

THE AVERAGE SMARANDACHE FUNCTION

Florian Luca

Mathematical Institute, Czech Academy of Sciences

Žitná 25, 115 67 Praha 1

Czech Republic

For every positive integer n let $S(n)$ be the minimal positive integer m such that $n \mid m!$. For any positive number $x \geq 1$ let

$$A(x) = \frac{1}{x} \sum_{n \leq x} S(n) \quad (1)$$

be the average value of S on the interval $[1, x]$. In [6], the authors show that

$$A(x) < c_1 x + c_2 \quad (2)$$

where c_1 can be made rather small provided that x is enough large (for example, one can take $c_1 = .215$ and $c_2 = 45.15$ provided that $x > 1470$). It is interesting to mention that by using the method outlined in [6], one gets smaller and smaller values of c_1 for which (2) holds provided that x is large, but at the cost of increasing c_2 ! In the same paper, the authors ask whether it can be shown that

$$A(x) < \frac{2x}{\log x} \quad (3)$$

and conjecture that, in fact, the stronger version

$$A(x) < \frac{x}{\log x} \quad (4)$$

might hold (the authors of [6] claim that (4) has been tested by Ibstedt in the range $x \leq 5 \cdot 10^6$ in [4]. Although I have read [4] carefully, I found no trace of the aforementioned computation!).

In this note, we show that $\frac{x}{\log x}$ is indeed the correct order of magnitude of $A(x)$.

For any positive real number x let $\pi(x)$ be the number of prime numbers less than or equal to x ,

$$B(x) = xA(x) = \sum_{1 \leq n \leq x} S(n), \quad (5)$$

$$E(x) = 2.5 \log \log(x) + 6.2 + \frac{1}{x}. \quad (6)$$

We have the following result:

Theorem.

$$.5(\pi(x) - \pi(\sqrt{x})) < A(x) < \pi(x) + E(x) \quad \text{for all } x \geq 3. \quad (7)$$

Inequalities (7), combined with the prime number theorem, assert that

$$.5 \leq \liminf_{x \rightarrow \infty} \frac{A(x)}{\frac{x}{\log x}} \leq \limsup_{x \rightarrow \infty} \frac{A(x)}{\frac{x}{\log x}} \leq 1,$$

which says that $\frac{x}{\log x}$ is indeed the right order of magnitude of $A(x)$. The natural conjecture is that, in fact,

$$A(x) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right). \quad (8)$$

Since

$$\frac{x}{\log x} \left(1 + \frac{1}{2 \log x}\right) < \pi(x) < \frac{x}{\log x} \left(1 + \frac{3}{2 \log x}\right) \quad \text{for } x \geq 59,$$

it follows, by our theorem, that the upper bound on $A(x)$ is indeed of the type (8). Unfortunately, we have not succeeded in finding a lower bound of the type (8) for $A(x)$.

The Proof

We begin with the following observation:

Lemma.

Suppose that $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ is the decomposition of n in prime factors (we assume that the p_i 's are distinct but not necessarily ordered). Then:

1.
$$S(n) \leq \max_{i=1}^k (\alpha_i p_i). \quad (9)$$

2. Assume that $\alpha_1 p_1 = \max_{i=1}^k (\alpha_i p_i)$. If $\alpha_1 \leq p_1$, then $S(n) = \alpha_1 p_1$.

3.
$$S(n) > \alpha_i (p_i - 1) \quad \text{for all } i = 1, \dots, k. \quad (10)$$

Proof.

For every prime number p and positive integer k let $e_p(k)$ be the exponent at which p appears in $k!$.

1. Let $m \geq \max_{i=1}^k (\alpha_i p_i)$. Then

$$e_p(m) = \sum_{s \geq 1} \left\lfloor \frac{m}{p^s} \right\rfloor \geq \left\lfloor \frac{m}{p} \right\rfloor \geq \alpha_i \quad \text{for } i = 1, \dots, k.$$

This obviously implies $n \mid m!$, hence $m \geq S(n)$.

2. Assume that $\alpha_1 \leq p_1$. In this case, $S(n) \geq \alpha_1 p_1$. By 1 above, it follows that in fact $S(n) = \alpha_1 p_1$.

3. Let $m = S(n)$. The asserted inequality follows from

$$\alpha_i \leq e_{p_i}(m) = \sum_{s \geq 1} \left\lfloor \frac{m}{p_i^s} \right\rfloor < m \sum_{s \geq 1} \frac{1}{p_i^s} = \frac{m}{p_i - 1}.$$

The Proof of the Theorem.

In what follows p denotes a prime. We assume $x > 1$. The idea behind the proof is to find good bounds on the expression

$$B(x) - B(\sqrt{x}) = \sum_{\sqrt{x} < n \leq x} S(n). \quad (11)$$

Consider the following three subsets of the interval $I = (\sqrt{x}, x]$:

$$\begin{aligned} C_1 &= \{n \in I \mid S(n) \text{ is not a prime}\}, \\ C_2 &= \{n \in I \mid S(n) = p \leq \sqrt{x}\}, \\ C_3 &= \{n \in I \mid S(n) = p > \sqrt{x}\}. \end{aligned}$$

Certainly, the three subsets above are, in general, not disjoint but their union covers I . Let

$$D_i(x) = \sum_{n \in C_i} S(n) \quad \text{for } i = 1, 2, 3.$$

Clearly,

$$\max(D_i(x) \mid i = 1, 2, 3) \leq B(x) - B(\sqrt{x}) \leq D_1(x) + D_2(x) + D_3(x). \quad (12)$$

We now bound each D_i separately.

The bound for D_1 .

Assume that $m \in C_1$. By the Lemma, it follows that $S(m) \leq \alpha p$ for some $p^\alpha \parallel m$ and $\alpha > 1$. First of all, notice that $S(m) \leq \alpha\sqrt{m}$. Indeed, this follows from the fact that

$$S(m) \leq \alpha p \leq \alpha p^{\alpha/2} \leq \alpha\sqrt{m} \quad \text{for } \alpha \geq 2.$$

In particular, from the above inequality it follows that $p \leq \sqrt{m} \leq \sqrt{x}$. Write now $m = p^\alpha k$. Since $m \leq x$, it follows that $k \leq x/p^\alpha$. These considerations show that

$$D_1(x) < \sum_{p \leq \sqrt{x}} \sum_{\alpha \geq 2}^{\infty} \alpha p \cdot \frac{x}{p^\alpha} = x \sum_{p \leq \sqrt{x}} \sum_{\alpha \geq 2}^{\infty} \frac{\alpha}{p^{\alpha-1}} = x \sum_{p \leq \sqrt{x}} \frac{2p-1}{(p-1)^2}. \quad (13)$$

In the above formula (13), we used the fact that

$$\sum_{\alpha \geq 2} \alpha z^{\alpha-1} = \frac{d}{dz} \left(\frac{1}{1-z} \right) - 1 = \left(\frac{1}{1-z} \right)^2 - 1 = \frac{2z - z^2}{(1-z)^2} \quad \text{for } |z| < 1$$

with $z = 1/p$. Since

$$\frac{2p-1}{(p-1)^2} \leq \frac{5}{4p} \quad \text{for } p \geq 3,$$

it follows that

$$D_1(x) < x \left(3 - \frac{5}{8} + \frac{5}{4} \sum_{p \leq \sqrt{x}} \frac{1}{p} \right) = x \left(2.375 + 1.25 \sum_{p \leq \sqrt{x}} \frac{1}{p} \right). \quad (14)$$

From a formula from [5], we know that

$$\sum_{p \leq y} \frac{1}{p} < \log \log y + 1.27 \quad \text{for all } y > 1.$$

Hence, inequality (14) implies

$$D_1(x) < x \left(2.375 + 1.25 \left(\log \log \sqrt{x} + 1.27 \right) \right) < x \left(3.1 + 1.25 \log \log x \right). \quad (15)$$

The bound for D_2

Assume that $S(m) = p$. Then $m = py$ where p does not divide y . Since $m > \sqrt{x}$, it follows that

$$\frac{\sqrt{x}}{p} < y \leq \frac{x}{p}$$

Since $p \leq \sqrt{x}$, it follows that at least one integer in the above interval is a multiple of p ; hence, cannot be an acceptable value for y . This shows that there are at most

$$\left\lfloor \frac{x - \sqrt{x}}{p} \right\rfloor \leq \frac{x - \sqrt{x}}{p}$$

possible values for y . Hence,

$$D_2(x) \leq \sum_{p \leq \sqrt{x}} p \cdot \left(\frac{x - \sqrt{x}}{p} \right) \leq (x - \sqrt{x}) \pi(\sqrt{x}). \quad (16)$$

Bounds for D_3

Assume $S(m) = p$ for some $p > \sqrt{x}$. Then, $m = py$ for some $y < x/p$. Hence,

$$D_3(x) = \sum_{\sqrt{x} < p \leq x} p \cdot \left\lfloor \frac{x}{p} \right\rfloor. \quad (17)$$

Notice that, unlike in the previous cases, (17) is in fact an equality. Since $z \geq [z] > .5z$ for all real numbers $z > 1$, it follows, from formula (17), that

$$.5x(\pi(x) - \pi(\sqrt{x})) < D_3(x) < x(\pi(x) - \pi(\sqrt{x})). \quad (18)$$

Denote now by

$$F(x) = 3.1 + 1.25 \log \log(x)$$

From inequalities (12), (15), (16) and (17), it follows that

$$\begin{aligned} .5x(\pi(x) - \pi(\sqrt{x})) < D_3(x) < B(x) - B(\sqrt{x}) < D_1(x) + D_2(x) + D_3(x) < \\ xF(x) + (x - \sqrt{x})\pi(\sqrt{x}) + x(\pi(x) - \pi(\sqrt{x})) = x\pi(x) - \sqrt{x}\pi(\sqrt{x}) + xF(x). \end{aligned} \quad (19)$$

The left inequality (7) is now obvious since

$$B(x) > B(\sqrt{x}) + .5x(\pi(x) - \pi(\sqrt{x})) \geq 1 + .5x(\pi(x) - \pi(\sqrt{x})).$$

For the right inequality (7), let $G(x) = x\pi(x)$. Formula (19) can be rewritten as

$$B(x) - B(\sqrt{x}) < G(x) - G(\sqrt{x}) + xF(x). \quad (20)$$

Applying inequality (20) with x replaced by \sqrt{x} , $x^{1/4}$, ..., $x^{1/2^s}$ until $x^{1/2^s} < 2$ and summing up all these inequalities one gets

$$B(x) - B(1) < G(x) + \sum_{i=0}^s x^{1/2^i} F(x^{1/2^i}). \quad (21)$$

The function $F(x)$ is obviously increasing. Hence,

$$B(x) < 1 + G(x) + F(x) \sum_{i=0}^s x^{1/2^i}. \quad (22)$$

To finish the argument, we show that

$$x \geq \sum_{i=1}^s x^{1/2^i}. \quad (23)$$

Proceed by induction on s . If $s = 0$, there is nothing to prove. If $s = 1$, this just says that $x > \sqrt{x}$ which is obvious. Finally, if $s \geq 2$, it follows that $x \geq 4$. In particular, $x \geq 2\sqrt{x}$ or $x - \sqrt{x} \geq \sqrt{x}$. Rewriting inequality (23) as

s

\dots

which is precisely inequality (23) for \sqrt{x} . This completes the induction step. Via inequality (22), inequality (23) implies

$$B(x) < 1 + x\pi(x) + 2xF(x) = 1 + x\pi(x) + 2x(3.1 + 1.25 \log \log x) \quad (24)$$

or

$$A(x) < \pi(x) + \frac{1}{x} + 6.2 + 2.5 \log \log x = \pi(x) + E(x).$$

Applications

From the theorem, it follows easily that for every $\epsilon > 0$ there exists x_0 such that

$$A(x) < (1 + \epsilon) \frac{x}{\log x}. \quad (25)$$

In practice, finding a lower bound on x_0 for a given ϵ , one simply uses the theorem and the estimate

$$\pi(x) < \frac{x}{\log x} \left(1 + \frac{3}{2 \log x} \right) \quad \text{for } x > 1. \quad (26)$$

(see [5]). By (7) and (26), it now follows that (25) is satisfied provided that

$$\frac{x}{\log x} > \frac{1}{\epsilon} \left(\frac{3}{2 \log^2 x} + E(x) \right).$$

For example, when $\epsilon = 1$, one gets

$$A(x) < 2 \frac{x}{\log x} \quad \text{for } x \geq 64, \quad (27)$$

for $\epsilon = .5$, one gets

$$A(x) < 1.5 \frac{x}{\log x} \quad \text{for } x \geq 254 \quad (28)$$

and for $\epsilon = 0.1$ one gets

$$A(x) < 1.1 \frac{x}{\log x} \quad \text{for } x \geq 3298109. \quad (29)$$

Of course, inequalities (27)-(29) may hold even below the smallest values shown above but this needs to be checked computationally.

In the same spirit, by using the theorem and the estimation

$$\pi(x) > \frac{x}{\log x} \left(1 + \frac{1}{2 \log x} \right) \quad \text{for } x \geq 59$$

(see [5]) one can compute, for any given ϵ , an initial value x_0 such that

$$A(x) > (.5 - \epsilon) \frac{x}{\log x} \quad \text{for } x > x_0.$$

For example, when $\epsilon = 1/6$ one gets

$$A(x) > \frac{1}{3} \frac{x}{\log x} \quad \text{for } x \geq 59. \quad (30)$$

Inequality (30) above is better than the inequality appearing on page 62 in [2] which asserts that for every $\alpha > 0$ there exists x_0 such that

$$A(x) > x^{\alpha/x} \quad \text{for } x > x_0 \quad (31)$$

because the right side of (31) is bounded and the right side of (30) isn't!

A diophantine equation

In this section we present an application to a diophantine equation. The application is not of the theorem per se, but rather of the counting method used to prove the theorem.

Since S is defined in terms of factorials, it seems natural to ask how often the product $S(1) \cdot S(2) \cdot \dots \cdot S(n)$ happens to be a factorial.

Proposition.

The only solutions of

$$S(1) \cdot S(2) \cdot \dots \cdot S(n) = m! \quad (32)$$

are given by $n = m \in \{1, 2, \dots, 5\}$.

Proof.

We show that the given equation has no solutions for $n \geq 50$. Assume that this is not so. Let P be the largest prime number smaller than n . By Tchebysheff's theorem, we know that $P \geq n/2$. Since $S(P) = P$, it follows that $P \mid m!$. In particular, $P \leq m$. Hence, $m \geq n/2$.

We now compute an upper bound for the order of 2 in $S(1) \cdot S(2) \cdot \dots \cdot S(n)$. Fix some $\beta \geq 1$ and assume that k is such that $2^\beta \parallel S(k)$. Since

$$S(k) = \max(S(p^\alpha) \mid p^\alpha \parallel k),$$

it follows that $2^\beta \parallel S(p^\alpha)$ for some $p^\alpha \parallel k$.

We distinguish two situations:

Case 1.

p is odd. In this case, $2^\beta p \mid S(p^\alpha)$. If $\beta = 1$, then $\alpha = 2$. If $\beta = 2$, then $\alpha = 4$. For $\beta \geq 3$, one can easily check that $\alpha \geq 2^\beta - \beta + 1$ (indeed, if $\alpha \leq 2^\beta - \beta$, then one can check that $p^\alpha \mid (2^\beta p - 1)!$ which contradicts the definition of S). In particular, $p^{2^\beta - \beta + 1} \mid k$. Since $2^{x-1} \geq x + 1$ for $x \geq 3$, it follows that $\alpha \geq 2^{\beta-1} + 2$. Since $k \leq n$, the above arguments show that there are at most

$$\frac{n}{p^{2^\beta}} \quad \text{for } \beta = 1, 2$$

and

$$\frac{n}{p^{2^{\beta-1}+2}} \quad \text{for } \beta \geq 3$$

integers k in the interval $[1, n]$ for which $p \mid k$, $S(k) = S(p^\alpha)$, where α is such that $p^\alpha \parallel k$ and $2^\beta \parallel S(k)$.

Case 2.

$p = 2$. If $\beta = 1$, then $k = 2$. If $\beta = 2$, then $k = 4$. Assume now that $\beta \geq 3$. By an argument similar to the one employed at Case 1, one gets in this case that $\alpha \geq 2^\beta - \beta$. Since $2^\alpha \parallel k$, it follows that $2^{2^\beta - \beta} \mid k$. Since $k \leq n$, it follows that there are at most

$$\frac{n}{2^{2^\beta - \beta}}$$

such k 's.

From the above analysis, it follows that the order at which 2 divides $S(1) \cdot S(2) \cdot \dots \cdot S(n)$ is at most

$$e_2 < 3 + n \sum_{\substack{p \leq n \\ p \text{ odd}}} \left(\frac{1}{p^2} + \frac{2}{p^4} + \sum_{\beta \geq 3} \frac{\beta}{p^{2^{\beta-1}+2}} \right) + n \sum_{\beta \geq 3} \frac{\beta}{2^{2^\beta - \beta}}. \quad (38)$$

(the number 3 in the above formula counts the contributions of $S(2) = 2$ and $S(4) = 4$). We now bound each one of the two sums above.

For fixed p , one has

$$\frac{1}{p^2} + \frac{2}{p^4} + \sum_{\beta \geq 3} \frac{\beta}{p^{2^{\beta-1}+2}} = \frac{1}{p^2} + \frac{2}{p^4} + \frac{3}{p^6} + \frac{4}{p^{10}} + \dots < \sum_{\gamma \geq 1} \frac{\gamma}{p^{2^\gamma}} = \frac{p^2}{(p^2 - 1)^2}. \quad (39)$$

Hence,

$$\sum_{\substack{p \leq n \\ p \text{ odd}}} \left(\frac{1}{p^2} + \frac{2}{p^4} + \sum_{\beta \geq 3} \frac{\beta}{p^{2^{\beta-1}+2}} \right) < \sum_{p \text{ odd}} \frac{p^2}{(p^2-1)^2} < .245 \quad (40)$$

We now bound the second sum:

$$\begin{aligned} \sum_{\beta \geq 3} \frac{\beta}{2^{2^{\beta}-\beta}} &= \frac{3}{2^5} + \frac{4}{2^{12}} + \frac{5}{2^{27}} + \dots < \frac{3}{2^6} + \sum_{\beta \geq 3} \frac{\beta}{2^{2+4(\beta-2)}} = \\ &= \frac{3}{2^6} + \frac{1}{4} \left(\sum_{\gamma \geq 1} \frac{\gamma+2}{16^\gamma} \right) = \frac{3}{2^6} + \frac{1}{4} \left(\frac{15}{16} + \frac{31}{225} \right) < .099 \end{aligned} \quad (41)$$

From inequalities (38), (40) and (41), it follows that

$$e_2 < 3 + .344n. \quad (42)$$

We now compute a lower bound for e_2 . Since $e_2 = e_2(m!)$, it follows, from Lemme 1 in [1] and from the fact that $m \geq n/2$, that

$$e_2 \geq m - \frac{\log(m+1)}{\log 2} \geq \frac{n}{2} - \frac{\log(n/2+1)}{\log 2}. \quad (43)$$

From inequalities (42) and (43), it follows that

$$3 + .344n \geq .5n - \frac{\log(.5n+1)}{\log 2},$$

which gives $n \leq 50$. One can now compute $S(1) \cdot S(2) \cdot \dots \cdot S(n)$ for all $n \leq 50$ to conclude that the only instances when these products are factorials are $n = 1, 2, \dots, 5$.

We conclude suggesting the following problem:

Problem.

Find all positive integers n such that $S(1), S(2), \dots, S(n^2)$ can be arranged in a latin square.

The above problem appeared as Problem 24 in SNJ 9, (1994) but the range of solutions was restricted to $\{2, 3, 4, 5, 7, 8, 10\}$. The published solution was based on the simple observation that the sum of all entries in an $n \times n$ latin square has to be a multiple of n . By computing the sums $B(x^2)$ for x in the above range, one concluded that $B(x^2) \not\equiv 0 \pmod{x}$ which meant that there is no solution for such x 'ses. It is unlikely that this argument can be extended to cover the general case. One should notice that from our theorem, it follows that if a solution exists for some $n > 1$, then the size of the common sums of all entries belonging to the same row (or column) is $\cong n\pi(n^2)$.

Addendum

After this paper was written, it was pointed out to us by an anonymous referee that Finch [3] proved recently a much stronger statement, namely that

$$\lim_{x \rightarrow \infty} \frac{\log(x)}{x} \cdot A(x) = \frac{\pi^2}{12} = 0.82246703\dots \quad (44)$$

Finch's result is better than our result which only shows that the limsup of the expression $\log(x)A(x)/x$ when x goes to infinity is in the interval $[0.5, 1]$.

References

- [1] Y. Bugeaud & M. Laurent, "Minoration effective de la distance p -adique entre puissances de nombres algébriques", *J. Number Theory* **61** (1996), pp. 311-342.
- [2] C. Dumitrescu & V. Seleacu, "The Smarandache Function", Erhus U. Press, 1996.
- [3] S.R. Finch, "Moments of the Smarandache Function", *SNJ* **11**, No. 1-2-3 (2000), p. 140-142.
- [4] H. Ibsted, "Surfing on the ocean of numbers", Erhus U. Press, 1997.
- [5] J. B. Rosser & L. Schoenfeld, "Approximate formulas for some functions of prime numbers", *Illinois J. of Math.* **6** (1962), pp. 64-94.
- [6] S. Tabirca & T. Tabirca, "Two functions in number theory and some upper bounds for the Smarandache's function", *SNJ* **9** No. 1-2, (1998), pp. 82-91.

1991 AMS Subject Classification: 11A25, 11L20, 11L26.