PRIME NUMBER CONJECTURES FROM THE SHAPIRO CLASS STRUCTURE

HARTOSH SINGH BAL AND GAURAV BHATNAGAR*

ABSTRACT. The height $H(n)$ of n, introduced by Pillai in 1929, is the smallest positive integer i such that the ith iterate of Euler's totient function at n is 1. H. N. Shapiro (1943) studied the structure of the set of all numbers at a height. We state a formula for the height function due to Shapiro and use it to list steps to generate numbers at any height. This turns out to be a useful way to think of this construct. In particular, we extend some results of Shapiro regarding the largest odd numbers at a height. We present some theoretical and computational evidence to show that H and its relatives are closely related to the important functions of number theory, namely $\pi(n)$ and the nth prime p_n. We conjecture formulas for $\pi(n)$ and p_n in terms of the height function.

1. INTRODUCTION

The principal object of our investigation is a number theoretic function H that we call the *height function*. It is defined as follows. Let $H(1) := 0$, and

$$H(n) := H(\varphi(n)) + 1, \tag{1.1}$$

for $n = 2, 3, 4, \ldots$. Here $\varphi(n)$ denotes Euler's totient function, the number of positive integers less than n which are co-prime to n. The first few values of H are: $0, 1, 2, 2, 3, 2$. Let C_k be the set of numbers at height k, that is,

$$C_k := \{n : H(n) = k\}.$$

We call C_k the *Shapiro classes* in honor of Harold N. Shapiro [9] who studied these classes first. In this paper, we examine the Shapiro class structure, and show that the height function and its relatives are very closely related to the important functions of number theory, namely $\pi(n)$, the number of primes less than or equal to n, and p_n the nth prime number.

Shapiro arrived at these classes by considering iterates of Euler's totient function. For $i \geq 1$, we denote by

$$\varphi^i(n) := \underbrace{\varphi(\varphi(\cdots))}_{i \text{ times}}(n)$$

the ith iterate of φ. Then the height function can also be defined as follows. Let $H(1) := 0$. For $n > 1$, let $H(n)$ be the smallest number i such that $\varphi^i(n) = 1$. The height function (with this definition) was studied first by Pillai [7]. The Shapiro classes (by other names) have been studied earlier by Shapiro [9], Erdős, Granville,

Date: April 9, 2019.

2010 *Mathematics Subject Classification.* Primary 11A25; Secondary 11A41, 11N37.

Key words and phrases. Euler's totient function, Iterated totient function, Prime numbers, Multiplicative Functions, Chebyshev's theorem, Shapiro Classes.

*Research partially supported by the Austrian Science Fund (FWF), grant F50-N15, in the framework of the Special Research Program "Algorithmic and Enumerative Combinatorics".

Pomerance and Spiro [3], and others [2, 4, 5, 6]. We have mildly modified Shapiro's formulation.

In Section 2 we find an alternative, inductive, approach to generate Shapiro classes. To do so, we require a formula for $H(n)$ due to Shapiro. In Section 3, we apply our ideas to extend a theorem of Shapiro about the largest odd numbers at a given height. This rests upon a property that is obvious from our construction, but not observed earlier, regarding the largest possible prime number at a height. This, and results of Erdős et.al. [3], led us to look for number theoretic information from this structure.

Since $H(n)$ is not an increasing function, we consider the sum of heights function, defined as: $S(0) := 0$, and

$$S(n) := \sum_{k=1}^{n} H(k). \tag{1.2}$$

The structure consisting of Shapiro classes allows us to obtain number theoretic information quite easily. It appears that elementary techniques, such as those found in the textbooks of Apostol [1] and Shapiro [10], can be modified to express classical theorems in terms of functions related to the height function. We illustrate this idea in Section 4, by proving Chebyshev-type theorems, that is, inequalities for $\pi(n)$ and p_n in terms of n and $S(n)$.

Our results, and the results/conjectures in Erdős, Granville, Pomerance and Spiro [3], motivate the experimental work presented in Section 5. As a sample, consider Figure 1, a plot of $S(n)$ and p_n on the same set of axes. The remarkable agreement of these graphs (upto $n = 5000$) suggests the importance of this function to the theory of prime numbers. From numerical computations (see Tables 1 and 5), it appears that $S(n)$ is a better approximation to p_n than $n \log n$, at-least until $n = 6 \times 10^7$.

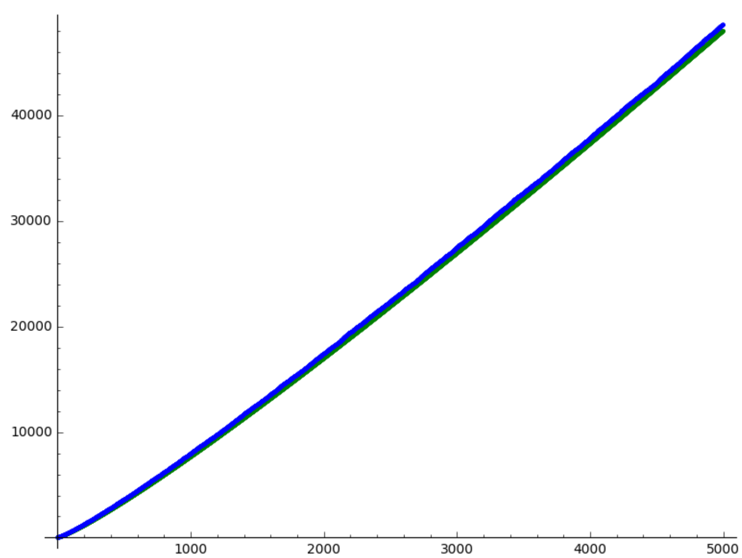

FIGURE 1. Comparison of the nth prime (top) with $S(n) = \sum_{k=1}^{n} H(k)$.

n	p_n	$S(n)$	$S(n) - p_n$	$\lfloor n \log n \rfloor$	$\lfloor n \log n \rfloor - p_n$	$\frac{S(n)-p_n}{p_n}$
10	29	22	-7	23	-6	-24.1%
10^2	541	486	-55	460	-81	-10.2%
10^3	7919	7640	-279	6907	-1012	-3.52%
10^4	104729	104488	-241	92103	-12626	-0.230%
10^5	1299709	1325890	26181	1151292	-148417	2.01%
10^6	15485863	16069024	583161	13815510	-1670353	3.77%
10^7	179424673	188786066	9361393	161180956	-18243717	5.22%

Table 1. $S(n)$ against $\lfloor n \log n \rfloor$

2. Listing Shapiro classes

We begin our study of the height function and Shapiro classes. The objective of this section to build some useful intuition, by writing a set of rules to generate Shapiro classes inductively. Towards this end, we first prove a formula due to Shapiro for calculating the height of a function. As a corollary, we show the additive nature of the height function. We illustrate the usefulness of these rules by obtaining several elementary properties of the Shapiro classes.

It is instructive to compute the first few classes. Table 2 gives the first four Shapiro classes. We begin with 1 in C_0 and 2 in C_1. Now since $\varphi(3) = 2$, we find that $H(3) = H(2) + 1 = 2$. So $3 \in C_2$. Similarly, we see that 4 has height 2, 5 has height 3 and 6 has height 2. By computing the values of $H(n)$ for a few more

k	C_k
3	$5, 7, 8, 9, 10, 12, 14, 18$
2	$3, 4, 6$
1	2
0	1

Table 2. C_k for $k = 0, 1, 2, 3$.

values, a few rules for finding members of the Shapiro class become evident. For example, consider the following rules for finding members of C_k from C_{k-1}. We have

- If m is an odd number, then the height of $2m$ is the same as m.
- If m is an even number, then $H(2m) = H(m) + 1$.
- If m is an odd number, then $H(3m) = H(m) + 1$.
- If p is a prime number, then $H(p) = H(p - 1) + 1$. Thus $p - 1$ is an even number at one height lower than p.

We can use these rules to generate the members of C_4. If we multiply all the even numbers of C_3 by 2, we obtain 16, 20, 24, 28, 36. Now multiplying all the odd numbers by 3, we find that 15, 21, 27 are in C_4. Next, consider $8 + 1 = 9$, $10 + 1 = 11$, $12 + 1 = 13$, $14 + 1 = 15$, $18 + 1 = 19$. Of these, 11, 13, and 19, are primes, and are thus at height 4. Finally, the odd numbers already obtained are: $11, 13, 15, 19, 21, 27$. Multiplying them by 2, we find that $22, 26, 30, 38, 42$, and 54 are also in C_4. The reader may verify that we have obtained all the numbers at height 4.

The rules work to generate all the elements of C_4 from C_3, but they are not comprehensive. The complete set of rules will appear shortly, as an application of a formula for $H(n)$.

Theorem 2.1 (Shapiro [9]). *Let $n > 1$ with prime factorization $n = 2^\alpha p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$, where $p_1, p_2, p_3, \dots, p_r$ are distinct odd primes, and $\alpha, \alpha_1, \alpha_2, \dots, \alpha_r \geq 0$. Then,*

$$H(n) = \begin{cases} \alpha + \sum_{i=1}^{r} \alpha_i (H(p_i) - 1) & \text{if } \alpha > 0 \\ \sum_{i=1}^{r} \alpha_i (H(p_i) - 1) + 1 & \text{if } \alpha = 0. \end{cases} \tag{2.1}$$

Before giving a proof of this theorem, we obtain a corollary which indicates the additive nature of the height function.

Corollary 2.2 (Shapiro [9]). *Let n and m be natural numbers with prime factorizations $m = 2^\alpha p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ and $n = 2^\beta p_1^{\beta_1} p_2^{\beta_2} \dots p_r^{\beta_r}$, where $\alpha, \beta, \alpha_i, \beta_i \geq 0$. Then we have*

$$H(nm) = \begin{cases} H(m) + H(n), & \text{if } \alpha, \beta > 0; \\ H(m) + H(n) - 1, & \text{if } \alpha = 0 \text{ or } \beta = 0. \end{cases} \tag{2.2}$$

Remark. Shapiro [9] considered Corollary 2.2 as his fundamental theorem. The formula (2.1) follows immediately. The proof presented below is closely related to Shapiro's proof of Corollary 2.2.

Proof. Note that if both α and β are strictly positive, then

$$\begin{aligned} H(mn) &= \alpha + \beta + \sum_{i=1}^{r} (\alpha_i + \beta_i) \left(H(p_i) - 1 \right) \\ &= H(n) + H(m), \end{aligned}$$

as required. Similarly, the second part of the formula follows by considering the cases where only one of α and β is 0, and where both α and β are 0. □

Proof of Theorem 2.1. We prove this formula by induction on the number of primes in the prime factorization of n. Remarkably, we will require the computation in the corollary to prove the theorem.

First we prove the formula for numbers n of the form $n = 2^\alpha$, by induction. Clearly, the formula holds for $\alpha = 1$. Suppose it holds for $\alpha = k$. Then

$$H(2^{k+1}) = H\big(\varphi(2^{k+1})\big) + 1 = H(2^k) + 1 = k + 1$$

by the induction hypothesis. This proves the formula for powers of 2.

Next, we consider numbers n of the form $n = 2^\alpha \cdot 3^{\alpha_1}$. Consider first the case $\alpha > 0$. Again we will prove this formula by induction, this time on α_1. It is easy to verify the formula for $\alpha_1 = 1$. Next suppose the formula is true for $\alpha_1 = k$. For $n = 2^\alpha \cdot 3^{k+1}$, we have

$$H(2^\alpha \cdot 3^{k+1}) = H\big(\varphi(2^\alpha \cdot 3^{k+1})\big) + 1 = H(2^\alpha \cdot 3^k) + 1 = \alpha + (k+1) \cdot (H(3) - 1)$$

as required.

Next we consider the case $\alpha = 0$. In this case, we can obtain the formula from the previous case. Since $\alpha_1 \geq 1$, We have

$$H(n) = H\bigl(\varphi(n)\bigr) + 1 = H\bigl(2 \cdot 3^{\alpha_1 - 1}\bigr) + (H(3) - 1) = 1 + \alpha_1(H(3) - 1)$$

which is the required formula.

Next, as the induction hypothesis, we assume that the formula is valid for numbers of the form

$$2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r},$$

where $p_1 < p_2 < \cdots < p_r$ are the first r odd primes. We will show the formula for

$$n = 2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r} p_{r+1}^{\alpha_{r+1}}.$$

We first consider the case $\alpha > 0$. Again we will prove this by induction on α_{r+1}. Let

$$\begin{aligned} n' &= \varphi(p_{r+1}) = 2^\beta p_1^{\beta_1} \cdots p_r^{\beta_r} \\ m' &= \varphi(n/p_{r+1}^{\alpha_{r+1}}) = 2^\gamma p_1^{\gamma_1} \cdots p_r^{\gamma_r}. \end{aligned}$$

We will first prove the formula for $\alpha_{r+1} = 1$. Then,

$$\varphi(n) = \varphi(p_{r+1})\varphi(n/p_{r+1}),$$

so

$$\begin{aligned} H(n) &= H(\varphi(n)) + 1 \\ &= H(m'n') + 1 \\ &= H(m') + H(n') + 1. \end{aligned}$$

The last equality follows by the argument in Corollary 2.2 applied to even numbers that have only 2, $p_1, \ldots, p_r$ in their prime factorization. (By the induction hypothesis, we have assumed that (2.1) holds for such numbers, and so this argument can be used in our proof.) Thus we find that

$$\begin{aligned} H(n) &= H(m') + H(n') + 1 \\ &= H\bigl(2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r}\bigr) - 1 + H(p_{r+1}) - 1 + 1 \\ &= \alpha + \sum_{i=1}^{r} \alpha_i\bigl(H(p_i) - 1\bigr) + H(p_{r+1}) - 1, \end{aligned}$$

as required.

Next, we suppose the formula is true for $\alpha_{r+1} = k$. We will prove it for

$$n = 2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r} p_{r+1}^{k+1}.$$

Again, let m', n', β_i, γ_i be as before. Then we have

$$\begin{aligned} H(n) &= H(\varphi(n)) + 1 \\ &= H\bigl(2^{\beta+\gamma} p_1^{\beta_1+\gamma_1} \cdots p_r^{\beta_r+\gamma_r} p_{r+1}^k\bigr) + 1 \\ &= \beta + \gamma + \sum_{i=1}^{r} (\beta_i + \gamma_i)\bigl(H(p_i) - 1\bigr) + k(H(p_{r+1} - 1)) + 1 \\ &= H(\varphi(p_{r+1})) + H\bigl(\varphi(2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r})\bigr) + k(H(p_{r+1}) - 1) + 1 \\ &= H\bigl(2^\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r}\bigr) + (k+1)(H(p_{r+1}) - 1), \end{aligned}$$

which proves the formula for $\alpha > 0$.

Finally, we prove the formula for $\alpha = 0$ and $\alpha_{r+1} \geq 1$. Let

$$n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} p_{r+1}^{\alpha_{r+1}}.$$

Now

$$\begin{aligned} H(n) &= H(\varphi(n)) + 1 \\ &= H\big(2^{\beta+\gamma} p_1^{\beta_1+\gamma_1} \cdots p_r^{\beta_r+\gamma_r} p_{r+1}^{\alpha_{r+1}-1}\big) + 1 \\ &= H(p_1^{\alpha_1} \cdots p_r^{\alpha_r}) + \alpha_{r+1}(H(p_{r+1}) - 1) + 1, \end{aligned}$$

where the last step is obtained as in the $\alpha > 0$ case. From here we obtain the formula by applying the induction hypothesis. □

We return to our set of rules to generate the numbers at a given height from the previous ones. Corollary 2.2 tells us what happens if we multiply a number by 5. Recall that the height of 5 is 3. So if m is any number, then $H(5m) = H(m) + 2$. Thus to obtain numbers at height k we have to multiply numbers at height $k-2$ by 5 and the other prime at height 3, namely 7. In general, to obtain all the numbers at a height, we have to consider all the primes at lower heights. We are now in a position to create a comprehensive list of rules to generate C_k.

Let Q_k denote the set of primes at height k. That is, for $k = 1, 2, 3, \ldots$, define

$$Q_k := \{p : p \text{ a prime and } H(p) = k\}.$$

For example:

$$Q_1 = \{2\}, Q_2 = \{3\}, Q_3 = \{5, 7\}, Q_4 = \{11, 13, 19\}, Q_5 = \{17, 23, 29, 31, 37, 43\}.$$

Theorem 2.3. *Suppose C_1, C_2, ..., C_{k-1} are known. The steps for obtaining C_k, that is, the elements at height k are as follows*

(1) *Multiply each even element of C_{k-1} by 2.*
(2) *Multiply each odd element of C_{k-1} by 3.*
(3) *Multiply each element of C_{k-k_1} by elements of Q_{k_1+1}, where $k_1 = 2, 3, \ldots$.*
(4) *If $m+1$ is a prime where m is an even number in C_{k-1}, then $m + 1 \in C_k$.*
(5) *Multiply each odd number obtained already in C_k by 2.*

Remark. Shapiro considered a slightly different function $C(n)$, where $C(1) = 0$ and $C(n) = H(n) - 1$, for $n > 1$. So $C(2) = 0$ whereas $H(2) = 1$. Theorem 2.3 is the primary reason why we have deviated from Shapiro's formulation.

Proof. The proof is a formalization of our earlier discussion. Consider the following statements.

(1) If $x \in C_{k-1}$ is even, then $H(2x) = k$.
(2) if $x \in C_{k-1}$, then $H(3x) = k$.
(3) if $x \in C_{k-k_1}$ and $p \in Q_{k_1+1}$ then $H(px) = k$.

All these statements follow from Corollary 2.2.

Step 1 generates all even numbers of the form $2^a m$ where $a > 1$, and m is odd, i.e., numbers divisible by 4. The steps (2) and (3) will generate all composite, odd, numbers in C_k.

Since $\phi(p) = p - 1$, and for all odd primes $p - 1$ is even, thus all prime numbers are obtained by adding 1 to even numbers in C_{k-1}. This explains (4).

It remains to generate numbers of the form $2m$ where m is odd. Since $H(2n) = H(n)$ when n is odd, we must multiply each odd number obtained by steps (1)-(4) by 2. This will generate all the elements of C_k. □

Remark. In Step (3), we need to multiply only those elements of C_{k-k_1} that are odd and not divisible by 3.

Now it is easy to generate some more sets C_k. The elements of the first few sets C_k are given in Table 3. The prime numbers are given in bold.

k	C_k
6	**41**, **47**, 51, **53**, 55, **59**, **61**, 64, 65, **67**, 68, 69, **71**, **73**, 75, 77, **79**, 80, 82, 87, 88, 91, 92, 93, 94, 95, 96, 99, 100, 102, 104, 105, 106, **109**, 110, 111, 112, 116, 117, 118, 120, 122, 124, **127**, 129, 130, 132, 133, 134, 135, 138, 140, 142, 144, 146, 147, 148, 150, 152, 154, 156, 158, **163**, 168, 171, 172, 174, 180, 182, 186, 189, 190, 196, 198, 210, 216, 218, 222, 228, 234, 243, 252, 254, 258, 266, 270, 294, 324, 326, 342, 378, 486
5	**17**, **23**, 25, **29**, **31**, 32, 33, 34, 35, **37**, 39, 40, **43**, 44, 45, 46, 48, 49, 50, 52, 56, 57, 58, 60, 62, 63, 66, 70, 72, 74, 76, 78, 81, 84, 86, 90, 98, 108, 114, 126, 162
4	**11**, **13**, 15, 16, **19**, 20, 21, 22, 24, 26, 27, 28, 30, 36, 38, 42, 54
3	**5**, **7**, 8, 9, 10, 12, 14, 18
2	**3**, 4, 6
1	**2**
0	1

TABLE 3. Numbers with height ≤ 6. The primes are in bold.

Theorems 2.1 and 2.3 are very useful to think about the C_k. We illustrate this in the following observations. These have been found previously by other authors.

Observations.

(1) The number 2^k comes at height k.

(2) The smallest even number at height k is 2^k. To see this, consider the following argument.

An even number in C_k can arise in two ways. If it is obtained by multiplying an element of C_{k-1} by 2, it is bigger than or equal to 2^k by induction. The other possibility is that it is of the form $2m$, where m is an odd number at height k.

If m is a prime number, then it is obtained by adding 1 to an even number at height $k-1$, so it is bigger than 2^{k-1} by induction. This implies $2m > 2^k$.

Suppose m is an odd, composite number with prime factorization $m = p_1^{\alpha_1} \dots p_r^{\alpha_2}$, with $H(p_i) = k_i + 1$. Then by Theorem 2.1, $H(m) = k$ implies that

$$k - 1 = \sum_{i=1}^{r} \alpha_i k_r.$$

But by induction, as above, we must have $p_i > 2^{k_i}$ for $i = 1, 2, \dots, r$. Thus

$$m = \prod_i p_i^{\alpha_i} > 2^{\sum_i \alpha_i k_i} = 2^{k-1}.$$

Again, this implies that $2m > 2^k$.

(3) Any odd number in C_k is bigger than 2^{k-1}. This follows from the above argument.
(4) If $n \leq 2^k$, then $H(n) \leq k$. This is what the two items above amount to; this was implicit in Pillai [7], and stated by Shapiro [9].
(5) (Shapiro [9]) The numbers at height h that are less than 2^h are all odd.
(6) (Shapiro [9]) The largest odd number at height h is 3^{h-1}. This follows from induction and Theorem 2.3.
(7) (Pillai [7]) The largest even number at height h is $2 \cdot 3^{h-1}$.
(8) (Pillai [7]) Since any number at height h is between 2^{h-1} and $2 \cdot 3^{h-1}$, we have the inequalities:

$$\frac{\log n/2}{\log 3} + 1 \leq H(n) \leq \frac{\log n}{\log 2} + 1. \tag{2.3}$$

All the observations above were noted by previous authors. One can ask whether Theorem 2.3 gives any new information. Indeed, there is one very important observation missed by previous authors.

The largest prime at a level k is less than or equal to $2 \cdot 3^{k-2} + 1$. This is obvious from Step (4) and Pillai's observation that the largest even number at each height is $2 \cdot 3^{k-2}$.

In the next section, we show how this observation can be used to obtain information about the numbers appearing at the end of each class, thus extending some of Shapiro's results.

3. On the Shapiro Class structure

The objective of this section is to illustrate the application of Theorems 2.1 and 2.3, by extending some results of Shapiro [9]. Our theorem in this section is a characterization of the last few numbers at each height.

As noted above, the largest prime at a level k is less than or equal to $2 \cdot 3^{k-2} + 1$. This upper bound is met for many k. The smallest such examples are obtained when $k = 2, 3, 4, 6, 7, 8, 11, 18, 19$. The first few examples, corresponding to these values of k, are $3, 7, 19, 163, 487, 1458, 39367, 86093443, 258280327$. Primes of this kind (cf. OEIS [11, A003306]) play an important role in our theorem. Let $\widehat{P}$ denote the set of primes of this form, that is,

$$\widehat{P} := \{p : p \text{ a prime}, p = 2 \cdot 3^{k-2} + 1 \text{ for some } k \geq 2\}.$$

To prove our main result, we require a useful proposition.

Proposition 3.1. *Let $k > 2$. Let m be an odd, composite number not divisible by 3 at height k. Then,*

$$m < 2 \cdot 3^{k-2} + 1.$$

Before proving the proposition, we prove a special case, where m is of the form pq or p^2.

Lemma 3.2. *Let $k > 2$. Let p and q be (possibly the same) primes. Suppose $p, q \neq 2, 3$, and $H(pq) = k$. Then*

$$pq < 2 \cdot 3^{k-2} + 1.$$

Proof. Let $H(p) = a$ and $H(q) = b$. Since p and q are not $2, 3$, we must have $a, b > 2$. Since $p \leq 2 \cdot 3^{a-2} + 1$ and $q \leq 2 \cdot 3^{b-2} + 1$, we must have

$$pq \leq 4 \cdot 3^{a+b-4} + 2 \cdot 3^{a-2} + 2 \cdot 3^{b-2} + 1.$$

Now from Corollary 2.2, we see that $k = H(pq) = a + b - 1$. Now consider

$$\begin{aligned} 2 \cdot 3^{a+b-3} + 1 &= 6 \cdot 3^{a+b-4} + 1 \\ &= 4 \cdot 3^{a+b-4} + 3^{b-2} \cdot 3^{a-2} + 3^{a-2} \cdot 3^{b-2} + 1 \\ &> 4 \cdot 3^{a+b-4} + 2 \cdot 3^{a-2} + 2 \cdot 3^{b-2} + 1, \end{aligned}$$

since $a > 2$ implies that $3^{a-2} > 2$. This completes the proof of the lemma. □

Proof of Proposition 3.1. We use induction on k. For $k = 3, 4$, the statement is vacuously true. Let $k > 5$. Let $p|m$, $H(p) = a$ and $H(m/p) = b$. Then $a, b > 2$, otherwise m has to be divisible by 3. Further $b < k$ since $a + b - 1 = k$. Thus by the induction hypothesis, we must have

$$m/p < 2 \cdot 3^{b-2} + 1.$$

Of course, since p is a prime, $p < 2 \cdot 3^{a-2} + 1$. Now by using the same argument as in Lemma 3.2, we see that

$$m < 2 \cdot 3^{a+b-3} + 1 = 2 \cdot 3^{k-2} + 1,$$

as required. □

Next, we determine all the numbers in the set

$$R_k := \{n \in C_k : 2^2 3^{k-2} < n \leq 2 \cdot 3^{k-1}\}.$$

These numbers are the largest elements at height k.

Theorem 3.3. *Let $k > 2$. The set R_k comprises all the numbers of the form m, where*

$$m = 2 \cdot 3^{k-a} p,$$

where $p \in \widehat{P}$, with $H(p) = a$ and $a \leq k$.

Proof. We let S_k denote the set

$$S_k := \{m : m = 2 \cdot 3^{k-a} p \text{ for some } p \in \widehat{P}, \text{ with } H(p) = a \text{ and } a \leq k\}.$$

We want to show that $R_k = S_k$.

First observe from (2.1) that if $m \in S_k$, then $H(m) = k$. Further, if $m \in S_k$, then

$$2^2 3^{k-2} < m \leq 2 \cdot 3^{k-1}.$$

This follows from

$$2 \cdot 3^{k-a} p = 2 \cdot 3^{k-1} \Big(\frac{2}{3} + \frac{1}{3^{a-1}} \Big). \tag{3.1}$$

Thus $S_k \subset R_k$.

To show the converse, we apply an inductive argument using Theorem 2.3.

For $k = 3$, $R_3 = \{14, 18\} = S_3$. So let $k > 3$.

Observe that R_k has only even numbers. This is because all odd numbers at height k are less than or equal to 3^{k-1}, and $3^{k-1} < 4 \cdot 3^{k-2}$.

Even numbers are obtained from Step 1 or Step 5 in Theorem 2.3. However, since there are no numbers in C_{k-1} that are bigger than $2 \cdot 3^{k-2}$, none of the numbers in R_k are obtained from Step 1. Thus all the numbers in R_k are of the form $2r$, where r is an odd number at height k, and

$$2 \cdot 3^{k-2} < r \leq 3^{k-1}.$$

By Proposition 3.1, all odd composite numbers not divisible by 3 are less than $2 \cdot 3^{k-2} + 1$. Thus there are only two possibilities for r.

(1) r is a prime of the form $2 \cdot 3^{k-2} + 1$ with $H(r) = k$, i.e., $2r \in S_k$ as required.
(2) r is divisible by 3.

In case r is divisible by 3, it is obtained from Step 2, and there is an s such that $2 \cdot 3^{k-3} < s \leq 3^{k-2}$, with $r = 3s$. So $2s \in R_{k-1}$. By the induction hypothesis, $2s = 2 \cdot 3^{k-1-a}p$, for some $p \in \widehat{P}$ with $H(p) = a \leq k - 1$. So

$$2r = 2 \cdot 3^{k-a}p,$$

and $2r \in S_k$. This completes the proof of $S_k \subset R_k$. □

To state our next result, we require some notation. Let p_1, p_2, $p_3, \ldots$ be the elements of $\widehat{P}$ listed in increasing order, where $H(p_i) = a_i$, and $a_1 < a_2 < a_3 < \cdots$. For example, the first few pairs (p_i, a_i) are $(3, 2)$, $(7, 3)$, $(19, 4)$, and $(163, 6)$.

Corollary 3.4. *Let $k > 2$. Let $p_i \in \widehat{P}$, $i = 1, 2, \ldots, r$, be as above, with r the largest such that $a_r \leq k$. At height k, the largest numbers, in decreasing order, are:*

$$2 \cdot 3^{k-1}, 2 \cdot 3^{k-3} \cdot 7, 2 \cdot 3^{k-4} \cdot 19, 2 \cdot 3^{k-6} \cdot 163, \ldots, 2 \cdot 3^{k-a_r} \cdot p_r, 2^2 \cdot 3^{k-2}.$$

Proof. Let $m_i = 2 \cdot 3^{k-a_r}p_i$. Note that from (3.1), it follows that if $a_i > a_j$, then $m_i < m_j$. Thus $2 \cdot 3^{k-a_i}p_i$, $i = 1, 2, \ldots, r$, are in decreasing order. □

This immediately implies a similar result for the largest odd numbers at a height.

Corollary 3.5. *Let $k > 2$. Let $p_i \in \widehat{P}$, $i = 1, 2, \ldots, r$, be as above, with r the largest such that $a_r \leq k$. At height k, the largest odd numbers, in decreasing order, are:*

$$3^{k-1}, 3^{k-3} \cdot 7, 3^{k-4} \cdot 19, 3^{k-6} \cdot 163, 3^{k-a_5} \cdot p_5, \ldots, 3^{k-a_r} \cdot p_r.$$

Remark. Corollary 3.5 extends Theorems 10, 11 and their corollary from Shapiro [9, §5]. Shapiro considers only two primes in $\widehat{P}$, namely 7 and 19.

To summarize our work so far, we have found in Theorem 2.3 an alternative way of thinking about the Shapiro classes. We saw above how a rather obvious observation, about the largest possible prime in a class, can be used to obtain more information about the numbers that appear at a height.

At this point, we would like to venture a comment of a philosophical nature, motivated by another innocuous observation about primes in Shapiro classes.

Theorem 2.1 suggests that $H(n)$ is a "measure of complexity" of a number. The prime numbers can be considered the "atoms" of numbers. A number is built from 1 by successive multiplication by prime numbers, so the number of prime powers dividing a number says something about how complicated a number is. However, this construction does not distinguish between two primes. On the other hand, the Shapiro class structure naturally distinguishes between the primes. On looking at Table 3, we see that primes don't come in order. For example, 19 appears at height 4 and 17 at height 5. Thus, the height function gives a "measure of complexity" to each prime, and indeed, to each number. That is why we expect this construct will say something about prime numbers.

4. Chebyshev-type theorems

In this section, we explore one strategy to discover what Shapiro classes imply for prime numbers. The strategy is to study elementary methods explained in Shapiro [10, Chapter 9] and Apostol [1, Chapter 4], and express classical results using $H(n)$ and $S(n)$. The objective of this preliminary investigation is to arrive at suitable functions that are related to the prime number functions $\pi(n)$ and p_n.

We derive results analogous to Chebyshev's theorem, which states that there are constants $0 < a < A$ such that, for $n > 1$,

$$a\frac{n}{\log n} < \pi(n) < A\frac{n}{\log n}. \tag{4.1}$$

According to Apostol [1, Theorem 4.6, (14) and (18)], we can take $a = 1/4$ and $A = 6$, when n is an even number.

We will use this result to provide an alternate formulation of Chebyshev's theorem in terms of $S(n)$. In addition, we find inequalities for p_n by modifying the proof of Apostol [1, Theorem 4.7] appropriately.

We begin with two preliminary lemmas.

Lemma 4.1. *For $n > 1$, we have*

$$\frac{S(n)}{n} \geq \frac{\log n/2}{\log 3}.$$

Proof. We use the following refinement of Stirling's formula due to Robbins [8, (1) and (2)].

$$\sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n e^{\frac{1}{12n}}. \tag{4.2}$$

From (2.3) it is immediate that

$$S(n) = \sum_{k=2}^{n} H(k) \geq \frac{1}{\log 3}\left(\log n! - (n-1)\log 2\right) + n - 1.$$

Thus, using (4.2), we obtain, for $n > 1$,

$$\begin{aligned}\frac{S(n)}{n} &\geq \frac{\log n}{\log 3} - \frac{\log 2}{\log 3} + \frac{1}{n}\cdot\Big(\frac{\log(2\pi n)}{2\log 3} + \frac{\log 2}{\log 3} - 1\Big)\\ &\qquad + \frac{1}{(\log 3)n(12n+1)} + \Big(1 - \frac{1}{\log 3}\Big)\\ &\geq \frac{\log n}{\log 3} - \frac{\log 2}{\log 3}\end{aligned}$$

as required. □

We require one more lemma.

Lemma 4.2. *Let n be such that $2^{k-1} < n \leq 2^k$. Let $\beta = \log 2/2\log 3 \approx 0.31546\ldots$. Then*

$$\beta(k-2) \leq \frac{S(n)}{n} \leq k. \tag{4.3}$$

Proof. Since $n > 2^{k-1}$, we have

$$\frac{S(n)}{n} \geq \frac{S(2^{k-1})}{n} \geq \frac{S(2^{k-1})}{2^k} \geq \frac{1}{2}\Big(\frac{\log 2^{k-2}}{\log 3}\Big),$$

where we have used Lemma 4.1. In this manner, we obtain:

$$\frac{S(n)}{n} \geq \beta(k-2),$$

where $\beta = \log 2/2\log 3$. This proves the first inequality.

Since n is such that $n \leq 2^k$, then in view of Observation (4) in §2, we must have

$$H(m) \leq k \text{ for all } m \leq n.$$

The second inequality follows immediately from this. □

Theorem 4.3 (A Chebyshev-type Theorem)**.** *For $n > 2$, there are constants a and A such that*

$$a\frac{n^2}{S(n)} \leq \pi(n) \leq A\frac{n^2}{S(n)}.$$

Proof. Let $n > 2$, and k be such that $2^{k-1} < n \leq 2^k$. The inequalities (4.1) imply that there are constants a' and A' such that:

$$a'\frac{2^{k-1}}{k-1} \leq \pi(2^{k-1}) \leq \pi(n) \leq \pi(2^k) \leq A'\frac{2^k}{k}. \tag{4.4}$$

Now using (4.3), we obtain

$$\pi(n) \geq a'\Big(\frac{k-2}{k-1}\Big)\frac{2^{k-1}}{k-2} \geq \frac{a'\beta}{4}\cdot\frac{n^2}{S(n)}$$

and,

$$\pi(n) \leq 2A'\frac{n^2}{S(n)}.$$

This completes the proof of the theorem. □

Remarks.

(1) By following the proof of (4.1) in Shapiro [10, Chapter 9] or Apostol [1, Chapter 4], we can obtain (4.1) in the special case when n is a power of 2; and from there, for all values of n.

(2) We can obtain values for a and A in the statement of Theorem 4.3 by taking $a' = 1/4$ and $A' = 6$ (or perhaps even better values, closer to 1). However, the purpose here is to find a suitable function that can be related to $\pi(n)$. The function is evidently $F(n) = n^2/S(n)$.

Theorem 4.4. *For $n > 2$, there are constants a, A_1 and A_2, such that*

$$aS(n) \leq p_n \leq A_1 S(n) + A_2 n.$$

Proof. The proof is analogous to the proof of the inequalities

$$\frac{1}{6}n\log n < p_n < 12\big(n\log n + n\log(12/e)\big),$$

given by Apostol [1, Theorem 4.7].

Let $m = p_n$, so $\pi(p_n) = n$. Let K be such that $2^{K-1} < p_n = m \leq 2^K$, and k such that $2^{k-1} < n \leq 2^k$. Clearly, $k \leq K$ (since $n < p_n$).

We begin with the first inequality. The inequalities (4.4) imply that there is a constant A' such that

$$n = \pi(p_n) \leq A'\frac{2^K}{K} \leq 2A'\frac{m}{K},$$

or

$$p_n = m \geq \frac{Kn}{2A'} \geq \frac{kn}{2A'}.$$

But $k \geq S(n)/n$ by (4.3), so we obtain

$$p_n \geq aS(n),$$

where $a = 1/2A'$.

For the second inequality, (4.4) implies that for some a',

$$p_n = m \leq \frac{2n(K-1)}{a'}.$$

Next, using $K - 1 \leq \log m/\log 2$, we obtain

$$K - 1 \leq \frac{\log m}{\log 2} \leq \frac{2\sqrt{m}}{e \log 2},$$

so

$$m \leq \frac{4n\sqrt{m}}{a' e \log 2},$$

which yields,

$$\sqrt{m} \leq \frac{4n}{a' e \log 2}.$$

Taking logs, we find that

$$\log m \leq 2 \log n + 2 \log(4/a' e \log 2) \leq 2k \log 2 + 2 \log(4/a' e \log 2).$$

Finally, we put all the above together, to find that

$$\begin{aligned} p_n (= m) &\leq \frac{2n(K-1)}{a'} \\ &\leq \frac{2n \log m}{a' \log 2} \\ &\leq \frac{4nk}{a'} + \frac{4 \log(4/a' e \log 2)}{a' \log 2} n \\ &\leq \frac{4n}{a'} \left(\frac{S(n)}{\beta n} + 1 \right) + \frac{4 \log(4/a' e \log 2)}{a' \log 2} n \\ &= A_1 S(n) + A_2 n, \end{aligned}$$

for some constants A_1 and A_2. This completes the proof. □

To summarize, we obtained two Chebyshev-type theorems, one for $\pi(n)$ and the other for p_n. Of course, the first such theorem came up in response to Gauss' conjecture which said the constants a and A in (4.1) are both 1. The question arises: How good are these functions in approximating $\pi(n)$ and p_n? These questions are considered in the next section.

5. (Conjectural) formulas for prime numbers

In this section we note some conjectural formulas that are motivated by Theorems 4.3 and 4.4 and present some computational evidence. We note here a particular constant B that appears in our study:

$$B := \frac{\gamma}{\log 2} \approx 0.832746\ldots, \tag{5.1}$$

where γ is the Euler-Mascheroni constant.

A formula for $\pi(n)$. In view of Theorem 4.3, the first question we investigate is: If $\pi(n)$ is approximately a constant multiple of $F(n) = n^2/S(n)$, then what should that constant be? However, initial experiments on Sage [12] with various numerical guesses did not match the data as n became large. However, on graphing the difference of $\pi(n)$ with $F(n)$, the error term appears to be of the same type as $F(n)$ itself. This leads to the following conjecture:

Conjecture 1. *Let $G(n)$ be defined as*

$$G(n) = \frac{n^2}{S(n)} + \frac{(bn)^2}{S(\lfloor bn \rfloor)}. \tag{5.2}$$

Then $G(n) \sim \pi(n)$, for a constant b, where b is (approximately)

$$b \approx e^B/10 \approx 0.22996\ldots.$$

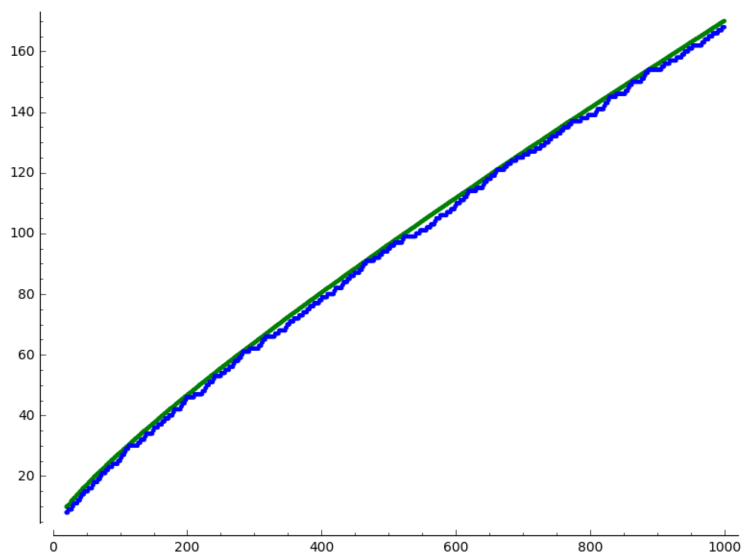

FIGURE 2. Graph of $G(n)$ (top) against $\pi(n)$

Notes.

(1) Figure 2 shows the graphs of $\pi(n)$ and $G(n)$ on the same set of axes for $n \leq 1000$. The agreement is quite striking.
(2) Figure 3 shows two curves $\pi(n)/G(n)$ (top) and $\pi(n)/\operatorname{Li}(n)$ (bottom) for $n = 20$ to $n = 90000$. It appears that $G(n)$ is a better estimate than $\operatorname{Li}(n)$ for "small" values of n.
(3) Table 4 compares the performance of $G(n)$ with that of $\operatorname{Li}(n)$ for some values of n.

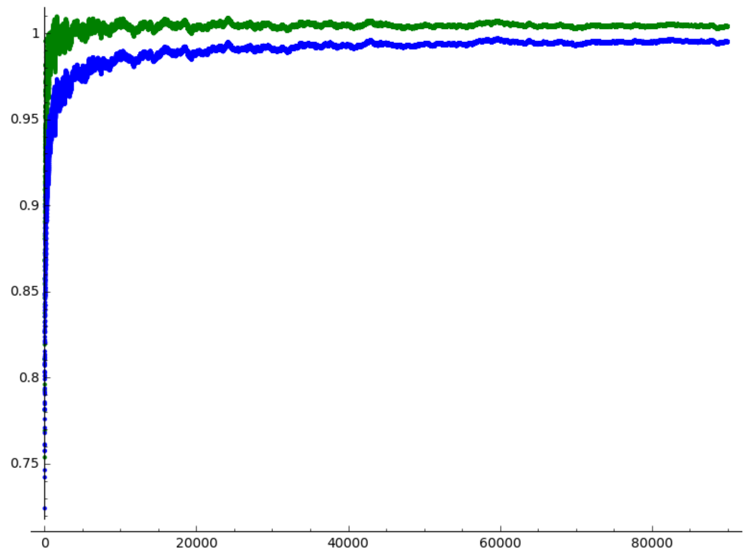

FIGURE 3. Graphs of $\pi(n)/G(n)$ (top) and $\pi(n)/\operatorname{Li}(n)$

n	$\pi(n)$	$\lfloor G(n) \rfloor$	$\pi(n) - \lfloor G(n) \rfloor$	$\lfloor \operatorname{Li}(n) \rfloor$	$\pi(n) - \lfloor \operatorname{Li}(n) \rfloor$	$\frac{\pi(n)-G(n)}{\pi(n)}$
10	4	8	−4	6	−2	−113.64%
10^2	25	27	−2	30	−5	−10.78%
10^3	168	170	−2	177	−9	−1.22%
10^4	1229	1222	7	1246	−17	0.50%
10^5	9592	9547	45	9629	−37	0.46%
10^6	78498	78340	158	78627	−129	0.20%
10^7	664579	664297	282	664918	−339	0.04%

TABLE 4. $G(n)$ against $\operatorname{Li}(n)$

A formula for the n^{th} prime. Given Theorem 4.4, one can ask how well is p_n approximated by a constant times $S(n)$. It turns out that even with the constant equal to 1, the approximation is quite good. Indeed, it appears that

$$p_n \approx S(n). \tag{5.3}$$

Here the values of n we have computed are until $n \approx 6 \times 10^7$.

Notes.

(1) See Figure 1 mentioned in the introduction for a graph of $S(n)$ and p_n, for $n \leq 5000$.
(2) Table 1 indicates that $S(n)$ is a better approximation to p_n than $n \log n$.
(3) Table 5 contains the values of $S(n)$ and p_n at some large random values of n, and the relative error of the approximation. It appears that the relative error is increasing.

n	p_n	$S(n)$	$S(n)-p_n$	$\frac{S(n)}{n\log n}$	$\frac{S(n)-p_n}{p_n}$
3874958	65619413	68671533	3052120	1.168215	4.65%
17594789	326260271	344294853	18034582	1.172923	5.53%
29742315	568063631	601049024	32985393	1.174364	5.81%
32970915	633319879	670440504	37120625	1.174637	5.86%
46262236	905219069	959827638	54608569	1.175509	6.03%
54074749	1066983163	1132214258	65231095	1.175901	6.11%

TABLE 5. Comparison of $S(n)$ with p_n: Random values

From the above, it seems that the approximation is slightly off. In view of Conjecture 1, we expect the following.

Conjecture 2. *Let p_n denote the nth prime. Then*

$$p_n \sim \frac{1}{(1+b)} S(n),$$

for a constant b, where $b \approx e^B/10 \approx 0.22996\ldots$.

Remark. Erdos et.al. [3] conjectured that there is a constant δ such that $S(n) \sim \delta n \log n$. These authors showed that a certain form of the Elliot–Halberstam conjecture implies their conjecture. Conjecture 2 implies that $\delta \approx 1 + b \approx 1.22996\ldots$. Note that $1/(1+b) \approx 0.813\ldots$.

Professor Pomerance pointed out that (5.3) is inconsistent with Conjecture 1, and commented that it would be interesting to perform numerical computations to conjecture the value of δ. From Table 5, it appears that the number is bigger than 1.175. But we were only able to compute upto $n = 6 \times 10^7$. We expect the limit to be larger.

Below, we briefly outline the steps to obtain the value of δ that follows from Conjecture 1. This motivates our statement of Conjecture 2.

Define $S(x)$ now as $\sum_{k\leq x} H(k)$. Take $G(x) = 0$ when $x < 2$. Then (5.2) can be written as

$$G(n) = \frac{n^2}{S(n)} + \frac{(bn)^2}{S(bn)}$$

On inverting this, we obtain the approximation

$$\frac{n^2}{S(n)} \approx \sum_{k\geq 0} (-1)^k G(b^k n).$$

Note that for any fixed n this is a finite sum. Upon replacing $G(x)$ by $\pi(x)$, we see that from Conjecture 1, we have

$$\frac{n^2}{S(n)} \approx \sum_{k\geq 0} (-1)^k \pi(b^k n). \tag{5.4}$$

Let $\widehat{S}(n)$ denote the approximation to $S(n)$ obtained from (5.4), that is,

$$\widehat{S}(n) := \frac{n^2}{\sum_{k\geq 0} (-1)^k \pi(b^k n)}.$$

The computations (with $b = 0.229962525551838$), in Table 6 indicate that $\widehat{S}(n)$ is quite an accurate way to estimate the value of $S(n)$. This further supports Conjecture 1. Note that we may replace $\pi(x)$ by $\mathrm{Li}(x)$ to estimate $S(n)$.

n	$S(n)$	$\widehat{S}(n)$	$S(n) - \lfloor \widehat{S}(n) \rfloor$	$\frac{S(n)-\lfloor \widehat{S}(n) \rfloor}{S(n)}$
992	7569	7628	−59	−0.77949%
7524	76008	76089	−81	−0.10656%
56762	713411	710300	3111	0.43607%
596319	9206082	9181418	24664	0.2679%
17594789	344294853	344263181	31672	0.009199%
32970915	670440504	670698724	−258220	−0.03851%
54074749	1132214258	1133070822	−856564	−0.07565%

TABLE 6. Comparison of $S(n)$ with $\widehat{S}(n)$: Random values

Multiplying both sides of (5.4) by $\log(n)/n$, and taking (formal) limits, one can obtain

$$\lim_{n\to\infty} \frac{n \log n}{S(n)} \approx \frac{1}{1+b} \approx 0.813\ldots,$$

which suggests that $\delta \approx 1 + b \approx 1.22996\ldots$.

From the above we see that Conjecture 2 is consistent with Conjecture 1.

Prime gaps. We end this section with a few remarks about the prime gap. Let g_n denote the prime gap, that is, $g_n = p_{n+1} - p_n$. Given that $S(n)$ approximates p_n, it is natural to ask whether g_n is approximated by $H(n+1)$. On the other hand, the prime gap is notorious for its irregularity, and one cannot expect much in this regard. Nevertheless, it seems that on average, $H(n+1)$ does quite a good job of approximating g_n. Indeed, Table 7 gives a few values of the following function:

$$S_\Delta(k) = \frac{1}{2^k} \sum_{2^k < m \leq 2^{k+1}} \frac{g_m}{H(m+1)}. \tag{5.5}$$

In view of Conjecture 2, we expect the limit to be $1/(1+b) \approx 0.813\ldots$.

k	$S_\Delta(k)$	k	$S_\Delta(k)$
1	1.16666666666667	10	1.02073447927940
2	1.08333333333333	11	1.01134933317550
3	1.23333333333333	12	1.00132388905581
4	1.11041666666667	13	0.994501291351865
5	1.08541666666667	14	0.988054230925281
6	1.08377976190476	15	0.982638771673053
7	1.03947792658730	16	0.976617044103504
8	1.05543154761905	17	0.971171169038482
9	1.04229290674603	18	0.966205359427567

TABLE 7. $S_\Delta(k)$ for $k = 1, \ldots, 18$

Acknowledgements. We thank Professor Carl Pomerance (Dartmouth) and Manjil P. Saikia (Vienna) for useful comments. All numerical computations and graphics have been generated using Sage [12].

References

[1] T. M. Apostol. *Introduction to analytic number theory.* Springer-Verlag, New York-Heidelberg, 1976. Undergraduate Texts in Mathematics.

[2] P. A. Catlin. Concerning the iterated Φ function. *Amer. Math. Monthly*, 77:60–61, 1970.

[3] P. Erdős, A. Granville, C. Pomerance, and C. Spiro. On the normal behavior of the iterates of some arithmetic functions. In *Analytic number theory (Allerton Park, IL, 1989)*, volume 85 of *Progr. Math.*, pages 165–204. Birkhäuser Boston, Boston, MA, 1990.

[4] W. H. Mills. Iteration of the ϕ function. *Amer. Math. Monthly*, 50:547–549, 1943.

[5] T. D. Noe. Primes in classes of the iterated totient function. *J. Integer Seq.*, 11(1):Article 08.1.2, 10, 2008.

[6] J. C. Parnami. On iterates of Euler's ϕ-function. *Amer. Math. Monthly*, 74:967–968, 1967.

[7] S. S. Pillai. On a function connected with $\phi(n)$. *Bull. Amer. Math. Soc.*, 35(6):837–841, 1929.

[8] H. Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62:26–29, 1955.

[9] H. Shapiro. An arithmetic function arising from the ϕ function. *Amer. Math. Monthly*, 50:18–30, 1943.

[10] H. N. Shapiro. *Introduction to the theory of numbers.* Pure and Applied Mathematics. John Wiley & Sons, Inc., New York, 1983. A Wiley-Interscience Publication.

[11] N. J. A. Sloane. The on-line encyclopedia of integer sequences. `https://oeis.org`, 2003.

[12] W. Stein et al. *Sage Mathematics Software (Version 8.3).* The Sage Development Team, 2018. `http://www.sagemath.org`.

The Caravan, Jhandewalan Extn., New Delhi 110001, India
E-mail address: `hartoshbal@gmail.com`

Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria.
E-mail address: `bhatnagarg@gmail.com`