A Counting Function for the Sequence of Perfect Powers

M. A. Nyblom
School of Mathematics and Geospatial Science,
RMIT University,
GPO Box 2476V, Melbourne,
Victorial 3001, Australia
e-mail: michael.nyblom@rmit.edu.au

1 Introduction

A natural number of the form m^n where m is a positive integer and $n \geq 2$ is called a perfect power. Unsolved problems concerning the set of perfect powers abound throughout much of number theory. The most famous of these is known as the Catalan conjecture, which states that the only perfect powers which differ by unity are the integers 8 and 9. It is of interest to note that this particular problem has only recently been solved using rather deep results from the theory of cyclotomic fields (see [4]). The set of perfect powers can naturally be arranged into an increasing sequence of distinct integers, in which those perfect powers expressible with different exponents are treated as a single element of the sequence. The first few terms of this sequence of perfect powers without duplication are

$$1, 4, 8, 9, 16, 25, 27, 32, 36, 49, 64, 81, 100, 125, 128 \ldots\ , \tag{1}$$

and is listed in the On-Line Encyclopedia of Integer Sequences under Sloane A001597. The sequence in (1) has many properties, one being that the infinite sum of its reciprocals is convergent (see [3])–a clear indication of the scarcity of the perfect powers amongst the set of natural numbers. This latter fact is naturally reflected in the well known result that the sequence of perfect powers has zero asymptotic density that is, if $N(x)$ denotes the number of elements of (1) less than a positive real x, then $\lim_{x\to\infty} N(x)/x = 0$. In view of this result, one may question what is the precise nature of the growth rate of the counting function $N(x)$, in particular can an asymptotic estimate for $N(x)$ be found. We shall establish such a distributional result for the sequence of perfect powers by proving that $N(x) \sim \sqrt{x}$ as $x \to \infty$. As will be seen, this asymptotic formula can be interpreted as stating that the perfect squares dominate the count of the sequence elements in (1) as $x \to \infty$. To contrast the main result, we shall in addition develop a closed-form expression for $N(x)$ using elementary sieve methods.

Citation:
Nyblom, M 2006, 'A counting function for the sequence of perfect powers', Australian Mathematical Society Gazette, vol. 33, no. 5, pp. 338-343.

As will be seen, this formula is some what reminiscent to Legendre's counting function for the number of primes in the interval $(\sqrt{x}, x]$. In what follows we denote the integer part of x by $\lfloor x \rfloor$.

2 An Asymptotic Formula

To help establish the main results of this paper we shall first need to formally introduce the following family of sets.

Definition 2.1 *Suppose $x \geq 1$ and $n \in \mathbb{N} \backslash \{1\}$, then let $A_n(x)$ denote the set of perfect powers having exponent n and which are less than or equal to x that is, $A_n(x) = \{k^n : k \in \mathbb{N}, k^n \leq x\}$*

We now establish the asymptotic formula for the counting function $N(x)$.

Theorem 2.1 *If $N(x)$ denotes the number of sequence elements of (1) that are less than or equal to x, then $N(x) \sim \sqrt{x}$ as $x \to \infty$.*

Proof: The first step of the argument will be to obtain upper and lower functional bounds for $N(x)$. Assuming without loss of generality that $x \geq 4$ observe $1 \in A_n(x)$ for each $n \in \mathbb{N} \backslash \{1\}$ but for n sufficiently large $A_n(x) \backslash \{1\} = \emptyset$. Defining the auxiliary function $M(x) = \max\{n \in \mathbb{N} \backslash \{1\} : A_n(x) \backslash \{1\} \neq \emptyset\}$ we clearly see $M(x) \geq 2$, as $A_2(x) \backslash \{1\} \neq \emptyset$, and that $N(x)$ is equal to the number of elements of the set $A = \bigcup_{n=2}^{M(x)} A_n(x)$. Furthermore from the inequality $2^{\lfloor \log_2 x \rfloor} \leq x < 2^{\lfloor \log_2 x \rfloor + 1}$, it is immediately deduced that $M(x) = \lfloor \log_2 x \rfloor$. Since for large x the family of sets $\{A_n(x)\}_{n=2}^{\lfloor \log_2 x \rfloor}$ are not mutually disjoint it follows that

$$N(x) = |A| \leq \sum_{n=2}^{\lfloor \log_2 x \rfloor} |A_n(x)| \ , \tag{2}$$

and since $A_2(x) \subseteq A$, one also has

$$|A_2(x)| \leq |A| = N(x) \ . \tag{3}$$

Now as $A_n(x) \neq \emptyset$ there must exist a largest integer $m \geq 1$ such that $m^n \leq x < (m+1)^n$. By taking the n-th root through the previous inequality we deduce $m \leq \sqrt[n]{x} < m+1$, that is $m = \lfloor \sqrt[n]{x} \rfloor$ and so $A_n(x)$ must contain $\lfloor \sqrt[n]{x} \rfloor$ elements. Consequently (2) and (3) together yields that

$$\lfloor \sqrt{x} \rfloor \leq N(x) \leq \sum_{n=2}^{\lfloor \log_2 x \rfloor} \lfloor \sqrt[n]{x} \rfloor \ . \tag{4}$$

Using the upper and lower bounds in (4) we can establish required the asymptotic estimate for $N(x)$ as follows. Dividing (4) by $\sqrt{x}$ observe for large x the following train of inequalities

$$\frac{\lfloor\sqrt{x}\rfloor}{\sqrt{x}} \leq \frac{N(x)}{\sqrt{x}} \leq \frac{\lfloor\sqrt{x}\rfloor}{\sqrt{x}} + \sum_{n=3}^{\lfloor\log_2 x\rfloor} \frac{\lfloor\sqrt[n]{x}\rfloor}{\sqrt{x}} \leq \frac{\lfloor\sqrt{x}\rfloor}{\sqrt{x}} + \sum_{n=3}^{\lfloor\log_2 x\rfloor} \frac{\sqrt[n]{x}}{\sqrt{x}}$$
$$\leq \frac{\lfloor\sqrt{x}\rfloor}{\sqrt{x}} + \sum_{n=3}^{\lfloor\log_2 x\rfloor} \frac{\sqrt[3]{x}}{\sqrt{x}}$$
$$= \frac{\lfloor\sqrt{x}\rfloor}{\sqrt{x}} + \frac{(\lfloor\log_2 x\rfloor - 2)}{\sqrt[6]{x}} . \qquad (5)$$

Via an application of L'Hopitals rule, it is easily seen that

$$0 \leq \frac{\lfloor\log_2 x\rfloor - 2}{\sqrt[6]{x}} < \frac{\log_2 x}{\sqrt[6]{x}} \to 0$$

as $x \to \infty$, moreover by recalling $\lim_{x\to\infty}\lfloor x\rfloor/x = 1$, we finally deduce from (5) that $N(x)/\sqrt{x} \to 1$ as $x \to \infty$.

■

Remark: 2.1 *Since the number of perfect squares less than or equal to x is given by $\lfloor\sqrt{x}\rfloor$ and as $\lfloor\sqrt{x}\rfloor \sim \sqrt{x}$ we can interpret Theorem 2.1 as stating that the perfect squares dominate the count of the sequence elements of (1) as $x \to \infty$.*

3 An Exact Formula

One of the earliest known sieve methods was a simple effective procedure for finding all prime numbers up to a certain bound x. This procedure which involves the systematic deletion of all multiples of primes less than or equal to $\sqrt{x}$ was captured succinctly by Legendre using a theoretical analog of the sifting process, known today as the Inclusion-Exclusion Principal, to study the prime counting function $\pi(x) = |\{p \leq x : p \text{ a prime }\}|$. His method led to an exact formula for the number of primes in the interval $(\sqrt{x}, x]$ in particular, if $\mu(\cdot)$ denotes the Möbius function then

$$\pi(x) - \pi(\sqrt{x}) = -1 + \sum_{d|P_x} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor , \qquad (6)$$

where the sum is taken over all divisors of $P_x = \prod_{p\leq\sqrt{x}} p$ (see [2, pg.15]). In this section we shall employ the same elementary sieve method of Legendre to establish an exact formula for the counting function $N(x)$ which is similar in form to (10). We begin with a technical lemma for the sets of Definition 2.1.

Lemma 3.1 *For any set of m positive integers $\{n_1, \ldots, n_m\}$ all greater than unity*

$$\bigcap_{i=1}^{m} A_{n_i}(x) = A_{[n_1,\ldots,n_m]}(x) \ , \tag{7}$$

where $[n_1, \ldots, n_m]$ denotes the least common multiple of the m integers $n_1, \ldots, n_m$.

Proof: We begin by demonstrating that $A_n(x) \cap A_m(x) = A_{[n,m]}(x)$ for any $n, m \in \mathbb{N}\backslash\{1\}$, which is the base step of our inductive argument. Now since $n|[n,m]$ and $m|[n,m]$ any number of the form $k^{[n,m]}$ where $k \in \mathbb{N}$ can be rewritten as a perfect power having an exponent n and m, thus $A_{[n,m]}(x) \subseteq A_n(x) \cap A_m(x)$. Let $s \in A_n(x) \cap A_m(x)$ with $s \neq 1$, then $s = k_1^n = k_2^m$ for some $k_1, k_2 \in \mathbb{N}\backslash\{1\}$. We have to produce a $k \in \mathbb{N}\backslash\{1\}$ such that $s = k^{[n,m]}$. As $k_1^n = k_2^m$ both k_1 and k_2 must have the same prime divisors. Writing $k_1 = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ and $k_2 = p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}$ we deduce from the equality $k_1^n = k_2^m$ that $n\alpha_i = m\beta_i$ for each $i = 1, 2, \ldots, r$. Consequently $n|n\alpha_i$ and $m|n\alpha_i$ and so $n\alpha_i = [n,m]\gamma_i$ for some $\gamma_i \in \mathbb{N}$. Thus $s = k^{[n,m]}$ where $k = p_1^{\gamma_1} p_2^{\gamma_2} \cdots p_r^{\gamma_r}$ which establishes that $A_n(x) \cap A_m(x) \subseteq A_{[n,m]}(x)$.
Now suppose for $m > 1$ the set identity in (11) holds for an arbitrary set of m positive integers $\{n_1, \cdots, n_m\}$ all greater than unity. Then as $[[n_1, n_2, \ldots, n_m], n_{m+1}] = [n_1, \ldots, n_{m+1}]$ observe from the inductive assumption and the base step that

$$\begin{aligned}
\bigcap_{i=1}^{m+1} A_{n_i}(x) = (\bigcap_{i=1}^{m} A_{n_i}(x)) \cap A_{n_{m+1}}(x) &= A_{[n_1,\ldots,n_m]}(x) \cap A_{n_{m+1}}(x) \\
&= A_{[[n_1,\ldots,n_m],n_{m+1}]}(x) \\
&= A_{[n_1,\ldots,n_{m+1}]}(x) \ .
\end{aligned}$$

Hence (11) holds for $m+1$ arbitrary positive integers greater than unity and so the result is established by the principal of mathematical induction.

■

Theorem 3.1 *If $x \geq 4$ then the counting function for the sequence in (1) is given by the explicit expression*

$$N(x) = \lfloor x \rfloor - \sum_{d|P_x} \mu(d) \lfloor x^{\frac{1}{d}} \rfloor \ , \tag{8}$$

where the sum is taken over all divisors of $P_x = \prod_{p \leq \lfloor \log_2 x \rfloor} p$.

Proof: We begin by establishing a slight reformulation for the set A of Theorem 2.1. Recalling that $A = \cup_{n=2}^{\lfloor \log_2 x \rfloor} A_n(x)$, we claim if $p_1, \ldots, p_m$ are the first m primes less than or equal to $\lfloor \log_2 x \rfloor$, then in fact $A = B$ where

$$B = \bigcup_{r=1}^{m} A_{p_r}(x) \ .$$

The inclusion $B \subseteq A$ follows automatically by definition as each set $A_{p_k}(x)$ is included in the union of sets which form A. To establish the reverse inclusion $A \subseteq B$, first observe that as $p_1, \ldots, p_m$ represent the complete list of primes less than or equal to $\lfloor \log_2 x \rfloor$, every integer $n \in \{2, 3, \ldots, \lfloor \log_2 x \rfloor\}$ must be divisible by at least one of these primes since otherwise, by the fundamental theorem of arithmetic, n would be divisible by a prime $p' > \lfloor \log_2 x \rfloor$ and so $n > \lfloor \log_2 x \rfloor$, a contradiction. Consequently if given any $s \in A_n(x)$, then $s = k^n$ and one may write $n = p_r \gamma$ for some $r \in \{1, 2, \ldots, m\}$ and $\gamma \in \mathbb{N}$. Thus $s = (k^\gamma)^{p_r} \in A_{p_r}(x)$ and so every element of A is contained in the set B.

Now $N(x) = |A| = |B|$ and since for x large the family of sets $\{A_{p_i}(x)\}_{i=1}^m$ are not mutually disjoint we deduce from an application of the Inclusion-Exclusion Principal applied to the set B that

$$N(x) = \sum_{k=1}^{m} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le m} |A_{p_{i_1}}(x) \cap \cdots \cap A_{p_{i_k}}(x)| \ , \tag{9}$$

where the expression $1 \le i_1 < \cdots < i_k \le m$ indicates that the sum is taken over all ordered k-element subsets $\{i_1, \ldots, i_k\}$ of the set $\{1, 2, \ldots, m\}$. As the least common multiple of the k prime numbers $p_{i_1}, \ldots, p_{i_k}$ is clearly the product $d = p_{i_1} p_{i_2} \cdots p_{i_k}$, observe from Lemma 3.1 that $|A_{p_{i_1}}(x) \cap \cdots \cap A_{p_{i_k}}(x)| = |A_d(x)| = \lfloor x^{\frac{1}{d}} \rfloor$, noting here we have again used the fact that the number of elements in the set $A_n(x)$ is $\lfloor \sqrt[n]{x} \rfloor$. Defining $P_x = \prod_{p \le \lfloor \log_2 x \rfloor} p$ we see that for each $k \in \{1, 2, \ldots, m\}$ the inner summation in (9) consists of adding $\binom{m}{k}$ terms of the form $\lfloor x^{\frac{1}{d}} \rfloor$, where $d = p_{i_1} p_{i_2} \cdots p_{i_k}$ is a divisor of P_x having k distinct prime factors. Consequently as $\mu(p_{i_1} p_{i_2} \cdots p_{i_k}) = (-1)^k$ the double summation in (9) must sum terms of the form $-\mu(d) \lfloor x^{\frac{1}{d}} \rfloor$ over all divisors $d$where d of P_x excluding $d = 1$. Finally by recalling that $\mu(1) = 1$ we deduce that the right hand side of (9) reduces to the right hand side of (8).

■

Remark: 3.1 *An immediate consequence of Theorem 3.1 is that the number of non-perfect powers less than or equal to x is equal to $\sum_{d|P_x} \mu(d) \lfloor x^{\frac{1}{d}} \rfloor$.*

4 Numerical Example

We examine now how the explicit expression for $N(x)$ in (8) can be practically implemented to compute the number of perfect powers less than or equal to a given large positive real x. For notational convenience let the inner summation of (9) be denoted by

$$S_k(x) = \sum_{1 \le i_1 < \cdots < i_k \le m} \left\lfloor x^{(p_{i_1} \cdots p_{i_k})^{-1}} \right\rfloor \ .$$

Observe that in order to evaluate each $S_k(x)$, one must sum the terms $\lfloor x^{(p_{i_1}\cdots p_{i_k})^{-1}}\rfloor$ over those subscripts $i_1 < \cdots < i_k$ whose values are chosen from the ordered k-element subsets of $\{1, 2, \ldots, m\}$, consequently the number of summands is $\binom{m}{k}$. Thus on first acquaintance, it would appear that the calculation of $S_k(x)$ would involve having to determine for each $1 \le k \le m$, all $\binom{m}{k}$ combinations of prime numbers from the set $\{p_1, \ldots, p_m\}$. However, for sufficiently large x this may not be necessary since for certain values of k one can show that $S_k(x) = \binom{m}{k}$ as follows.

To begin consider for any $x > 2$ the arithmetic function $k(x) = \min\{k \in \mathbb{N} : p_1p_2\cdots p_k > x\}$, where again p_i denotes the i-th prime number. We wish to first show that if there are m primes less than or equal to $\lfloor \log_2 x\rfloor$, then $k(\lfloor \log_2 x\rfloor)$ will be at most $m-2$ when $m > 5$. Recalling for any $n \ge 2$, there exists a prime strictly between n and $2n$ (Bertrand's Postulate), observe as each $p_i \ge 2$, that

$$p_{m-5}p_{m-4}(p_{m-3}p_{m-2}) > p_{m-5}(p_{m-4}p_{m-1}) > p_{m-5}p_m > p_{m+1} > \lfloor \log_2 x\rfloor \ .$$

Thus when $m > 5$ we have $p_1\cdots p_{m-2} > \lfloor \log_2 x\rfloor$ and so $k(\lfloor \log_2 x\rfloor) \le m-2$. Now for $\lfloor \log_2 x\rfloor > p_5 = 11$ and $k \ge k(\lfloor \log_2 x\rfloor)$ we note that in the summation $S_k(x)$ all $\binom{m}{k}$ combinations of products $p_{i_1}\cdots p_{i_k} \ge p_1\cdots p_{k(\lfloor \log_2 x\rfloor)} > \lfloor \log_2 x\rfloor$. Consequently from the inequality $2^{\lfloor \log_2 x\rfloor} \le x < 2^{\lfloor \log_2 x\rfloor+1}$ it is immediate that

$$1 < 2^{\lfloor \log_2 x\rfloor (p_{i_1}\cdots p_{i_k})^{-1}} \le x^{(p_{i_1}\cdots p_{i_k})^{-1}} < 2^{(\lfloor \log_2 x\rfloor+1)(p_{i_1}\cdots p_{i_k})^{-1}} \le 2 \ .$$

Thus $\lfloor x^{(p_{i_1}\cdots p_{i_k})^{-1}}\rfloor = 1$ and so the summation $S_k(x)$ must consist of adding $\binom{m}{k}$ terms all of which are identically 1, that is $S_k(x) = \binom{m}{k}$. Hence for $x > 2^{p_5} = 2^{11}$ the number of perfect powers less than or equal to x can be calculated by the alternate expression

$$N(x) = \sum_{k=1}^{k(\lfloor \log_2 x\rfloor)-1} (-1)^{k+1} S_k(x) + \sum_{k=k(\lfloor \log_2 x\rfloor)}^{m} (-1)^{k+1} \binom{m}{k} \ . \tag{10}$$

For $x > 2^{11}$ the value of the arithmetic function $k(\lfloor \log_2 x\rfloor)$ will in practice be much smaller than the number of primes less than or equal to $\lfloor \log_2 x\rfloor$, consequently in calculating $N(x)$, we shall only have to evaluate $S_k(x)$ for the few values of $1 \le k < k(\lfloor \log_2 x\rfloor)$. In what follows the reader may wish to consult the table of perfect powers less than or equal to 10^9 by Serhart Sevki Dincer in [1].

Example 4.1 *Consider $x = 2^{18} = 262144$. From the table of perfect powers one can by inspection deduce that $N(x) = 583$. To demonstrate the use of (8) we shall apply the alternate* *expression in (**??**) to verify the number of perfect powers less than or equal to* x *is 583. Now* $\lfloor \log_2 x \rfloor = 18$ *and so there are* $m = 7$ *primes, namely* $2, 3, 5, 7, 11, 13, 17$ *less than* $\lfloor \log_2 x \rfloor$*. As* $2 \cdot 3 \cdot 5 > 18 > 2 \cdot 3$ *we have that* $k(\lfloor \log_2 x \rfloor) = 3$ *and so from (10)*

$$N(x) = S_1(x) - S_2(x) + \sum_{k=3}^{7} (-1)^{k+1} \binom{7}{k} \ . \tag{11}$$

Using a calculator one finds in this instance that

$$\begin{aligned} S_1(x) &= \lfloor \sqrt{2^{18}} \rfloor + \lfloor \sqrt[3]{2^{18}} \rfloor + \lfloor \sqrt[5]{2^{18}} \rfloor + \lfloor \sqrt[7]{2^{18}} \rfloor + \lfloor \sqrt[11]{2^{18}} \rfloor + \lfloor \sqrt[13]{2^{18}} \rfloor + \lfloor \sqrt[17]{2^{18}} \rfloor \\ &= 512 + 64 + 12 + 5 + 3 + 2 + 2 = 600 \ . \end{aligned}$$

To evaluate $S_2(x)$ *first recall from definition*

$$S_2(x) = \sum_{1 \le i_1 < i_2 \le 7} \lfloor 2^{18(p_{i_1} p_{i_2})^{-1}} \rfloor \ .$$

Now if $p_{i_1} p_{i_2} > 18$ *then* $\lfloor 2^{18(p_{i_1} p_{i_2})^{-1}} \rfloor = 1$*. However, of the* $\binom{7}{2} = 21$ *combinations of products* $p_{i_1} p_{i_2}$ *with* $1 \le i_1 < i_2 \le 7$*, the only products less than* 18 *are* $2 \cdot 3$*,* $2 \cdot 5$*,* $2 \cdot 7$ *and* $3 \cdot 5$*. Thus the summation* $S_2(x)$ *will consist of adding* $21 - 4 = 17$ *terms all of which are identically* 1*, together with the sum of the terms* $\lfloor \sqrt[6]{2^{18}} \rfloor$*,* $\lfloor \sqrt[10]{2^{18}} \rfloor$*,* $\lfloor \sqrt[14]{2^{18}} \rfloor$ *and* $\lfloor \sqrt[15]{2^{18}} \rfloor$*, which are* 8*,* 3*,* 2 *and* 2 *respectively. Consequently* $S_2(x) = 17 + 8 + 3 + 2 + 2 = 32$ *and so finally adding in the alternating sum of binomial coefficients in (11) yields*

$$N(x) = 600 - 32 + 35 - 35 + 21 - 7 + 1 = 583 \ ,$$

as required.

References

[1] S. S. Dincer, http://www.ug.bcc.bilkent.edu.tr/~sevki/powers_up_to_1e9.html

[2] G. Tenenbaum and M. France, *The Prime Numbers and Their Distribution*, Student Mathematics Library Volume 6, AMS.

[3] H. W. Gould, Problem H-170, *Fibonacci Quarterly*, Vol. 8, pg. 268, 1970.

[4] P. Mihailescu, A class number free criterion for Catalan's conjecture, *J. Number Theory*, 99 (2003), 225-231.