A construction of A. Schinzel– many numbers in a short interval without small prime factors

Sergei Konyagin

November 29, 2022, Moscow

Introduction

For $x \geqslant 0$. let $\pi(x)$ denote the number of primes $p \leqslant x$. For example, $\pi(1) = 0$, $\pi(2.5) = 1$,$\pi(5) = 3$. The prime number theorem describes the asymptotic distribution of primes. It is known (J. Hadamard, Sh. de la Vàlléé-Poussin, 1896) that

$$\pi(x) \sim x/\log x \quad (x \to \infty).$$

The sharpest known version of the prime number theorem (I.M. Vinogradov, N.M. Korobov, 1958) says that for $x \geqslant 3$

$$\pi(x) - \int_2^x \frac{dt}{\log t} = O\left(x \exp\left(-c(\log x)^{3/5}(\log\log x)^{-1/5}\right)\right).$$

One can conclude the asymptotic expansion for $\pi(x)$. In particular,

$$\pi(x) = x/\log x + x/\log^2 x + o(x/\log^2 x) \quad (x \to \infty).$$

We will discuss the local behavior of $\pi(x)$.

A general question is: how can we estimate $\pi(x+y)-\pi(y)$ in terms of x? We assume that $x \geqslant 2$ and $y \geqslant 2$ are integers. We consider several simple examples.

1. $x = 2$. Then $\pi(y+2)-\pi(y)$ is the number of primes less than or equal to $y+2$ but greater than y, i.e. the number of primes in the set $\{y+1, y+2\}$. One of these eqo numbers is even and the other is odd. Only the odd number can be a prime. Hence,
$\pi(y+2)-\pi(y) \leqslant 1$.

2. $x = 3$ or $x = 4$. Then $\pi(y+4)-\pi(y)$ is the number of primes in the set $\{y+1, y+2, y+3, y+4\}$. Two of these number are even and two other numbers are odd. Therefore,
$\pi(y+3)-\pi(y) \leqslant \pi(y+4)-\pi(y) \leqslant 2$.

Observe that for $x = 2, 3, 4$ and any $y \geqslant 2$ we have

$$\pi(x+y)-\pi(y) \leqslant \pi(x).$$

The conjecture of Hardy and Littlewood on subbaditivity

Based on extensive numerical experiments and heuristical arguments, G. Hardy and E. Littlewood (1923) conjectured the following.

Conjecture 1

For any integers $x, y \geqslant 2$ the inequality

$$\pi(x+y) \leqslant \pi(x) + \pi(y). \tag{1}$$

holds.

Observe that due to the symmetry we can assume WLOG that $y \geqslant x$. We have seen that the conjecture holds for $x = 2, 3, 4$. Now it has been verified for $\min(x, y) \leqslant 1371$ (D.M. Gordon and G. Rodemich, 1998).

On the other hand, its validity is known if x is not much less than y. In particular, take $x = y$. The equality

$$\pi(z) = z/\log z + z/\log^2 z + o(z/\log^2 z) \quad (z \to \infty)$$

implies that

$$\pi(2x) \leqslant 2\pi(x) - (2\log 2 + o(1))x(\log x)^{-2} \quad (x \to \infty).$$

On the other hand, its validity is known if x is not much less than y. In particular, take $x = y$. The equality

$$\pi(z) = z/\log z + z/\log^2 z + o(z/\log^2 z) \quad (z \to \infty)$$

implies that

$$\pi(2x) \leqslant 2\pi(x) - (2\log 2 + o(1))x(\log x)^{-2} \quad (x \to \infty).$$

Moreover, using explicit estimates for $\pi(x)$, one can check the inequality $\pi(2x) \leqslant 2\pi(x)$ for all integers $x \geqslant 2$.
It is known that there is $c < 1$ such that $\pi(x + y) \leqslant \pi(x) + \pi(y)$ holds for any integers x, y with $x \geqslant 2$, $y^c \leqslant x \leqslant y$.
However, it is likely to expect that in general the conjecture is false.

Prime k-tuple conjecture

We know that $\pi(x+y)-\pi(y)$ is the number of primes in the set $\{y+1,\dots,y+x\}$. Now we will discuss a related question. Let $\{b_1<\dots<b_k\}$ be a sequence of integers. Are there infinitely many such y that all numbers $y+b_1,\dots,y+b_k$ are primes? For some sets $\{b_1<\dots<b_k\}$ this is impossible due to local obstacles: one of the numbers $y+b_1,\dots,y+b_k$ must be divisible by a small prime. For example, take $k=2, b_1=1, b_2=2$. Then one of the numbers $y+1, y+2$ is even and both numbers are primes only for $y=1$. In general, assume that for some prime p any congruent class $\bmod p$ contains a number from the set $\{b_1,\dots,b_k\}$. Then one of the numbers $y+b_1,\dots,y+b_k$ must be divisible by p. Therefore, if all the numbers are primes, then one of them is equal to p.

Let us call a set $\{b_1, \dots, b_k\}$ of integers admissible if for each prime p there is some congruence class $\bmod p$ which contains none of the integers b_i. We see that there are infinitely many such y that all numbers $y + b_1, \dots, y + b_k$ are primes only if the set $\{b_1, \dots, b_k\}$ is admissible. G. Hardy and E. Littlewood conjectured that the condition of admissibility is sufficient as well.

Let us call a set $\{b_1, \dots, b_k\}$ of integers admissible if for each prime p there is some congruence class $\bmod p$ which contains none of the integers b_i. We see that there are infinitely many such y that all numbers $y + b_1, \dots, y + b_k$ are primes only if the set $\{b_1, \dots, b_k\}$ is admissible. G. Hardy and E. Littlewood conjectured that the condition of admissibility is sufficient as well.

Conjecture 2

(Prime k-tuple conjecture.) Let $b_1 < \dots < b_k$ be any admissible sequence. Then there exist infinitely many integers n for which all the numbers $n + b_1, \dots, n + b_k$ are primes.

(Strictly speaking, the prime k-tuple conjecture is a more general hypothesis, but for our purpose we need the particular case we have formulated.) It is widely believed that Conjecture 2 is true.
In spite of the spectacular recent progress in the investigation of this conjecture, we still do not know any set $\{b_1, \dots, b_k\}$, $k \geqslant 2$, for which it is known to be true.

By $\rho^*(x)$ we denote the maximal cardinality of an admissible subset of the set $\{1, \dots, x\}$.
We observe that if $1 \leqslant b_1 < \dots < b_k \leqslant x \leqslant y$ and the numbers $y + b_1, \dots, y + b_k$ are primes then the set $\{b_1, \dots, b_k\}$ is admissible. Indeed, if $p \leqslant x$ is a prime then $(y + b_i, p) = 1$ for $i = 1, \dots, k$. Hence, $\{b_1, \dots, b_k\}$ misses the residue class $-y$ modulo p. If $p > x$ then the number p of the residue classes modulo p is greater than k and again there are residue classes free of numbers $b_1, \dots, b_k$.

By $\rho^*(x)$ we denote the maximal cardinality of an admissible subset of the set $\{1, \dots, x\}$.
We observe that if $1 \leqslant b_1 < \dots < b_k \leqslant x \leqslant y$ and the numbers $y + b_1, \dots, y + b_k$ are primes then the set $\{b_1, \dots, b_k\}$ is admissible. Indeed, if $p \leqslant x$ is a prime then $(y + b_i, p) = 1$ for $i = 1, \dots, k$. Hence, $\{b_1, \dots, b_k\}$ misses the residue class $-y$ modulo p. If $p > x$ then the number p of the residue classes modulo p is greater than k and again there are residue classes free of numbers $b_1, \dots, b_k$. Hence,

$$\pi(x + y) - \pi(y) \leqslant \rho^*(x).$$

Moreover, $\max_{y \geqslant x}(\pi(x + y) - \pi(y)) = \rho^*(x)$ if Conjecture 2 is true. Conjecture 1 would follow from the inequality $\rho^*(x) \leqslant \pi(x)$ for $x \geqslant 2$ and is equivalent to this inequality if Conjecture 2 holds.
In principal the quantity $\rho^*(x)$ can be evaluated for any positive integer x. It has been verified that $\rho^*(x) \leqslant \pi(x)$ for $2 \leqslant x \leqslant 1371$ implying Conjecture 1 for $2 \leqslant \min(x, y) \leqslant 1371$.

The inequality of Hensley and Richards

D. Hensley and I. Richards (1974) showed that Conjecture 1 is not compatible with the prime k-tuple conjecture. Namely,

$$\rho^*(x) - \pi(x) \geqslant (\log 2 - o(1))x(\log x)^{-2} \quad (x \to \infty). \tag{2}$$

In particular, for large x

$$\rho^*(x) > \pi(x). \tag{3}$$

It is known from D.A. Clark and N.C. Jarvic (2003) that (3) holds for $x = 4916$: $\rho^*(4916) \geqslant 657 > 656 = \pi(4916)$.

The inequality of Hensley and Richards

D. Hensley and I. Richards (1974) showed that Conjecture 1 is not compatible with the prime k-tuple conjecture. Namely,

$$\rho^*(x) - \pi(x) \geqslant (\log 2 - o(1))x(\log x)^{-2} \quad (x \to \infty). \tag{2}$$

In particular, for large x

$$\rho^*(x) > \pi(x). \tag{3}$$

It is known from D.A. Clark and N.C. Jarvic (2003) that (3) holds for $x = 4916$: $\rho^*(4916) \geqslant 657 > 656 = \pi(4916)$.
Therefore, assuming the prime k-tuple conjecture we get

$$\max_{y \geqslant x}(\pi(x+y) - \pi(x) - \pi(y)) \geqslant (\log 2 - o(1))x(\log x)^{-2} \quad (x \to \infty)$$

and, in particular, there is y such that $\pi(4916+y) > \pi(4916) + \pi(y)$.

If Conjecture 2 holds, then there is a theoretical possibility to find y satisfying the last inequality and to disprove Conjecture 1. However, probably it is an absolutely untractable project. The proportion of such numbers y is too small. We do not know a good way for the search. One needs to try a huge number of candidates to have a reasonable chance to catch y. It is likely that the expected number of trials have thousands of digits.

If Conjecture 2 holds, then there is a theoretical possibility to find y satisfying the last inequality and to disprove Conjecture 1. However, probably it is an absolutely untractable project. The proportion of such numbers y is too small. We do not know a good way for the search. One needs to try a huge number of candidates to have a reasonable chance to catch y. It is likely that the expected number of trials have thousands of digits.
It is naturally to expect that

$$\rho^*(x) \leqslant (1 + o(1))\pi(x) \quad (x \to \infty).$$

implying a weak form of Conjecture 1

$$\pi(x + y) \leqslant (1 + o(1))\pi(x) + \pi(y) \quad (x, y \to \infty).$$

Now we know from H.L. Montgomery and R.C. Vaughan (1973) that

$$\rho^*(x) \leqslant 2\pi(x) \quad (x \geqslant 2). \tag{4}$$

Even a small improvement of the constant 2 in (4) would be a great result.

The construction of Hensley and Richards

We will assume that x is a sufficiently large integer. Since the property of admissibility is invariant under translations, one can construct it as a subset of any set of x consecutive integers. WLOG we can consider that x is an odd number. Hensley and Richards construct an admissible set $\mathcal{B} \subset [(1-x)/2, (x-1)/2]$ as

$$\mathcal{B} = \{\pm p : \ y < p \leqslant x/2\},$$

where p runs over primes and y is an appropriate number, $0 < y < x/2$. We have

$$|\mathcal{B}| = 2\pi(x/2) - 2\pi(y).$$

To prove the required inequality, we have to show that for some y the set $\mathcal{B}$ is admissible and

$$|\mathcal{B}| - \pi(x) = (\log 2 + o(1))x(\log x)^{-2}. \qquad (5)$$

Since for $x \to \infty$

$$\pi(x) = x(\log x)^{-1} + (1 + o(1))x(\log x)^{-2},$$

$$\pi(x/2) = \frac{x/2}{\log x - \log 2} + (1 + o(1))(x/2)(\log(x/2))^{-2},$$

we can deduce that

$$2\pi(x/2) - \pi(x) = (\log 2 + o(1))x(\log x)^{-2}.$$

If $y = o(x/\log x)$, then $\pi(y) = o(x/\log^2 x)$ and the required inequality (5) holds. We take

$$y = x(\log x)^{-1}(\log\log x)^{-1/2}.$$

Since for $x \to \infty$

$$\pi(x) = x(\log x)^{-1} + (1 + o(1))x(\log x)^{-2},$$

$$\pi(x/2) = \frac{x/2}{\log x - \log 2} + (1 + o(1))(x/2)(\log(x/2))^{-2},$$

we can deduce that

$$2\pi(x/2) - \pi(x) = (\log 2 + o(1))x(\log x)^{-2}.$$

If $y = o(x/\log x)$, then $\pi(y) = o(x/\log^2 x)$ and the required inequality (5) holds. We take

$$y = x(\log x)^{-1}(\log\log x)^{-1/2}.$$

If $p \leqslant y$ then the congruence class $0 (\mathrm{mod}\, p)$ contains none of the elements of $\mathcal{B}$. Moreover, clearly that for $p > |\mathcal{B}|$ some congruence class $\mathrm{mod}\, p$ which contains none of the elements of $\mathcal{B}$. We have to examine the primes

$$y < p \leqslant |\mathcal{B}|. \tag{6}$$

The method of Erdős and Rankin

For any prime p satisfying (6) we want to find an integer $z \in (-x/2 - p, -x/2)$ such that no element of $\mathcal{B}$ is $z (\mathrm{mod}\, p)$. Any number $u \in (-x/2, x/2)$ congruent to z modulo p can be written as $z + jp$ where $jp < p + x$. Hence,

$$j \leqslant 1 + x/p \leqslant 1 + x/y,$$

Denote $Y = 1 + [x/y]$. Thus,

$$1 \leqslant j \leqslant Y.$$

No number $z + jp$ should be equal to $\pm p'$ where $p' > y$ is a prime. It is sufficient to prove that any number $z + jp$, $1 \leqslant j \leqslant Y$, is divisible by some prime $q \leqslant X$ where $X = [(\log x)/2]$.

By the classical result of R.A. Rankin (1938) improving earlier results by E. Westzynthius (1931) and P. Erdős (1935) for large positive integers X and

$$Y \leqslant cX(\log X)(\log\log X)^{-2}(\log\log\log X) \tag{7}$$

there is $u \in \mathbb{Z}$ such that each number $u+1, \dots, u+Y$ has a prime factor $\leqslant X$. Now we know that the bound can be multiplied by $\log\log X$ (K. Ford, B. Green, SK, J. Maynard, T. Tao, 2018), but this improvement is not essential for our purposes.

By the classical result of R.A. Rankin (1938) improving earlier results by E. Westzynthius (1931) and P. Erdős (1935) for large positive integers X and

$$Y \leqslant cX(\log X)(\log\log X)^{-2}(\log\log\log X) \tag{7}$$

there is $u \in \mathbb{Z}$ such that each number $u+1, \dots, u+Y$ has a prime factor $\leqslant X$. Now we know that the bound can be multiplied by $\log\log X$ (K. Ford, B. Green, SK, J. Maynard, T. Tao, 2018), but this improvement is not essential for our purposes. By our choice,

$$X = [\log x/2],\ Y = [x/y]+1 = \log x(\log\log x)^{1/2}+1 \leqslant 3X(\log X)^{1/2},$$

and (7) holds. Thus, there is $u \in \mathbb{Z}$ such that each number $u+1, \dots, u+Y$ has a prime factor $\leqslant X$.

Therefore, each number $up + jp$, $j = 1, \dots, Y$ has a prime factor $\leqslant X$. Moreover, let

$$P = \prod_{q \leqslant X} q$$

where the product is taken over primes q, and $z \equiv up \pmod{P}$. Then for $z \equiv up \pmod{P}$ each number $z + jp$, $j = 1, \dots, Y$ has a prime factor $\leqslant X$. It suffices to find $z \in (-x/2 - p, -x/2)$.

Therefore, each number $up + jp$, $j = 1, \dots, Y$ has a prime factor $\leqslant X$. Moreover, let

$$P = \prod_{q \leqslant X} q$$

where the product is taken over primes q, and $z \equiv up (\mathrm{mod}\, P)$. Then for $z \equiv up (\mathrm{mod}\, P)$ each number $z + jp$, $j = 1, \dots, Y$ has a prime factor $\leqslant X$. It suffices to find $z \in (-x/2 - p, -x/2)$.
We have $\log P \sim (\log x)/2$. Hence, $P = x^{1/2+o(1)} < y$. One can take a required number z as the number from $(-x/2 - P, -x/2) \subset (-x/2 - p, -x/2)$ congruent to up modulo P.

This result led to the questions: is it true that $\rho^*(x) - \pi(x) \ll x/\log^2 x$? Or even,

$$\rho^*(x) \leqslant 2\pi(x/2)? \tag{8}$$

A.Schizel offered a construction against these conjectures. Although he did not prove corresponding rigorous results, (8) has been disproved by D.A. Clark and N.C. Jarvic (2003) for $x = 130808636$.

The construction of A. Schinzel

/

A. Schinzel (1961/62) suggested the following construction. Let m be a positive integer and $p_1, \dots, p_m$ be the least primes. We will consider that $m < \sqrt{\log\log x}$. Apply the "hard" sieve of Eratosthenus to the interval $[1, x]$ eliminating all multiples of the primes $p \leqslant y$, except: for the distinguished primes $p_1, \dots, p_m$, we eliminate the congruence classes $n \equiv 1 (\mathrm{mod}\, p_i)$ instead of the classes $n \equiv 0 (\mathrm{mod}\, p_i)$. By U we denote the residual set. We consider that

$$y > x(\log x)^{-2} > \sqrt{x}.$$

The set U is the set of numbers

$$u = \prod_{r \in R} r^{\alpha(r)} P \not\equiv 1 (\mathrm{mod} s)$$

for all $s \in S$ where R and S are disjoint subsets of $\{p_1, \dots, p_m\}$, $R \cup S = \{p_1, \dots, p_m\}$, $\alpha(r) > 0$ for $r \in R$ and P is either 1 or a prime greater than y.
If we skip "greater than y" in the last sentence, then we get a larger set U_0. It can be shown that

$$|U_0| = \left(\sum_{i=1}^{m} \frac{(\log p_i) p_i}{(p_i - 1)^2} + o(1) \right) x (\log x)^{-2}.$$

For $m \to \infty$

$$\sum_{i=1}^{m} \frac{(\log p_i) p_i}{(p_i - 1)^2} \sim \log m.$$

Hence,

$$|U_0| - \pi(x) \sim (\log m) x (\log x)^{-2} \quad (m \to \infty,\ m < \sqrt{\log \log x}).$$

It is possible to show that

$$|U_0 \setminus U| \ll y(\log x)^{-1}(\log\log x)^m. \tag{9}$$

It is possible to show that

$$|U_0 \setminus U| \ll y(\log x)^{-1}(\log\log x)^m. \tag{9}$$

So, to have

$$|U| - \pi(x) \sim (\log m)x(\log x)^{-2} \quad (m \to \infty)$$

we need

$$y = o\left(x(\log x)^{-1}(\log\log x)^{-m}\log m\right).$$

It is possible to show that

$$|U_0 \setminus U| \ll y(\log x)^{-1}(\log\log x)^m. \tag{9}$$

So, to have

$$|U| - \pi(x) \sim (\log m)x(\log x)^{-2} \quad (m \to \infty)$$

we need

$$y = o\left(x(\log x)^{-1}(\log\log x)^{-m}\log m\right).$$

Thus, we could get

$$\left(\rho^*(x) - \pi(x)\right)\left(x(\log x)^{-2}\right) \to \infty \quad (x \to \infty)$$

if for some

$$y \leqslant x(\log x)^{-1}(\log\log x)^{-m}, \quad m \to \infty$$

the corresponding set U is admissible. Again, this is probably true, but the problem looks hopeless.

Roughly speaking, a local version of the problem is the following. Fix m. Then for a sufficiently large X and $Y = [X(\log X)^m]$ there is $u \in \mathbb{Z}$ such that each number $u+1, \dots, u+Y$ has a prime factor $\leqslant X$. It is supposed that this DOES NOT hold for $m > 2$. Let me recall that current results on large gaps between primes are based on corresponding local results.

Roughly speaking, a local version of the problem is the following. Fix m. Then for a sufficiently large X and $Y = [X(\log X)^m]$ there is $u \in \mathbb{Z}$ such that each number $u+1, \dots, u+Y$ has a prime factor $\leqslant X$. It is supposed that this DOES NOT hold for $m > 2$. Let me recall that current results on large gaps between primes are based on corresponding local results.

However, it is possible to use Schinzel's approach to improve a lower bound for $\rho^*(x) - \pi(x)$. Although we can't prove that the set U is admissible, we can remove a few elements from it to get an admissible set.

Theorem 1

For large x

$$\rho^*(x) - \pi(x) \gg x(\log x)^{-2} \log\log\log x.$$

Corollary 1

Assuming prime k- tuple conjecture, for large x there is y such that

$$\pi(x+y) - \pi(x) - \pi(y) \gg x(\log x)^{-2} \log\log\log x.$$

Justification of the title

The construction of A. Schinzel– many numbers in a short interval without small prime factors.
By $\rho^*(x)$ we denote the maximal cardinality of an admissible subset of the set $\{1, \dots, x\}$. We have estimated $\rho^*(x)$ from below.
An equivalent definition: $\rho^*(x)$ is the maximum over y of the numbers from $\{y+1, \dots, y+x\}$ without prime factors $\leqslant x$.

Thank you for your attention!