\frac{x^{1/2}}{\sqrt{\log x}}\right).\end{aligned}$$

A *perfect power* is a number of the form n^m, where m and n are positive integers such that $m \geq 2$. Therefore, all perfect powers are square-full numbers. Let $\mathcal{P}$ denote the set of perfect powers and let $N(x)$ be the cardinality of the set of perfect powers not exceeding x. The following formula is well-known [4, Thm. 5]

$$N(x) = x^{1/2} + f(x)x^{1/3},$$

where $\lim_{x\longrightarrow\infty} f(x) = 1$.

We prove the following result.

Theorem 4. *Let $\mathcal{P}$ be the set of perfect powers. The following asymptotic formula holds*

$$\sum_{\substack{n\le x\\ n\in\mathcal{P}}}\Omega(n)=2x^{1/2}\log\log x+2(B_2-\log 2)x^{1/2}+O\left(\frac{x^{1/2}}{\log x}\right),$$

and therefore,

$$\sum_{\substack{n\le x\\ n\in\mathcal{P}}}\Omega(n)\sim 2N(x)\log\log x.$$

The proofs of Theorems 1, 2, and 4 are detailed in Sections 2, 3, and 4 respectively.

2 Proof of Theorem 1

We start by considering the following Euler product (converging at $\mathrm{Re}(s)>1$), which gives the generating Dirichlet series for the h-free numbers:

$$\frac{\zeta(s)}{\zeta(hs)}=\prod_p\left(\frac{1-\frac{1}{p^{hs}}}{1-\frac{1}{p^s}}\right)=\prod_p\left(1+\frac{1}{p^s}+\cdots+\frac{1}{p^{(h-1)s}}\right)=\sum_{n\in\mathcal{S}_h}\frac{1}{n^s}.$$

We will incorporate the coefficient $\Omega(n)$ by introducing it as the exponent of an extra variable, which we will later differentiate. This preserves the additive structure of $\Omega(n)$.

$$\prod_p\left(1+\frac{z}{p^s}+\cdots+\frac{z^{h-1}}{p^{(h-1)s}}\right)=\sum_{n\in\mathcal{S}_h}\frac{z^{\Omega(n)}}{n^s}. \tag{5}$$

We follow the Selberg–Delange method (see, for example, [6, Sect. 7.4] and [1, Sect. 1.4.4]). We will work with the Euler product

$$\begin{aligned}F(s,z):&=\prod_p\left(1-\frac{1}{p^s}\right)^z\left(1+\frac{z}{p^s}+\cdots+\frac{z^{h-1}}{p^{(h-1)s}}\right)\\&=\prod_p\left(1-\frac{1}{p^s}\right)^z\left(\frac{1-\frac{z^h}{p^{hs}}}{1-\frac{z}{p^s}}\right)=:\prod_p H(p,s,z).\end{aligned} \tag{6}$$

We claim that the product defining $F(s,z)$ converges absolutely and uniformly over compact subsets of $\mathrm{Re}(s)>1/2$, $|z|<\sqrt{2}$. Indeed, taking the logarithm of a single factor of (6), we have

$$\begin{aligned}\log H(p,s,z) &= -\log\left(1-\frac{z}{p^s}\right)+\log\left(1-\frac{z^h}{p^{hs}}\right)+z\log\left(1-\frac{1}{p^s}\right)\\ &= \sum_{k=1}^{\infty}\frac{z^k}{kp^{sk}}-\sum_{k=1}^{\infty}\frac{z^{hk}}{kp^{shk}}-z\sum_{k=1}^{\infty}\frac{1}{kp^{sk}}\\ &= \sum_{k=2}^{\infty}(z^k-z)\frac{1}{kp^{sk}}-\sum_{\substack{k=1\\ h\mid k}}^{\infty}\frac{hz^k}{kp^{sk}}. \end{aligned} \tag{7}$$

The sums in (7) converge for $\mathrm{Re}(s)>1/2$ and $|z|<\sqrt{2}$. More precisely, we have

$$\log H(p,s,z) \ll_h \sum_{k=2}^{\infty}\frac{\max(1,|z|^k)}{p^{k\,\mathrm{Re}(s)}} = \max\left(\frac{\frac{1}{p^{2\,\mathrm{Re}(s)}}}{1-\frac{1}{p^{\mathrm{Re}(s)}}},\frac{\frac{|z|^2}{p^{2\,\mathrm{Re}(s)}}}{1-\frac{|z|}{p^{\mathrm{Re}(s)}}}\right) \ll_{h,z} \frac{1}{p^{2\,\mathrm{Re}(s)}},$$

where we have used that the geometric series converges because $|z|/p^{\mathrm{Re}(s)}<1$. Summing over all the primes p, the above is bounded by $\zeta(2\,\mathrm{Re}(s))$ and is therefore finite for $\mathrm{Re}(s)>1/2$, $|z|<\sqrt{2}$. Thus we deduce that $\log F(s,z)$ converges absolutely and uniformly over compact subsets of $\mathrm{Re}(s)>1/2$, $|z|<\sqrt{2}$, and the same is true for $F(s,z)$.

Back to equation (5), we have

$$\sum_{n\in\mathcal{S}_h}\frac{z^{\Omega(n)}}{n^s}=\prod_p\left(1+\frac{z}{p^s}+\cdots+\frac{z^{h-1}}{p^{(h-1)s}}\right)=\zeta(s)^zF(s,z),$$

and this converges for $\mathrm{Re}(s)>1$, $|z|<\sqrt{2}$.

Now we apply Theorem 7.18 from [6] and obtain that

$$\sum_{\substack{n\le x\\ n\in\mathcal{S}_h}} z^{\Omega(n)}=\frac{F(1,z)}{\Gamma(z)}x\log^{z-1}x+O\left(x(\log x)^{\mathrm{Re}(z)-2}\right). \tag{8}$$

Differentiating both sides of (8) with respect to z, for z close to 1, we obtain

$$\sum_{\substack{n\le x\\ n\in\mathcal{S}_h}}\Omega(n)z^{\Omega(n)-1}=\left(\frac{F(1,z)}{\Gamma(z)}\right)'x\log^{z-1}x+\frac{F(1,z)}{\Gamma(z)}x(\log\log x)(\log^{z-1}x)+O\left(x\right), \tag{9}$$

where we have bounded the error term by using the mean value theorem and the fact that $(\log\log x)(\log^{\mathrm{Re}(z)-2}x)\ll 1$ when z is near 1. Evaluating (9) at $z=1$ gives

$$\sum_{\substack{n\le x\\ n\in\mathcal{S}_h}}\Omega(n)=\frac{F(1,1)}{\Gamma(1)}x(\log\log x)+O\left(x\right),$$

where the first term in (9) has been absorbed into the $O(x)$ error upon evaluation. Since

$$F(1,1)=\prod_p\left(1-\frac{1}{p^h}\right)=\frac{1}{\zeta(h)},$$

and $\Gamma(1)=1$, we obtain the desired result.

3 Proof of Theorem 2

Let n be an h-full number. We proceed to separate the primes in the factorization of n according to the congruence modulo h of their exponent. In other words, we write

$$n=\underbrace{q_{1,0}^{ht_{1,0}}\cdots q_{\ell_0,0}^{ht_{\ell_0,0}}}_{\exp\equiv 0(\bmod h)}\underbrace{q_{1,1}^{ht_{1,1}+1}\cdots q_{\ell_1,1}^{ht_{\ell_1,1}+1}}_{\exp\equiv 1(\bmod h)}\cdots\underbrace{q_{1,h-1}^{ht_{1,h-1}+h-1}\cdots q_{\ell_{h-1},h-1}^{ht_{\ell_{h-1},h-1}+h-1}}_{\exp\equiv h-1(\bmod h)}.$$

By setting

$$\begin{aligned}m&=q_{1,0}^{t_{1,0}}\cdots q_{\ell_0,0}^{t_{\ell_0,0}}q_{1,1}^{t_{1,1}-1}\cdots q_{\ell_1,1}^{t_{\ell_1,1}-1}\cdots q_{1,h-1}^{t_{1,h-1}-1}\cdots q_{\ell_{h-1},h-1}^{t_{\ell_{h-1},h-1}-1},\\ r_j&=q_{1,j}\cdots q_{\ell_j,j},\qquad (j=1,\ldots,h-1),\end{aligned}$$

we can finally write

$$n=m^hr_1^{h+1}\cdots r_{h-1}^{2h-1},$$

and we remark that the r_j are square-free and coprime in pairs.

After the above observation, the sum under consideration becomes

$$\sum_{\substack{n\le x\\ n\in\mathcal{N}_h}}\Omega(n)=\sum_{m^hr_1^{h+1}\cdots r_{h-1}^{2h-1}\le x}\mu^2(r_1\cdots r_{h-1})\Omega(m^hr_1^{h+1}\cdots r_{h-1}^{2h-1}).\tag{10}$$

Since Ω is totally additive and the r_j are square-free,

$$\begin{aligned}\Omega(m^hr_1^{h+1}\cdots r_{h-1}^{2h-1})&=h\Omega(m)+(h+1)\Omega(r_1)+\cdots+(2h-1)\Omega(r_{h-1})\\ &=h\Omega(m)+(h+1)\omega(r_1)+\cdots+(2h-1)\omega(r_{h-1}),\end{aligned}$$

where $\omega(n)$ denotes the number of distinct prime factors in the factorization of n (if $n=q_1^{s_1}\cdots q_r^{s_r}$, then $\omega(n)=r$).

Using this in (10), we obtain

$$\begin{aligned}\sum_{\substack{n\le x\\ n\in\mathcal{N}_h}}\Omega(n) &= h\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x}\mu^2(r_1\cdots r_{h-1})\sum_{m^h\le x/r_1^{h+1}\cdots r_{h-1}^{2h-1}}\Omega(m)\\ &\quad+(h+1)\sum_{r_1^{h+1}\le x}\omega(r_1)\sum_{m^h r_2^{h+2}\cdots r_{h-1}^{2h-1}\le x/r_1^{h+1}}\mu^2(r_1\cdots r_{h-1})\\ &\quad\vdots\\ &\quad+(2h-1)\sum_{r_{h-1}^{2h-1}\le x}\omega(r_{h-1})\sum_{m^h r_1^{h+1}\cdots r_{h-2}^{2h-2}\le x/r_{h-1}^{2h-1}}\mu^2(r_1\cdots r_{h-1})\\ &= hT_0(x)+(h+1)T_1(x)+\cdots+(2h-1)T_{h-1}(x).\end{aligned}\tag{11}$$

Eventually we will see that $T_0(x)$ gives the main term, while the $T_j(x)$ give secondary terms.

In order to continue we need some auxiliary results.

Lemma 5. *We have the following estimate*

$$\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x} 1 \ll_h x^{1/(h+1)}.\tag{12}$$

Proof of Lemma 5. We expand the sum in (12) as a combination of iterated sums, and we evaluate the successive terms starting from the innermost sum.

$$\begin{aligned}&\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x} 1\\ &=\sum_{r_1\le x^{1/(h+1)}}\ \sum_{r_2\le \frac{x^{1/(h+2)}}{r_1^{(h+1)/(h+2)}}}\cdots\sum_{r_{h-2}\le\frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}}\ \sum_{r_{h-1}\le\frac{x^{1/(2h-1)}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-2}^{(2h-2)/(2h-1)}}} 1\\ &\ll_h\sum_{r_1\le x^{1/(h+1)}}\ \sum_{r_2\le \frac{x^{1/(h+2)}}{r_1^{(h+1)/(h+2)}}}\cdots\sum_{r_{h-2}\le\frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}}\frac{x^{1/(2h-1)}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-2}^{(2h-2)/(2h-1)}}.\end{aligned}$$

In order to evaluate the innermost sum above, we approximate it with an integral

$$\begin{aligned}
&\sum_{r_{h-2}\le \frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}} \frac{x^{1/(2h-1)}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-2}^{(2h-2)/(2h-1)}}\\
&\ll_h \int_0^{\frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}} \frac{x^{1/(2h-1)}t^{-\frac{2h-2}{2h-1}}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-3}^{(2h-3)/(2h-1)}}dt\\
&\ll_h \left.\frac{x^{1/(2h-1)}t^{\frac{1}{2h-1}}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-3}^{(2h-3)/(2h-1)}}\right|_{t=0}^{t=\frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}}\\
&\ll_h \frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}.
\end{aligned}$$

Continuing in this way, we obtain

$$\begin{aligned}
&\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x} 1\\
&\ll_h \sum_{r_1\le x^{1/(h+1)}} \sum_{r_2\le \frac{x^{1/(h+2)}}{r_1^{(h+1)/(h+2)}}} \cdots \sum_{r_{h-3}\le \frac{x^{1/(2h-3)}}{r_1^{(h+1)/(2h-3)}\cdots r_{h-4}^{(2h-4)/(2h-3)}}} \frac{x^{1/(2h-2)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}\\
&\vdots\\
&\ll_h x^{1/(h+1)},
\end{aligned}$$

as desired. □

Lemma 6. *We have the following estimate*

$$\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x} \frac{\mu^2(r_1\cdots r_{h-1})}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} = \prod_p \left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right) + O_h\left(x^{-1/(h(h+1))}\right).$$

Proof of Lemma 6. First notice that we can approximate the sum up to x with the full sum as follows

$$
\begin{aligned}
&\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x} \frac{\mu^2(r_1\cdots r_{h-1})}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\\
&= \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\\
&+ O\left(\sum_{r_1\ge x^{1/(h+1)}} \frac{1}{r_1^{(h+1)/h}} \sum_{r_2} \frac{1}{r_2^{(h+2)/h}} \cdots \sum_{r_{h-2}} \frac{1}{r_{h-2}^{(2h-2)/h}} \sum_{r_{h-1}} \frac{1}{r_{h-1}^{(2h-1)/h}}\right.\\
&+ \sum_{r_1\le x^{1/(h+1)}} \frac{1}{r_1^{(h+1)/h}} \sum_{r_2\ge \frac{x^{1/(h+2)}}{r_1^{(h+1)/(h+2)}}} \frac{1}{r_2^{(h+2)/h}} \cdots \sum_{r_{h-2}} \frac{1}{r_{h-2}^{(2h-2)/h}} \sum_{r_{h-1}} \frac{1}{r_{h-1}^{(2h-1)/h}}\\
&\vdots\\
&+ \sum_{r_1\le x^{1/(h+1)}} \frac{1}{r_1^{(h+1)/h}} \sum_{r_2\le \frac{x^{1/(h+2)}}{r_1^{(h+1)/(h+2)}}} \frac{1}{r_2^{(h+2)/h}} \cdots \sum_{r_{h-2}\le \frac{x^{1/(2h-1)}}{r_1^{(h+1)/(2h-2)}\cdots r_{h-3}^{(2h-3)/(2h-2)}}} \frac{1}{r_{h-2}^{(2h-2)/h}}\\
&\left.\times \sum_{r_{h-1}\ge \frac{x^{1/(2h-1)}}{r_1^{(h+1)/(2h-1)}\cdots r_{h-2}^{(2h-2)/(2h-1)}}} \frac{1}{r_{h-1}^{(2h-1)/h}}\right). \qquad (13)
\end{aligned}
$$

We proceed to express the main term in (13) as an Euler product. This gives

$$
\begin{aligned}
\sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} &= \prod_p \left(1+\frac{1}{p^{(h+1)/h}}+\frac{1}{p^{(h+2)/h}}+\cdots+\frac{1}{p^{(2h-1)/h}}\right)\\
&= \prod_p \left(1+\frac{1}{p}\frac{p^{-1}-p^{-1/h}}{p^{-1/h}-1}\right)\\
&= \prod_p \left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right). \qquad (14)
\end{aligned}
$$

Now we treat the error term in (13). The inner sums with no condition over r or with a less-or-equal-than condition can be bounded by a constant, since

$$
\sum_{r=1}^{\infty} \frac{1}{r^{(h+j)/h}} = \zeta\left(\frac{h+j}{h}\right). \qquad (15)
$$

Therefore, the error term from (13) is

$$O_h\left(\sum_{j=1}^{h-1}\sum_{r\ge x^{1/(h+j)}}\frac{1}{r^{(h+j)/h}}\right)=O_h\left(\sum_{j=1}^{h-1}\int_{x^{1/(h+j)}}^{\infty}\frac{dt}{t^{(h+j)/h}}\right)=O_h\left(x^{-1/(h(h+1))}\right).$$

Combining the above with (14) in (13) yields the desired result. □

We can now return to the proof of Theorem 2. First we proceed to calculate $T_0(x)$ in (11). Using that $\Omega(m)=\Omega(1)=1$ when $x/2<r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x$ and estimate (2) when $r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2$, we obtain

$$\begin{aligned}T_0(x)=&\sum_{x/2<r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x}\mu^2(r_1\cdots r_{h-1})\\&+\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2}\mu^2(r_1\cdots r_{h-1})\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\log\log\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)\right.\\&\left.+B_2\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)+O_h\left(\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2}\frac{\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}}{\log\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)}\right).\end{aligned}\tag{16}$$

Notice that the first term in the sum above is $\ll_h x^{1/(h+1)}$ by Lemma 5. To bound the big-O error term, we fix $0<\varepsilon<\frac{1}{h+1}$. Since the function $\frac{t}{\log t}$ is strictly increasing for $t>e$ we have, for $\varepsilon>\frac{1}{\log(y/r)}$ and $1<r$,

$$\frac{\frac{y}{r}}{\log\left(\frac{y}{r}\right)}=\frac{\varepsilon\frac{y^\varepsilon}{r^\varepsilon}}{\log\left(\frac{y^\varepsilon}{r^\varepsilon}\right)}\cdot\frac{y^{1-\varepsilon}}{r^{1-\varepsilon}}\le\varepsilon\frac{y^\varepsilon}{\log\left(y^\varepsilon\right)}\cdot\frac{y^{1-\varepsilon}}{r^{1-\varepsilon}}=\frac{y}{\log y}\cdot\frac{1}{r^{1-\varepsilon}}.\tag{17}$$

On the one hand, we limit the sum so that $\frac{1}{\log(y/r)}\le\frac{1}{h+2}$ and therefore we can take $\frac{1}{h+2}<\varepsilon<\frac{1}{h+1}$. Taking $y=x^{1/h}$ and $r=r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}$, we have

$$\begin{aligned}&\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/e^{h(h+2)}}\frac{\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}}{\log\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)}\\&\le\frac{x^{1/h}}{\log\left(x^{1/h}\right)}\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/e^{h(h+2)}}\frac{1}{r_1^{(1-\varepsilon)(h+1)/h}\cdots r_{h-1}^{(1-\varepsilon)(2h-1)/h}}\\&\ll_h\frac{x^{1/h}}{\log x},\end{aligned}\tag{18}$$

where the final sum was bounded using a similar idea to the one we used in (15), since $\frac{(1-\varepsilon)(h+j)}{h} > 1$.

On the other hand,

$$\sum_{x/e^{h(h+2)}<r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2} \frac{\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}}{\log\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)} \ll_h \sum_{x/e^{h(h+2)}<r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2} \frac{he^{h+2}}{\log 2} \ll_h x^{1/(h+1)} \tag{19}$$

by Lemma 5.

Now we study the main term in (16). By writing

$$\log\log\left(\frac{y}{r}\right) = \log\log y + \log\left(1-\frac{\log r}{\log y}\right),$$

and applying this to $\log\log\left(\frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\right)$ we are led to consider

$$M := \sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2} \frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\left|\log\left(1-\frac{\log\left(r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}\right)}{\log\left(x^{1/h}\right)}\right)\right|.$$

We use that for $0\le t<1$, we have

$$-\log(1-t)\le \frac{t}{\sqrt{1-t}},$$

which can be proven by considering the derivative of $\log(1-t)+\frac{t}{\sqrt{1-t}}$. This gives

$$M \ll \sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2} \frac{x^{1/h}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}\cdot\frac{\log\left(r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}\right)}{\log\left(x^{1/h}\right)\sqrt{1-\frac{\log\left(r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}\right)}{\log\left(x^{1/h}\right)}}}$$

$$\ll \frac{x^{1/h}}{\log x}\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1}\le x/2} \frac{\log\left(r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}\right)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}\sqrt{1-\frac{\log\left(\frac{x^{1/h}}{2^{1/h}}\right)}{\log\left(x^{1/h}\right)}}}.$$

Taking advantage of the fact that

$$\sqrt{1-\frac{\log\left(\frac{x^{1/h}}{2^{1/h}}\right)}{\log\left(x^{1/h}\right)}} = \sqrt{\frac{\log 2}{\log x}},$$

and applying considerations similar to (15), we can simplify the bound to

$$M \ll_h \frac{x^{1/h}}{\sqrt{\log x}}. \tag{20}$$

By applying Lemma 6 in (16) and by incorporating the error terms (18) and (20) we obtain

$$\begin{aligned} T_0(x) &= \prod_p \left(1 + \frac{p - p^{1/h}}{p^2\left(p^{1/h} - 1\right)}\right) x^{1/h} \log\log x \\ &\quad + (B_2 - \log h) \prod_p \left(1 + \frac{p - p^{1/h}}{p^2\left(p^{1/h} - 1\right)}\right) x^{1/h} + O_h\left(\frac{x^{1/h}}{\sqrt{\log x}}\right). \end{aligned} \tag{21}$$

We proceed similarly to calculate the $T_j(x)$ in (11).

$$\begin{aligned} T_j(x) &= \sum_{r_j^{h+j} \le x} \omega(r_j) \sum_{r_1^{h+1} \cdots \widehat{r_j^{h+j}} \cdots r_{h-1}^{2h-1} \le x/r_j^{h+j}} \mu^2(r_1 \cdots r_{h-1}) \sum_{m^h \le \frac{x}{r_1^{h+1} \cdots r_{h-1}^{2h-1}}} 1 \\ &= \sum_{r_j^{h+j} \le x} \omega(r_j) \sum_{r_1^{h+1} \cdots \widehat{r_j^{h+j}} \cdots r_{h-1}^{2h-1} \le x/r_j^{h+j}} \mu^2(r_1 \cdots r_{h-1}) \left\lfloor \frac{x^{1/h}}{r_1^{(h+1)/h} \cdots r_{h-1}^{(2h-1)/h}} \right\rfloor, \end{aligned}$$

where the hat symbol indicates that the corresponding factor has been excluded.

We write

$$\left\lfloor \frac{x^{1/h}}{r_1^{(h+1)/h} \cdots r_{h-1}^{(2h-1)/h}} \right\rfloor = \frac{x^{1/h}}{r_1^{(h+1)/h} \cdots r_{h-1}^{(2h-1)/h}} + O(1)$$

and use the trivial bound

$$\omega(n) \le \frac{\log n}{\log 2} \qquad (n \ge 2) \tag{22}$$

to obtain

$$\begin{aligned} T_j(x) &= x^{1/h} \sum_{r_j^{h+j} \le x} \omega(r_j) \sum_{r_1^{h+1} \cdots \widehat{r_j^{h+j}} \cdots r_{h-1}^{2h-1} \le x/r_j^{h+j}} \frac{\mu^2(r_1 \cdots r_{h-1})}{r_1^{(h+1)/h} \cdots r_{h-1}^{(2h-1)/h}} \\ &\quad + O\left(\log x \sum_{r_1^{h+1} r_2^{h+2} \cdots r_{h-1}^{2h-1} \le x} 1\right). \end{aligned} \tag{23}$$

For the main term in (23), we complete the sum and proceed similarly as in the proof of Lemma 6. The difference here is that we have to work with sums of the form

$$\sum_{r=1}^{\infty} \frac{\omega(r)}{r^{(h+j)/h}} \text{ and } \sum_{r \ge x^{1/(h+j)}} \frac{\omega(r)}{r^{(h+j)/h}}.$$

We can bound these sums using the fact that given any arbitrarily small $\varepsilon > 0$, it follows from inequality (22) that $\omega(r) \ll r^\varepsilon$ provided r is sufficiently large. This results in

$$\sum_{r_1^{h+1}\cdots r_{h-1}^{2h-1} \le x} \frac{\mu^2(r_1\cdots r_{h-1})\omega(r_j)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} = \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})\omega(r_j)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} + O_h\left(x^{-1/(h(h+1))+\varepsilon}\right).$$

For the error term in (23) we can use Lemma 5. Combining the above with (12) in (23), we obtain

$$T_j(x) = x^{1/h} \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})\omega(r_j)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} + O_h\left(x^{1/(h+1)}\log x\right). \tag{24}$$

We aim at simplifying the main term in the above expression as much as possible. In order to do this, consider the following expression for a generating series with its corresponding Euler product:

$$\begin{aligned} G_j(z) &:= \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})z^{\omega(r_j)}}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} \\ &= \prod_p \left(1 + \frac{1}{p^{(h+1)/h}} + \cdots + \frac{z}{p^{(h+j)/h}} + \cdots + \frac{1}{p^{(2h-1)/h}}\right). \end{aligned}$$

The function $G_j(z)$ is holomorphic, since the product converges independently of z. For us, it suffices that $G_j(z)$ be holomorphic in a neighborhood of $z = 1$. Notice that

$$G_j'(1) = \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})\omega(r_j)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}}.$$

In order to compute $G_j'(1)$ from the Euler product, we consider the logarithmic derivative of $G_j(z)$ and obtain

$$\begin{aligned} \frac{G_j'(1)}{G_j(1)} &= \sum_p \frac{p^{-(h+j)/h}}{1 + \frac{1}{p^{(h+1)/h}} + \cdots + \frac{1}{p^{(2h-1)/h}}} \\ &= \sum_p \frac{p^{1+(1-j)/h} - p^{1-j/h}}{(p-1)\left(p^{1+1/h} + p^{1/h} - p\right)}. \end{aligned}$$

Multiplying by $G_j(1)$ and using (14), we obtain

$$\begin{aligned} G_j'(1) &= \sum_{r_1,\dots,r_{h-1}} \frac{\mu^2(r_1\cdots r_{h-1})\omega(r_j)}{r_1^{(h+1)/h}\cdots r_{h-1}^{(2h-1)/h}} \\ &= \prod_p \left(1 + \frac{p - p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right) \sum_p \frac{p^{1+(1-j)/h} - p^{1-j/h}}{(p-1)\left(p^{1+1/h} + p^{1/h} - p\right)}. \end{aligned}$$

Replacing the above in the main term of (24), and replacing this result, as well as (21), in (11), we get

$$\begin{aligned}\sum_{\substack{n\le x\\ n\in\mathcal{N}_h}}\Omega(n) &= h\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}\log\log x\\ &+\left(h(B_2-\log h)+\sum_p\frac{\sum_{j=1}^{h-1}(h+j)\left(p^{1+(1-j)/h}-p^{1-j/h}\right)}{(p-1)\left(p^{1+1/h}+p^{1/h}-p\right)}\right)\\ &\times\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}+O_h\left(\frac{x^{1/h}}{\sqrt{\log x}}\right)\\ &= h\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}\log\log x\\ &+\left(h(B_2-\log h)+\sum_p\frac{p\left(p^{1/h}-1\right)\sum_{j=1}^{h-1}(h+j)p^{-j/h}}{(p-1)\left(p^{1+1/h}+p^{1/h}-p\right)}\right)\\ &\times\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}+O_h\left(\frac{x^{1/h}}{\sqrt{\log x}}\right).\end{aligned}$$

We use that

$$\sum_{j=1}^{h-1}(h+j)x^j=\frac{(2h-1)x^{h+1}-2hx^h-hx^2+(h+1)x}{(x-1)^2}$$

in order to obtain

$$\begin{aligned}\sum_{\substack{n\le x\\ n\in\mathcal{N}_h}}\Omega(n) &= h\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}\log\log x\\ &+\left(h(B_2-\log h)+\sum_p\frac{(h+1)p^{1+1/h}-hp-2hp^{2/h}+(2h-1)p^{1/h}}{(p-1)\left(p^{1/h}-1\right)\left(p^{1+1/h}+p^{1/h}-p\right)}\right)\\ &\times\prod_p\left(1+\frac{p-p^{1/h}}{p^2\left(p^{1/h}-1\right)}\right)x^{1/h}+O_h\left(\frac{x^{1/h}}{\sqrt{\log x}}\right).\end{aligned}$$

This concludes the proof of Theorem 2.

4 Proof of Theorem 4

We first split the sum into a term involving squares and another term involving higher powers, as we expect that the squares are the only contributors to the main term.

$$\begin{aligned}\sum_{\substack{n\le x\\ n\in\mathcal{P}}}\Omega(n) &= \sum_{n^2\le x}\Omega(n^2)+\sum_{\substack{n^k\le x\\ k\ge 3}}\Omega(n^k)\\ &= 2\sum_{n\le x^{1/2}}\Omega(n)+\sum_{\substack{n^k\le x\\ k\ge 3}}\Omega(n^k)\\ &= S_1(x)+S_2(x). \end{aligned}\tag{25}$$

We can estimate the contribution from the squares by using (2),

$$\begin{aligned} S_1(x) &= 2x^{1/2}\log\log x^{1/2}+2B_2x^{1/2}+O\left(\frac{x^{1/2}}{\log x}\right)\\ &= 2x^{1/2}\log\log x+2(B_2-\log 2)x^{1/2}+O\left(\frac{x^{1/2}}{\log x}\right). \end{aligned}\tag{26}$$

Using (22), we obtain the following bound for the contribution of the higher powers:

$$S_2(x)=\sum_{\substack{n^k\le x\\ k\ge 3}}\Omega(n^k)\le\sum_{\substack{n^k\le x\\ k\ge 3}}\frac{\log n^k}{\log 2}\le\frac{\log x}{\log 2}x^{1/3}.$$

Combining this estimate with equation (26) in (25) gives the desired result.

5 Acknowledgments

The authors would like to thank the anonymous referee who provided extensive arguments and constructive remarks on an earlier version of this article, as well as helpful suggestions for the final version. The first author is grateful to Universidad Nacional de Luján. The second author is partially supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 355412-2013) and the Fonds de recherche du Québec - Nature et technologies (Projets de recherche en équipe 256442 and 300951).

References

1. S. R. Finch, *Mathematical Constants II*, Encyclopedia of Mathematics and its Applications, vol. 169, Cambridge University Press, 2019.

2. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, sixth ed., Oxford University Press, 2008.

3. A. Ivić and P. Shiu, The distribution of powerful integers, *Illinois J. Math.* **26** (1982), 576–590.

4. R. Jakimczuk, On the distribution of perfect powers, *J. Integer Seq.* **14** (2011), Article 11.8.5, 9.

5. R. Jakimczuk, The kernel of powerful numbers, *Int. Math. Forum* **12** (2017), 721–730.

6. H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

2020 *Mathematics Subject Classification*: Primary 11N37; Secondary 11A25, 11B99.
Keywords: $\Omega(n)$, integer sequence, h-free number, h-full number, perfect power.

(Concerned with sequences A005117, A001694, A004709, and A036966.)
