THE PRIME NUMBERS WITHOUT THE SIEVE OF ERATOSTHENES

FILIP SAIDAK

Abstract. The question of enumerating the sequence of the prime numbers is one of the oldest and most fruitful in all of mathematics. In this paper, we propose a simple new recursive technique (essentially different from all sieve methods, including the original one due to Eratosthenes) for explicitly listing the primes; and it is based only on a new general density hypothesis concerning their overall distribution. The currently conditional status of this iterative device is offset by at least one major advantage: formulation of a self-perpetuating principle unhindered by fixed magnitudes and bounds that limit all sieves.

1. Introduction

In Book IX of *Elements* [10] of Euclid, written sometime after 300 B.C., one finds the oldest extant proof of the infinitude of the prime numbers, now known as Euclid's Theorem (a fundamental result that has since been proved in many other ways, see [11], [22], [23], etc.) Less than a century later, also in Alexandria, Eratosthenes was the first to invent an ingenious sieve method for the evaluation of all prime numbers below a given magnitude x, a method that still carries his name (see [13], [12] or [17] for more details). For the longest time, the sieve of Eratosthenes provided the only way of computing complete lists of primes $\leq x$.

Following Legendre's idea of quantifying the sieve of Eratosthenes, i.e.

$$\pi(x) := \sum_{p\leq x} 1 = \sum_{d} \mu(d) \left[\frac{x}{d}\right] + \pi(\sqrt{x}) - 1,$$

(where the sum is extended over all positive integers d divisible by primes $\leq x$ only, and $\mu(d)$ denotes the versatile Möbius function: $\mu(1) = 1, \mu(d) = (-1)^m$ when d is the product of m distinct primes, and $\mu(d) = 0$ otherwise) and then Brun's revolutionary refinement of it (cf. [4], [3] and [2]), the sieve theory has developed dramatically during the 20th century, significantly expanding the scope of its applicability; it has helped establish many deep results concerning the prime gaps and differences, even though, in spite of all the efforts, certain questions addressing the finer aspects of the distribution of prime numbers – the existence of infinitely many *twin primes* being the most notorious example – still remain unsolved to this day.

In this short paper, we look at things from a different angle, and propose a new systematic method for listing the prime numbers. Not only does our prime number device not refer to a sieve, it is something of a reverse of it; defined recursively, it works forward by looking back, and is not limited by any fixed upper bounds on intervals of interest, its underlining principle possessing a useful self-perpetuating property: it could be started up once and would, henceforth, compute the list of consecutive primes forever. Moreover, it employs only the simplest congruence properties of the prime gaps, the sequence of which we define as:

$$G = \{2, 4, 2, 4, 2, 4, 6, \ldots\}, \quad \text{where} \quad g_n = p_{n+1} - p_n, \quad \text{for} \quad p_n \geq 5.$$

For convenience, we denote by G_n the truncation of G after the gap g_n.

2. Properties of Prime Gaps

One of the oldest unsolved problems in number theory remains the question of existence of infinitely many pairs of prime numbers p and $p+2$, that differ by 2 – e.g. $(3,5)$ and $(11,13)$ and $(2027,2029)$ – the so-called *twin primes*, mentioned above. In 1859, de Polignac [19] realized that there was nothing special about the prime gap having size 2 and he explicitly conjectured that every *even* integer $k \geq 2$ will appear infinitely often in G; in other words, if we define the counting function $g(x,k) := \#\{p_n \leq x : \ p_{n+1} - p_n = k\}$, then $g(x,k) \to \infty$, as $x \to \infty$. This important conjecture is still open for all fixed even integers k, despite that in 1919 Brun [3] proved that $g(x,2) \ll x/(\log x)^2$, for all x, and in 1923 Hardy and Littlewood [14] supplied convincing heuristics (their Conjecture B, on p. 42 of [14]) why we should expect

$$g(x,k) \ \sim \ C_k \frac{x}{(\log x)^2}, \tag{1}$$

where

$$C_k \ := \ 2 \prod_{p=3}^{\infty} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p|k \\ p>2}} \left(1 + \frac{1}{p-2}\right)$$

are well-defined constants; in particular, $C_2 = C_4 = C_8 = 1.32032\ldots, C_6 = 2.64064\ldots$, etc.

Studying more carefully the gap sequence G, one could ask an even more general question about frequencies of occurrences of various prime gap groupings, and quickly discover that, for example, the sequences $\{2,2\}$ and $\{4,4\}$ never occur in G, while $\{2,4\}$ (and $\{4,2\}$) are popular (and are the prototypical *prime triplets*). The reason for the impossibility of the former two cases is easily found to be due to a trivial congruence restriction: one of the integers $p, p+2, p+2+2$ (resp. $p, p+4, p+4+4$) must be divisible by 3, and thus cannot be a prime. In general, a sequence (or a constellation) of gaps is called *admissible*, if it does not in similar fashion cover the full residue system modulo any prime number. For instance, the grouping $\{6,6\}$ is admissible, but $\{6,6,6,6\}$ is not, since it fails modulo 5.

Note 1: A basic, but useful, observation is that if p is the smallest prime *not* dividing a given gap d, then $\{d, d, \cdots d\}$ will become *inadmissible* when the number of ds reaches $p-1$.

Many intriguing questions addressing the density and structure of admissibile prime constellations remain unsolved [9]. It is believed, but has never been proved, that if S is admissible, then it will occur infinitely often as a subsequence of the prime gap sequence G; i.e. $\pi(x,S) \to \infty$, as $x \to \infty$. This is known as the k-tuplet conjecture of Hardy and Littlewood [14, Theorem X, p. 61], and also follows from the polynomial Hypothesis H of Schinzel [25], the quantitative version of which was proposed by Bateman and Horn [1] in 1962. We have

$$\pi(x,S) \ \sim \ C_S \int_2^x \frac{dt}{(\log t)^{|S|}} \ \sim \ C_S \frac{x}{(\log x)^{|S|}}, \tag{2}$$

where (for each finite admissible sequence S) C_S are well-defined, explicitly given constants.

Concerning the extreme size of gaps between primes, in 1936 Cramér [5] proved, under the assumption of the Riemann Hypothesis, that we have

$$G(x) := \max_{p_n \leq x} (p_{n+1} - p_n) \ \ll \ \sqrt{x} \log x.$$

But this is far from expectation, and Cramér himself conjectured that, for all x,

$$G(x) \ll (\log x)^2. \tag{3}$$

3. Two New Hypotheses

Our main objective is to discuss a new prime-computing device. Instead of employing a sieving method, it will apply the following principle of maximal density of prime distribution.

Hypothesis 1. For all $n \geq n_0$, the prime gap sequence G can be constructed iteratively, where the nth step of the extension consists of appending to G_n the smallest even positive integer that does not violate any congruence conditions of admissibility.

The idea is simple, but it has some surprisingly far-reaching consequences. To illustrate the mechanism behind the construction, let us consider the sequence of gaps up to the prime 31: $G_{10} = \{2, 4, 2, 4, 2, 4, 6, 2\}$. To extend the sequence, we place 2 as the hypothetical next gap, and we try to see if it creates a disallowed full residue system modulo any prime. It does, immediately: $\{2, 2\}$ is an inadmissible pair $\pmod 3$. So 4 becomes the next gap candidate. But, looking backwards, 4 also creates an inadmissible subsequence $\{4 + 2, 6, 4 + 2, 4 + 2\} = \{6, 6, 6, 6\} \pmod 5$. Therefore, the gap must be at least as large as 6. In the case of 6, no congruence violation occurs, so 6 is placed as the next gap, extending the sequence to: $G_{11} = \{2, 4, 2, 4, 2, 4, 6, 2, 6\}$. This process is then repeated automatically: again, 2 is the first candidate for the extension, but it yields the inadmissible $\{2+6, 2+6\} = \{8, 8\} \pmod 3$. The next option, the gap 4, can be checked promptly not to violate any congruence restrictions, and is placed as the next term of the gap sequence, yielding: $G_{12} = \{2, 4, 2, 4, 2, 4, 6, 2, 6, 4\}$.

Note 2: The two steps just outlined added the primes 37 and 41 to the list ending with 31.

Plausibility of the recursive principle outlined in Hypothesis 1 will become clearer after it is shown to follow from the following Goldbach-type conjecture:

Hypothesis 2. For every *composite* number $n \geq n_0$, there exists an even $d \geq 2$, such that the set $\{n - d, n - 2d, \ldots, n - kd\}$ contains only prime numbers, where $k = \min_{p \nmid d} p$ - 1.

Theorem 3. Hypothesis 2 implies Hypothesis 1.

Proof. The argument is simple. As noted above, the sequence $\{d, d, \cdots, d\}$ becomes inadmissible once the set of integers $p, p + d, p + 2d, p + 3d, \cdots$ fills all congruence classes modulo some prime, and that happens exactly when the number of ds reaches the smallest prime that does not divide $d-1$. So, if Hypothesis 2 is true, then every sufficiently large composite number will be (through location of the specific d) discovered to violate the conditions of admissibility, by just following an arithmetic progression of prime numbers of maximal allowed length modulo the smallest non-divisor of d. This immediately implies that the extension principle of Hypothesis 1 can yield *only* gaps between prime numbers, and since it starts the search from the smallest possible candidates, it guarantees that p_{n+1} will always follow p_n. □

4. Heuristics

We have shown that Hypothesis 1 will be true if Hypothesis 2 is true. Hypothesis 2 is likely to work (by providing "witnesses" for rejection of all composite candidates), unless non-existence of arithmetic progressions of certain required lengths k prevents them from doing so. Unfortunately, the existence of such arithmetic progressions cannot be guaranteed (see Note 3 below); however, the two classical hypotheses (2) and (3), stated in the Introduction, provide a rationale for believing this to be true. First of all, Hypothesis 2 addresses admissible sequences of length k (in fact, only the special case of arithmetic progressions of length k), terminating at the composite integer n, so the k-tuplet conjecture and the estimates (2) should be able to tell us something about the distribution of such sequences. In our case, $k = |S|$ is not a fixed constant, which complicates matters slightly. However, the Cramér Hypothesis (3) ensures that k stays relatively small; since d is roughly of the same order as a gap between primes $\leq n$ (and thus bounded above by $G(n) \ll (\log n)^2$), we necessarily have

$$k \;<\; \min_{p \nmid d} p \;\ll\; \log d \;\ll\; \log\log n. \tag{4}$$

Therefore, it seems reasonable to conclude that, by extending the heuristics behind the general conjecture (2), for all "sufficiently large" numbers n, one should always expect existence of (at least some) arithmetic progressions of this length, since as $x \to \infty$ we clearly have:

$$\frac{x}{(\log x)^{|S|}} \;\gg\; \frac{x}{(\log x)^{c \log\log x}} \;=\; \exp(\log x - c(\log\log x)^2) \;\to\; \infty.$$

Note 3: A few small exceptions to the general rule described in Hypothesis 2 exist (e.g. the number 25 will show up as "prime," because of the unusual arithmetic progression $19, 13, 7, 1$), but these special cases can be enumerated. The choice of $n_0 = 200$ appears safe.

5. Final Remarks

I. The prime numbers are believed to be the densest set of pairwise-coprime positive integers. But, how does one define the meaning of "densest"? Erdős [8] formulated the idea as follows:

$$\sum_{a_n \leq x} \frac{1}{a_n} \;\leq\; \sum_{p \leq x} \frac{1}{p}, \tag{5}$$

where $\{a_n\}$ is any sequence of pairwise-coprime integers ≥ 2. However, certain complications arise when one tries to transfer (or localize) this property, as Hensley and Richards have shown in [15]: if the prime k-tuplet conjecture (from [14]) is true, then infinitely often we must have

$$\pi(x + y) > \pi(x) + \pi(y), \tag{6}$$

disproving a speculation of Hardy and Littlewood (also in [14], pp. 52–54), and implying that the initial subsequences of the prime sequence cannot be the densest possible. Therefore, it is somewhat surprising that it is actually possible to guarantee a maximal density of the prime numbers in an iterative manner, as we have outlined above.

II. The algorithm behind the original sieve of Eratosthenes finds all prime numbers below x after performing $O(x \log\log x)$ operations, but while doing so it requires $O(x)$ bits of storage. In computational applications, one always encounters a similar trade-off between the time and the space. Removing various repetitions and redundancies in the sieving process via the use of clever segmentation and modern primality tests can help one improve the overall results with respect to both the speed and the size. Today, the most time-efficient algorithms require only $O(x/\log\log x)$ operations, but need relatively large storage; and the most space-efficient algorithms use just $O(\sqrt{x})$ bits of storage, but are considerably slower (see [6], [20], and [26]).

Our device seems to compare favorably with the sieves. The mechanism of extension has $\pi(x)$ steps (up to x), and for each step it runs through the gap candidates (of average size $\log x$) and tests arithmetic progressions defined in Hypothesis 2. By equation (4), the length of these progressions is at most $O(\log\log x)$, so the time-complexity of the process is expected to be $O(x \log\log x)$. As far as space is concerned, the bulk is taken up by storing portions (the tail-ends) of the sequence of prime gaps required for testing the congruence conditions; in the worst case this would amount to $O(x/\log x)$, but, in reality, only very small portions are actually used, and the bound $O(\sqrt{x})$ appears much closer to the truth. Unfortunately, to say something more precise, one would need accurate information about 1) the extreme size (as a function of x) of the smallest d for which a candidate for the next prime gap is rejected, and 2) the average size of the corresponding progression lengths k. However, both of these depend on a better understanding of the *variable* extensions of (2), unavailable at present.

References

[1] P. Bateman and R. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp., **16** (1962), 1962, 363–367.

[2] V. Brun, *La série 1/5 + 1/7 + 1/11 + 1/13 + 1/17 + 1/19 + 1/23 + · · ·*, Bull. Sci. Math., **43** (1919), 100–104 and 124–128.

[3] V. Brun, *Le crible d'Eratosthène et le théorème de Goldbach*, C. R. Acad. Sci. Paris, **168** (1919), 544–546.

[4] V. Brun, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Archiv Math. Natur., **B 34** (8), 1915.

[5] H. Cramér, *On the order of magnitude of the difference between consecutive prime numbers*, Acta Arith., **2** (1936), 23–46.

[6] R. Crandall and C. Pomerance, *Prime Numbers: A Computational Perspective*, Second Edition, Chapter 3, Springer, New York, 2005.

[7] L. E. Dickson, *History of the Theory of Numbers*, Second Edition, Chelsea Publishing, New York, 1992.

[8] P. Erdős, *Seminar at the University of Limoges*, 1988.

[9] P. Erdős and H. Riesel, *On admissible constellations of consecutive primes*, BIT **28** (1988), 391–396.

[10] Euclid - *Elements*, Alexandria, c. 300 B. C. (for a modern edition, see T. L. Heath, *Thirteen Books of Euclid's Elements*, Dover, New York, 1956).

[11] L. Euler, *Variae observationes circa series infinitas* (1737), Commentarii Acad. Sci. Petropolitanae, **9** (1744), 160–188.

[12] J. Friedlander and H. Iwaniec, *Opera de Cribro*, AMS Colloquium Publications, No. 57, Providence, 2010.

[13] H. Halberstam and H. E. Richert, *Sieve Methods*, LMS Monographs, Vol. 4, Academic Press, London, 1974.

[14] G. H. Hardy and J. E. Littlewood, *Some problems of partitio numerorum (III): On the expression of a number as a sum of primes*, Acta Math., **44** (1923), 1–70.

[15] D. Hensley and I. Richards, *On the incompatibility of two conjectures concerning primes*, Proc. Sympos. Pure Math., Vol. XXIV, Saint Louis Univ., Saint Louis, Mo., 1972.

[16] D. Hensley and I. Richards, *Primes in intervals*, Acta Arith., **25** (1973/74), 375–391.

[17] H. Iwaniec, *Sieve Methods*, Lecture Notes, Rutgers University, 1996.

[18] W. Narkiewicz, *The Development of Prime Number Theory. From Euclid to Hardy and Littlewood*, Springer-Verlag, Berlin, 2000.

[19] A. de Polignac, *Recherches nouvelles sur les nombres premiers*, C. R. Acad. Sci. Paris, **29** (1859), 397–401.

[20] P. Pritchard, *Linear prime-number sieves: a family tree*, Sci. Comput. Programming, **9.1** (1987), 17–35.

[21] I. Richards, *On the incompatibility of two conjectures concerning primes*, Bull. Amer. Math. Soc., **80** (1974), 419-438.

[22] F. Saidak, *A new proof of Euclid's Theorem*, American Math. Monthly, **113.10** (2006), 937–938.

[23] F. Saidak, *A note on Euclid's Theorem concerning the infinitude of the primes*, Acta Univ. M. Belii, **24** (2016), 25–26.

[24] A. Schinzel, *Remarks on the paper "Sur certaines hypothéses concernant les nombres premiers"*, Acta Arith., **7** (1961), 1–8.

[25] A. Schinzel and W. Sierpiński, *Sur certaines hypothéses concernant les nombres premiers*, Acta Arith., **4** (1958), 185–208.

[26] J. Sorenson, *An Introduction to Prime Number Sieves*, Computer Sciences Technical Report #909, University of Wisconsin-Madison, 1990.

MSC2010: 11A41, 11A05, 11N35

Department of Mathematics, University of North Carolina, Greensboro, NC 27402, U.S.A.
E-mail address: saidak@protonmail.ch