PRIMES IN EXPLICIT SHORT INTERVALS ON RH

ADRIAN W. DUDEK, LOÏC GRENIÉ, AND GIUSEPPE MOLTENI

Abstract. On the assumption of the Riemann hypothesis, we give explicit upper bounds on the difference between consecutive prime numbers.

1. General setting and results

The computation of the maximal prime gaps given by Oliveira e Silva, Herzog and Pardi [8, Sec. 2.2] verifies that \( p_{k+1} - p_k < \log^2 p_k \) for all primes \( 11 \leq p_k \leq 4 \cdot 10^{18} \). This proves that \( \forall x \in [5, 4 \cdot 10^{18}] \), there is a prime in \( [x-0.5 \log^2 x, x+0.5 \log^2 x] \). It is the purpose of this article to furnish new explicit upper bounds on the difference between consecutive prime numbers with the assumption of the Riemann hypothesis. Specifically, we prove the following theorem.

Theorem 1.1. Assume RH. Let \( x \geq 2 \) and \( c := \frac{1}{2} + \frac{2}{\log x} \). Then there is a prime in \( (x-c\sqrt{x} \log x, x+c\sqrt{x} \log x) \) and at least \( \sqrt{x} \) primes in \( (x-(c+1)\sqrt{x} \log x, x+(c+1)\sqrt{x} \log x) \).

The mentioned conclusion coming from the computations of Oliveira e Silva et al. is stronger than the first part of our result for all \( x \leq 4 \cdot 10^{18} \). This allows one to use \( c = 0.55 \) for all \( x \geq 5 \) when only one prime is needed.

In a recent paper [3], the first author proved Theorem 1.1 with \( c = \frac{2}{\pi} = 0.6366... \) and an asymptotic result in the weaker form \( c = 0.5+\epsilon \) when \( x \geq x(\epsilon) \), without any information on the size of \( x(\epsilon) \).

In Appendix A we prove the same result with \( c = 0.6102 \). This value improves on the one in [3] and is stronger than Theorem 1.1 up to \( 2 \cdot 10^8 \). Despite its weakness, we believe that its method of proof is worthy of interest.

We first consider the setting in which we seek to establish Theorem 1.1. Throughout, we define the von Mangoldt function as

\[
\Lambda(n) := \begin{cases} 
\log p & \text{if } n = p^m, \text{ } p \text{ is prime, } m \in \mathbb{N}, \text{ } m \geq 1 \\
0 & \text{otherwise,}
\end{cases}
\]

\( \vartheta(x) := \sum_{p \leq x} \log p \), where the sum is restricted to primes, and \( \psi(x) := \sum_{n \leq x} \Lambda(n) \). It is often convenient to work with a smoothed version of \( \psi \), and so we define

\[
\psi^{(1)}(x) := \int_0^x \psi(u) \, du = \sum_{n \leq x} \Lambda(n)(x-n)
\]

through partial summation. One can recall the integral representation

\[
\psi^{(1)}(x) = -\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{\zeta'(s)}{\zeta(s)} \frac{x^{s+1}}{s(s+1)} \, ds \quad \forall x \geq 1
\]

which follows directly from an application of Perron’s formula (see, for example, Ingham’s classic text [6, Ch IV, Sec 4]). We let \( h \in \mathbb{R} \) such that \( 0 < h < x \). Then

\[
\psi^{(1)}(x+h) - 2\psi^{(1)}(x) + \psi^{(1)}(x-h) = \sum_n \Lambda(n) K(x-n; h)
\]

where \( K(u; h) := \max\{h-|u|, 0\} \); one can verify this by expanding the left hand side of the above identity. Note also that \( K(u; h) \) is supported on \( |u| \leq h \), positive in the open set, and has a unique
It is not a difficult task to bound $\Sigma_2$ with $\Sigma_1$ where for the last inequality we have also used that $h$ runs on the set of nontrivial zeros of the Riemann zeta-function, $r$ and $r'$ are constants, and $|R^{(1)}(x)| \leq 0.6/x$ (one can see [11, Lemma 3.3], though this is classical). Noting that

$$\sum \sum \{0 : j = 0, 1, 2, \}$$

we thus have that, assuming $h \leq x/\sqrt{3}$,

$$\sum_n \Lambda(n)K(x-n; h) = h^2 - \sum_{\rho} \frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(p+1)} \cdot \frac{3\theta}{x}$$

for some $\theta = \theta(x, h) \in [-1, 1]$, for then

$$0.6((x+h)^{-1} + 2x^{-1} + (x-h)^{-1}) \leq 3x^{-1}.$$

We split the sum over the zeros as

$$\sum_{\rho} \frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(p+1)} =: \Sigma_1 + \Sigma_2,$$

with $\Sigma_1$ and $\Sigma_2$ representing the sums on zeros with $|\Im(\rho)| \leq T$ and $|\Im(\rho)| > T$, respectively. It is not a difficult task to bound $\Sigma_2$ (here we repeat the argument in [9]). In fact, assuming RH,

$$|\Sigma_2| \leq 4(x+h)^{3/2} \sum_{|\Im(\rho)| > T} \frac{1}{|\rho(p+1)|},$$

and since $\sum_{|\Im(\rho)| > T} \frac{1}{|\rho(p+1)|} \leq \frac{\log T}{\pi T}$ (see [13, Lemma 1 (ii)]), one has

$$|\Sigma_2| \leq 4(x+h)^{3/2} \frac{\log T}{\pi T}.$$

Thus

$$\sum_n \Lambda(n)K(x-n; h) \geq h^2 - 4(x+h)^{3/2} \frac{\log T}{\pi T} \cdot \frac{3}{x}.$$

Now we remove the contribution from prime powers. Recalling that

$$0.9986\sqrt{x} \leq \psi(x) - \vartheta(x) \leq (1 + 10^{-6})\sqrt{x} + 3\sqrt{x}$$

for every $x \geq 121$ (see [12, Th. 6] and [10, Cor. 2]), we get

$$\sum_n \Lambda(n)K(x-n; h) \leq h \sum_{|n-x| < h} \Lambda(n) \leq h \left( \sum_{|p-x| < h} \log p + (\psi(x+h) - \vartheta(x+h)) - (\psi(x-h) - \vartheta(x-h)) \right) \leq h \left( \sum_{|p-x| < h} \log p + (1 + 10^{-6})\sqrt{x+h} + 3\sqrt{x+h} - 0.9986\sqrt{x-h} \right) \leq h \left( \sum_{|p-x| < h} \log p + 0.002\sqrt{x} + 3\sqrt{x} + \frac{2h}{\sqrt{x}} \right)$$

where for the last inequality we have also used that $h \leq x/\sqrt{3}$. Thus, when $x \geq 121$ we have

$$\sum_{|p-x| < h} \log p \geq h - \frac{1}{h}(\Sigma_1 - 4(x+h)^{3/2} \frac{\log T}{\pi h T} - 0.002\sqrt{x} - 3\sqrt{x} - \frac{2h}{\sqrt{x}} - \frac{3}{xh}).$$

It is clear that the positivity of the right hand side guarantees the existence of at least one prime in the interval $(x-h, x+h)$. 

(1.1)
From before, we have that
\[
\Sigma_1 = \sum_{|\text{Im}(\rho)| \leq T} \frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(\rho+1)}.
\]

There are essentially two ways to bound \(\Sigma_1\), both of them appearing already in [3].

The first one is based on the Taylor identity
\[
(1+\epsilon)^{d+i\gamma\epsilon} - 2 + (1-\epsilon)^{d+i\gamma\epsilon} = -4\sin^2(\gamma\epsilon) + O(\gamma^2\epsilon),
\]
while the second one is based on the identity
\[
\frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(\rho+1)} = \int_{x-h}^{x+h} K(x-u; h) u^{\rho-1} du.
\]

Thus, denoting \(\gamma\) the imaginary part of a nontrivial zero, on the assumption of RH one gets
\[
|\Sigma_1| \leq 4 \sum_{|\gamma| \leq T} \frac{\sin^2(\gamma\epsilon) + O(\gamma^2\epsilon)}{\gamma^2}
\]
from the first one, and
\[
|\Sigma_1| \leq \int_{x-h}^{x+h} K(x-u; h) \sum_{|\gamma| \leq T} \left|\frac{du}{\sqrt{u}}\right|
\]
from the second one. As a consequence, the first approach takes advantage of the cancellation due to the sum of the three functions \((1+\omega\epsilon)^{d+i\gamma\epsilon}\) with \(\omega \in \{0, \pm 1\}\) for the same zero, while the second approach takes advantage of the cancellation coming from the sum of values of the same function computed at different zeros.

The first approach is discussed in Section 2 while the second is discussed in Appendix A.

2. First bound for \(\Sigma_1\)

Let \(N(T)\) denote the number of nontrivial zeros of \(\zeta(s)\) in the upper half of the critical strip, where multiplicity is included. We state the estimate of \(N(T)\) done by Trudgian in [14]: let \(W(T) := \frac{2}{\pi} \log \left( \frac{T}{\log T} \right)\) denote what is essentially the main term of \(N(T)\) and let \(U(T) := N(T) - W(T)\), then the result says that
\[
|U(T)| \leq 0.112 \log T + 0.278 \log \log T + 2.51 + \frac{0.2}{T} + \frac{7}{8} =: R(T), \quad T \geq e.
\]

Note that \(U(2\pi) = 1\) because the imaginary part of the first zero is \(14.13\ldots\), and that \(dW(T) = \log\left(\frac{T}{e}\right)\frac{dT}{\sqrt{T}}\).

We introduce the notations
\[
T = \frac{\beta}{c \log x}, \quad h = c \sqrt{x} \log x
\]
for suitable \(\beta\) and \(c\).

Lemma 2.1. Let \(0 \leq h < x\). Then for every \(\gamma \in \mathbb{R}\) there exists \(\theta \in \mathbb{C}\) with \(|\theta| \leq 1\) such that
\[
\left(1 + \frac{h}{x}\right)^{\frac{1}{2} + i\gamma} - 2 + \left(1 - \frac{h}{x}\right)^{\frac{1}{2} + i\gamma} = -4\sin^2\left(\frac{\gamma h}{2x}\right) + \theta(2|\gamma| + 1)\frac{h^2}{x^2}.
\]

Proof. The proof is straightforward and follows from the Taylor expansion of \(\log(1+u)\) and some elementary inequalities. \(\square\)

Thus we get an explicit version of (1.2):
\[
|\Sigma_1| \leq 8x^{3/2} \sum_{0 < \gamma \leq T} \frac{\sin^2\left(\frac{\gamma h}{2x}\right)}{\gamma^2} + 2 \frac{h^2}{\sqrt{x}} \sum_{0 < \gamma \leq T} \frac{2\gamma+1}{\gamma^2}.
\]
Using the inequality $|\Sigma_1| \leq 8x^{3/2}$, we get

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \leq \sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} + \left( \frac{\log^2 T}{\pi} - \frac{1}{20} \right) \frac{h^2}{\sqrt{x}}.
\]

Let $\gamma_1 = 14.13 \ldots$ be the imaginary part of the first non-trivial zero of $\zeta(s)$. By partial summation we get

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} = \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} dN(\gamma) = \left[ \frac{\sin^2 (\frac{\gamma h}{T}) N(\gamma)}{\gamma^2} \right]_{\gamma_1}^{T} - \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} N(\gamma) d\gamma
\]

It then follows that

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} = \frac{\sin^2 (\frac{\gamma h}{T}) U(T)}{T^2} - \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \frac{\gamma}{2\pi e} log \left( \frac{\gamma_1}{2\pi e} \right) d\gamma - \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} U(\gamma) d\gamma.
\]

Recalling the upper bound $|U(T)| \leq R(T)$ and noticing that $\gamma_1 < 2\pi e$, we get

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \leq \frac{R(T)}{T^2} + \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \frac{\gamma}{2\pi e} log \left( \frac{\gamma_1}{2\pi e} \right) d\gamma - \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} U(\gamma) d\gamma
\]

Using the inequality $|\sin^2 \theta| < \frac{3}{4}$, we simplify to get

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \leq \frac{R(T)}{T^2} + \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \frac{\gamma}{2\pi} \frac{5h}{4x} \int_{\gamma_1}^{T} R(\gamma) \frac{1}{\gamma^2} d\gamma.
\]

Since $\int_{14^{\infty}} \frac{R(\gamma)}{\gamma^2} d\gamma \leq 0.297$, we get

\[
\sum_{0 < \gamma \leq T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} \leq \frac{1}{2\pi} \int_{\gamma_1}^{T} \frac{\sin^2 (\frac{\gamma h}{T})}{\gamma^2} d\gamma \log \left( \frac{T}{2\pi} \right) + \frac{R(T)}{T^2} + \frac{2.97h}{8x}
\]

We can bound the integral in the above equation with ease, for

\[
\int_{0}^{y} \frac{\sin^2 t}{t^2} dt = \frac{\pi}{2} - \int_{0}^{\infty} \frac{\sin^2 t}{t^2} dt = \frac{\pi}{2} - \int_{0}^{\infty} \frac{1 - \cos(2t)}{2t^2} dt = \frac{\pi}{2} - \int_{0}^{\infty} \frac{\cos(2t)}{2t^2} dt = \frac{\pi}{2} - \int_{0}^{\infty} \frac{\sin(2t)}{2t^3} dt + \int_{0}^{\infty} \frac{\sin(2t)}{2t^3} dt = \frac{\pi}{2} - \frac{1}{2y} - \frac{\sin(2y)}{4y^2} + \frac{\theta}{4y^2} = \frac{\pi}{2} - \frac{1}{2y} + \frac{\theta}{4y^2},
\]
for some \( \theta \in [-1, 1] \). As such, we can now use \( R(T) \leq 1.5 \log T \) for every \( T \geq \gamma_1 \) to get
\[
\sum_{0 < x \leq T} \frac{\sin^2 \left( \frac{\theta x}{\sqrt{\gamma}} \right)}{\gamma^2} \leq \frac{h}{8x} \left( 1 - \frac{2}{\pi} \frac{x}{hT} + \frac{4}{\pi} \frac{x^2}{h^2 T^2} \right) \log \left( \frac{T}{2\pi} \right) + 1.5 \frac{\log T}{T^2} + \frac{2.97h}{8x}
\]
so that (2.1) becomes
\[
|\Sigma_1| \leq 8x^{3/2} \left( \frac{h}{8x} \left( 1 - \frac{2}{\pi} \frac{x}{hT} + \frac{4}{\pi} \frac{x^2}{h^2 T^2} \right) \log \left( \frac{T}{2\pi} \right) + 1.5 \frac{\log T}{T^2} + \frac{2.97h}{8x} \right) + \left( \frac{\log^2 T}{\pi} + \frac{1}{20} \right) \frac{h^2}{\sqrt{x}}
\]
\[
= h \sqrt{x} \left( \frac{1 - \frac{2}{\pi} \frac{x}{hT} + \frac{4}{\pi} \frac{x^2}{h^2 T^2}}{\sqrt{x}} \right) \log \left( \frac{T}{2\pi} \right) + 12x^{3/2} \log T + 2.97h \sqrt{x} + \left( \frac{\log^2 T}{\pi} + \frac{1}{20} \right) \frac{h^2}{\sqrt{x}}
\]
(2.2) \[ + 2.97h \sqrt{x} + \log^2 T \frac{h^2}{\sqrt{x}} + 12x^{3/2} \log T - \frac{h^2}{20 \sqrt{x}}. \]

3. Proof of Theorem 1.1

Substituting (2.2) into (1.1) we get
\[
\sum_{|p-x| < h} \log p \geq h - \sqrt{x} \log \left( \frac{T}{2\pi} \right) - \frac{2}{\pi} \sqrt{x} - \frac{h}{\sqrt{x}} \log \left( \frac{T}{2\pi} \right) + \frac{4}{\pi} \sqrt{x} \frac{x^2}{h^2 T^2} \log \left( \frac{T}{2\pi} \right)
\]
\[ + 2.97 \sqrt{x} + \frac{\log^2 T}{\pi} \frac{h}{\sqrt{x}} + 12x^{3/2} \log T + \frac{h}{20 \sqrt{x}} - 4(x+h)^{3/2} \frac{\log T}{\pi h T} - 0.002 \sqrt{x} - 3 \sqrt{x} - \frac{2h}{x h} - \frac{3}{x h}.
\]
Recalling that we have set \( h = c \sqrt{x} \log x \), \( T = \frac{\pi}{x \log x} \) (so that \( hT = \beta x \)), and estimating \((x+h)^{3/2} \leq x^{3/2}(1+2h/x)\) (which holds whenever \( h/x \leq 1.6\)), we have that
\[
\sum_{|p-x| < h} \log p \geq h - \sqrt{x} \log \left( \frac{T}{2\pi} \right) + \frac{2}{\pi} \sqrt{x} - \frac{h}{\sqrt{x}} \log \left( \frac{T}{2\pi} \right) - 4 \frac{x^2}{\beta h} \log T - 3 \sqrt{x} - 2.05 \frac{h}{\sqrt{x}} - \frac{3}{x h},
\]
or, upon gathering like terms, that
\[
\sum_{|p-x| < h} \log p \geq h - \sqrt{x} \log T \left( \frac{2}{\pi} \sqrt{x} - \frac{h}{\sqrt{x}} \log \frac{T}{2\pi} - (3-\log(2\pi)) \sqrt{x} \right)
\]
\[ - \frac{2}{\pi} \log(2\pi) \frac{\sqrt{x}}{\beta} - 4 \frac{\sqrt{x}}{\beta^2} \log \left( \frac{T}{2\pi} \right) - 3 \sqrt{x} - 8 \log \frac{T}{\beta} \frac{h}{\sqrt{x}} - \log^2 \frac{T}{\beta} \frac{h}{\sqrt{x}} - \left( 2.05 + 12 \frac{\log T}{\beta^2} \right) \frac{h}{\sqrt{x}} - \frac{3}{x h}.
\]
For this computation it is convenient to take \( \beta = \beta(x) \) and diverging as \( x \) goes to \( \infty \). To ensure the best result we have to set \( \beta \) so that the sum \( \log T + \frac{2}{\pi} \frac{\log T}{\beta} \) is minimised. This sum is, up to terms of lower order in \( \beta \),
\[
\log \beta + \frac{\log x}{\pi \beta}.
\]
This last sum is minimum when
\[
\beta = \frac{1}{\pi x} \sqrt{x}.
\]
Thus we have \( T = \frac{1}{x \sqrt{x}} \). With this, the lower bound (3.1) becomes
\[
\sum_{|p-x| < h} \log p \geq h - \frac{1}{2} \sqrt{x} \log x - (4 - \log(2\pi^2)) \sqrt{x} + o(\sqrt{x}).
\]
Using the fact that \( c \geq 1/2 \), which is our best hope at the moment, in order to have a positive lower bound it is sufficient to take

\[
h \geq \left( \frac{1}{2} + \frac{d}{\log x} \right) \sqrt{x} \log x
\]

for any \( d > 4 - 2 \log \pi = 1.7105 \ldots \), when \( x \) is large enough. Actually, the choice \( d = 1.72 \) holds only for \( x \geq \exp(590) \approx 2 \cdot 10^{256} \). On the contrary, the choice \( d = 2 \) holds for \( x \geq 7.5 \cdot 10^8 \). Thus the claim asserting the existence of a prime when \( c = 1/2 + 2/\log x \) is proved for \( x \geq 7.5 \cdot 10^8 \).

Moreover, the upper bound \( \log(x+h) \sum_{|p-x|<h} 1 \geq \sum_{|p-x|<h} \log p \) and \( (3.1) \) prove the existence of \( \sqrt{x} \) primes in \( (x-(c+1)\sqrt{x} \log x, x+(c+1)\sqrt{x} \log x) \) for \( x \geq 1.4 \cdot 10^5 \). Lastly, for \( x \in [2,1.4 \cdot 10^9] \) it is sufficient to check that \( p_{k+1} - p_k \leq 2\varepsilon \sqrt{p_k} \log p_k \) (which gives the claim for \( x \in [p_k, p_{k+1}] \)) when \( k \leq 13010 \).

4. An Application

On the Riemann hypothesis, Cramér [2] was the first to prove the bound \( p_{n+1} - p_n \ll \sqrt{p_n} \log p_n \), and he noted the implication that there exists some constant \( \alpha > 0 \) such that there will be a prime in the interval

\((n^2, (n+\alpha \log n)^2)\)

for all sufficiently large \( n \). This was intended for comparison to Legendre’s conjecture that there is a prime in the interval \((n^2, (n+1)^2)\) for all \( n \). The following corollary of Theorem 1.1 states that one can take \( \alpha = \frac{1}{2} + o(1) \).

**Corollary 4.1.** Assume RH. Then for every integer \( n \geq 2 \) there is a prime in the interval

\((n^2, (n+\alpha \log n)^2)\)

where

\[\alpha := \frac{1}{2} + \frac{2}{\log n}\]

**Proof.** Let

\[\delta := d \left( 1 + \frac{4d \log n}{n} \right)^{1/2} \left( 1 + \frac{4d}{n} \right)\]

with \( d := \frac{1}{2} + \frac{1}{\log n} \). We will start by proving that there is a prime in the interval

\((n^2, (n+\delta \log n)^2)\).

For \( n = 2, \ldots, 6 \) the statement is true. We thus assume \( n \geq 7 \). We compare the interval \((n^2, (n+\delta \log n)^2)\) to that of Theorem 1.1 by setting \( x \) such that

\[x-d\sqrt{x} \log x = n^2\]

where \( d \) is as stated above. We can note that as we have \( n^2 < x \) and thus

\[1/\log n > \frac{2}{\log x}\]

it follows that \( d \geq c \) and so there is a prime in the interval \((x-d\sqrt{x} \log x, x+d\sqrt{x} \log x)\). Therefore, all we need to prove is that

\[(n+\delta \log n)^2 \geq x+d\sqrt{x} \log x\]

We can prove this by manipulating the expression for \( \delta \). Clearly, as \( e < n < \sqrt{x} \) we have that

\[\delta > d \left( 1 + \frac{2d \log x}{\sqrt{x}} \right)^{1/2} \left( 1 + \frac{4d}{\sqrt{x}} \right)\]

Noting that the inequality \( 1+2u > \frac{1}{1-u} \) holds for all \( u < 1/2 \), we have that

\[\delta > d \left( \frac{1}{1-\frac{\log x}{\sqrt{x}}} \right)^{1/2} \left( \frac{1}{1-\frac{2d}{\sqrt{x}}} \right)\]
We then apply the inequality \( \log(1-u) > -2u \) which holds in the range \( u < 1/2 \) to \( u = d \frac{\log x}{x} \) to get
\[
\delta > d \left( \frac{x}{x-d \sqrt{x \log x}} \right)^{1/2} \left( \frac{\log x}{\log(x-d \sqrt{x \log x})} \right).
\]
It follows now upon rearranging that
\[
2\delta n \log n > d \sqrt{x} \log x,
\]
or, adding \( n^2 \) to both sides, that
\[
n^2 + 2\delta n \log n > x + d \sqrt{x} \log x.
\]
The inequality (4.1) follows from this trivially and so the proof is complete for \( \delta \).

For \( n \geq 22 \), we have \( \alpha > \delta \) so that the corollary is proved in that case; for \( 2 \leq n \leq 21 \) the statement of the corollary is true. \( \square \)

Now, upon setting \( n = p_k \) in the above corollary, it follows that there is a prime in the interval
\[
(p_k^2, (p_k+\alpha \log p_k)^2)
\]
for all \( k \geq 1 \). It should be noted, as \( \delta = \frac{1}{2} + o(1) \) and the average gap between \( p_k \) and \( p_{k+1} \) is \( \log p_k \), that something can be said here about the existence of primes in the interval \( (p_k^2, p_{k+1}^2) \); this is related to the so-called Brocard conjecture predicting the existence of four primes at least in this interval (see for instance Ribenboim [11, p. 248]).

It was first proven by Cramér [2], on RH, that the number of \( n < x \) such that there is no prime in the interval \( (n^2, (n+1)^2) \) is \( \mathcal{O}(x^{2/3+\epsilon}) \) (improved to \( \mathcal{O}(x^{1/2+\epsilon}) \) unconditionally and to \( \mathcal{O}((\log x)^{2+\epsilon}) \) on RH in [1]), and from this it follows that there is a prime in almost all intervals of the form \( (p_{k+1}^2, p_{k+2}^2) \). However, there may still be infinitely many exceptions, though the following corollary assures us that the exceptions must occur when the prime gap is essentially less than half the average gap.

**Corollary 4.2.** Assume RH. Suppose that \( p_k \) and \( p_{k+1} \) are consecutive primes satisfying
\[
p_{k+1}-p_k \geq \alpha \log p_k + \frac{\alpha^2 \log^2 p_k}{2p_k}
\]
where \( \alpha := \frac{1}{2} + \frac{2}{\log p_k} \). Then there is a prime in the interval \( (p_k^2, p_{k+1}^2) \).

**Proof.** First, it follows that
\[
p_{k+1}-p_k > \frac{2\alpha p_k \log p_k + \alpha^2 \log^2 p_k}{p_k + p_{k+1}}.
\]
It is straightforward to rearrange this so that
\[
p_{k+1}^2 > (p_k + \alpha \log p_k)^2
\]
and, with reference to Corollary 4.1, this completes the proof. \( \square \)

**Appendix A. Second bound for \( \Sigma_1 \)**

Since \( N(T) \leq \frac{T}{2\pi} \log T \) (see [4] Corollary 1)), from (3.3) one has
\[
|\Sigma_1| \leq \frac{2N(T)}{\sqrt{x-h} \int_{x-h}^{x+h} K(x-u; h) \, du} = \frac{2h^2}{\sqrt{x-h}} N(T) \leq \frac{h^2 T}{\pi \sqrt{x-h}} \log T,
\]
which is the way this sum is estimated in [3]. We improve the result by proving the existence of a cancellation for the sum \( \sum_{|\gamma| \leq T} u^\gamma \). The structure of the counting function \( N(T) \) alone, that is, the fact that \( N(T) = \frac{T}{2\pi} \log T + O(\log T) \), is not sufficient to ensure a cancellation in \( \sum_{|\gamma| \leq T} u^\gamma \) for every \( u \). To see this, one can consider a set of points generated in this way: in the neighborhood of every integer \( n \) there is a cloud of \( \left[ \frac{1}{2} \log n \right] \) points which are placed very close to \( n \). Their counting function satisfies the same formula as \( N(T) \), size of the remainder included. For this set, however, one has \( \sum_{|\gamma| \leq T} u^\gamma \gg T \log T \) when \( u = e^{2\pi} \), and similarly for every \( u = e^{2\pi k} \) when \( k \in \mathbb{N} \) is small with respect to \( T \).
Thus, we can furnish a cancellation essentially in two ways: either we assume some hypothesis about the distribution of the imaginary parts of the zeroes of $\zeta(s)$ (for example the Pair Correlation Conjecture, as done in [5] and in [7]), or we try to prove a cancellation in some mean sense. The second possibility appears promising since in our computation the estimated object appears naturally in an integral and produces a result not depending on a further unproved hypothesis. In this way we can prove Theorem 1.1 with $c = 0.6102$.

**Cancellation in mean.** We let

$$S_\alpha(T) := \sum_{|\gamma| \leq T} e^{i\alpha \gamma},$$

keep the notations

$$T = \frac{\beta \sqrt{x}}{c \log x}, \quad h = c \sqrt{x} \log x,$$

and introduce

$$a := \log(x-h), \quad b := \log(x+h),$$

$$A := \frac{b-a}{2} = \frac{1}{2} \log \left( \frac{x+h}{x-h} \right), \quad B := \frac{a+b}{2} = \frac{1}{2} \log(x^2-h^2).$$

Notice that $A \sim h/x \approx T^{-1}$ and $B \sim \log x$ as $x$ diverges to infinity.

**Proposition A.1.** Assume RH. Suppose $\beta \geq 1$, $c \leq 1$ and $x \geq 2$. Then

$$\int_a^b |S_\alpha(T)|^2 \, d\alpha \leq \frac{1}{\pi^2} F(\alpha)T \log^2 \left( \frac{T}{2\pi} \right) + H(A,T)$$

with

$$F(y) := \frac{1}{y} \int_{-y}^{y} |\text{sinc}(u)| \, du \, du'$$

where $\text{sinc}(x) := \frac{\sin x}{x}$, and

$$H(A,T) := \frac{4}{\pi} (2+AT)AT(R(T)+1) \log \left( \frac{T}{2\pi} \right) + 8A \left( 1 + AT + \frac{1}{3} (AT)^2 \right) (R(T)+1)^2.$$

**Remark.** Once the orders of $h$ and $T$ as functions of $x$ are considered, the trivial bound and the new bound are respectively

$$\int_a^b |S_\alpha(T)|^2 \, d\alpha \leq \frac{2\beta}{\pi^2} T \log^2 T \quad \text{v.s.} \quad \int_a^b |S_\alpha(T)|^2 \, d\alpha \leq \frac{1+o(1)}{\pi^2} F(\beta)T \log^2 T,$$

and it is easy to see that the second one improves on the first one for every $\beta > 0$ as $T \to \infty$.

**Proof.** First, we have the series of working:

$$\int_a^b |S_\alpha(T)|^2 \, d\alpha = \Re \int_a^b |S_\alpha(T)|^2 \, d\alpha = \Re \sum_{|\gamma|,|\gamma'| \leq T} \int_a^b e^{i\alpha (\gamma - \gamma')} \, d\alpha$$

$$= \Re \sum_{|\gamma|,|\gamma'| \leq T} \frac{e^{ib(\gamma - \gamma')} - e^{ia(\gamma - \gamma')}}{i(\gamma - \gamma')} = 2\Re \sum_{|\gamma|,|\gamma'| \leq T} e^{iB(\gamma - \gamma')} \frac{\sin \left( A(\gamma - \gamma') \right)}{\gamma - \gamma'}$$

$$= 4 \sum_{\substack{0<\gamma<T \\ -T<\gamma'<T}} \cos \left( B(\gamma - \gamma') \right) \sin \left( A(\gamma - \gamma') \right) \leq 4A \sum_{\substack{0<\gamma<T \\ -T<\gamma'<T}} |\sin \left( A(\gamma - \gamma') \right)|. $$

In the following we use the following bounds for $\text{sinc}(x)$:

$$\| \text{sinc} \|_\infty \leq 1, \quad \| \text{sinc}' \|_\infty \leq 1/2, \quad \| \text{sinc}'' \|_\infty \leq 1/3.$$
These follow immediately from the representation $2 \text{sinc}(x) = \int_{-1}^{1} e^{ixy} dy$. We thus write the double sum on zeros as a Stieltjes integral. Recalling that the imaginary part of the first zero exceeds $2\pi$, we get

$$
\int_{a}^{b} |S_n(T)|^2 \, d\omega \leq 4A \int_{0 < \gamma \leq T} \int_{-T \leq \gamma' \leq T} |\text{sinc} \left( A(\gamma - \gamma') \right)| \, dN(\gamma) \, dN(\gamma')
$$

$$
= 4A \int_{2\pi < \gamma \leq T} \int_{-\gamma \leq \gamma' \leq \gamma} \left( |\text{sinc} (A(\gamma - \gamma'))| + |\text{sinc} (A(\gamma + \gamma'))| \right) dN(\gamma) \, dN(\gamma').
$$

To ease matters, we employ the notation

$$
f(t_1, t_2) := |\text{sinc}(A(t_1-t_2))| + |\text{sinc}(A(t_1+t_2))|
$$

which allows us to write that

$$
\int_{a}^{b} |S_n(T)|^2 \, d\omega = 4A \int_{2\pi < \gamma \leq T} \int_{-\gamma \leq \gamma' \leq \gamma} f(\gamma, \gamma') \, dW(\gamma) \, dW(\gamma') + 4A \int_{2\pi < \gamma \leq T} \int_{\gamma < \gamma' \leq \gamma} f(\gamma, \gamma') \, dW(\gamma) \, dU(\gamma')
$$

$$
+ 4A \int_{2\pi < \gamma \leq T} \int_{\gamma < \gamma' \leq \gamma} f(\gamma, \gamma') \, dU(\gamma) \, dU(\gamma').
$$

We write the sum of the above four integrals as $I + II + III + IV$ where the order is kept. It thus remains to estimate separately the contribution of each integral. The first one produces the main term, for

$$
I \leq \frac{4A}{(2\pi)^2} \left[ \int_{2\pi < \gamma \leq T} \int_{-\gamma \leq \gamma' \leq \gamma} f(\gamma, \gamma') \, d\gamma \, d\gamma' \right] \log^2 \left( \frac{T}{2\pi} \right)
$$

$$
\leq \frac{1}{\pi^2} \left[ \int_{0 < u \leq AT} \int_{u \leq u' \leq AT} |\text{sinc}(u-u')| \, du \, du' \right] \log^2 \left( \frac{T}{2\pi} \right) = \frac{1}{\pi^2} F(AT) T \log^2 \left( \frac{T}{2\pi} \right).
$$

In estimating the integral $II$, an application of integration by parts gives (note that $\partial_\gamma |\text{sinc}(A(\gamma \pm \gamma'))|$ has only jump singularities, so the formula still holds)

$$
II = \frac{4A}{2\pi} \int_{2\pi}^{T} \left[ \int_{2\pi}^{T} f(\gamma, \gamma') \log \left( \frac{\gamma}{2\pi} \right) \, d\gamma \right] \, dU(\gamma')
$$

$$
= \frac{2A}{\pi} \left[ \int_{2\pi}^{T} f(\gamma, \gamma') \log \left( \frac{\gamma}{2\pi} \right) \, d\gamma \right] U(\gamma') \bigg|_{2\pi}^{T} - \frac{2A}{\pi} \int_{2\pi}^{T} \left[ \int_{2\pi}^{T} \partial_\gamma f(\gamma, \gamma') \log \left( \frac{\gamma}{2\pi} \right) \, d\gamma \right] \, dU(\gamma')
$$

We can estimate it as (recall that $U(2\pi) = 1$)

$$
\leq \frac{2A}{\pi} \int_{2\pi}^{T} \| \text{sinc} \|_\infty \log \left( \frac{\gamma}{2\pi} \right) \, d\gamma (R(T)+1) + \frac{2A^2}{\pi} \int_{2\pi}^{T} \int_{2\pi}^{T} \| \text{sinc}' \|_\infty \log \left( \frac{\gamma}{2\pi} \right) |U(\gamma')| \, d\gamma \, d\gamma'
$$

$$
\leq \frac{4A}{\pi} AT (R(T)+1) \log \left( \frac{T}{2\pi} \right) + \frac{2A^2}{\pi} (AT)^2 R(T) \log \left( \frac{T}{2\pi} \right)
$$

$$
\leq \frac{2}{\pi} (2+AT) AT (R(T)+1) \log \left( \frac{T}{2\pi} \right).
$$

The contribution of $III$ equals that of $II$, for we note the symmetry of the integral under the transposition $\gamma \leftrightarrow \gamma'$. And so, lastly we have

$$
IV = 4A \int_{2\pi}^{T} \left[ \int_{2\pi}^{T} f(\gamma, \gamma') \, dU(\gamma') \right] \, dU(\gamma')
$$

$$
= 4A \int_{2\pi}^{T} \left[ f(\gamma, \gamma') U(\gamma') \right] \, dU(\gamma') - 4A \int_{2\pi}^{T} \left[ \int_{2\pi}^{T} \partial_\gamma (f(\gamma, \gamma')) U(\gamma) \, d\gamma \right] \, dU(\gamma')
$$

$$
= 4A \int_{2\pi}^{T} f(T, \gamma') U(T) \, dU(\gamma') - 4A \int_{2\pi}^{T} f(2\pi, \gamma') \, dU(\gamma').
$$
\[-4A \int_{2\pi}^{T} \left[ \int_{2\pi}^{T} \partial_{\gamma} \left( f(\gamma, \gamma') \right) U(\gamma) \, d\gamma \right] \, dU(\gamma'),\]

where a second integration by parts gives

\[ IV = 4A \int_{2\pi}^{T} R(T) U(\gamma') | d\gamma' | + 8A | U(T) | + 8A + 4A^2 \int_{2\pi}^{T} | U(\gamma') | \, d\gamma' \]

\[ + 4A^2 \int_{2\pi}^{T} | U(\gamma') | \, d\gamma' (R(T) + 1) + 4A^3 \int_{2\pi}^{T} \int_{2\pi}^{T} 2 | \text{sinc}'' \| \infty | U(\gamma') | \, d\gamma \, d\gamma'.\]

Estimating the integrals, one has that

\[ IV \leq 8A R^2(T) + 8A R(T) + 4A R^2(T) + 8A R(T) + 8A + 4A^2 T R(T) \]

\[ + 4A^2 T R(T) \left( R(T) + 1 \right) + 4A^3 \frac{2}{3} T^2 R^2(T) \]

\[ = 8A \left( (R(T) + 1)^2 + A T (R(T) + 1) R(T) + \frac{1}{3}(A T)^2 R^2(T) \right) \]

\[ \leq 8A \left( 1 + A T + \frac{1}{3}(A T)^2 \right) (R(T) + 1)^2.\]

Therefore, the contribution of \( II, III \) and \( IV \) is bounded by

\[ 2 \frac{T}{\pi} (2 + A T) A T (R(T) + 1) \log \left( \frac{T}{2\pi} \right) + 8A \left( 1 + A T + \frac{1}{3}(A T)^2 \right) (R(T) + 1)^2,\]

which is \( H(A, T) \).

\[ \square \]

**Estimation of \( \Sigma_1 \).** We now use the above result on cancellation to estimate the first sum over the zeroes. The Cauchy-Schwarz inequality yields the bound

\[ |\Sigma_1| \leq \int_{x-h}^{x+h} K(x-u; h) \left| \sum_{|\gamma| \leq T} e^{i\gamma \log u} \right| \, du \]

\[ = \int_{\log(x-h)}^{\log(x+h)} e^{\alpha/2} K(x-e^\alpha; h) \left| \sum_{|\gamma| \leq T} e^{i\gamma} \right| \, d\alpha \]

\[ \leq \left[ \int_{\log(x-h)}^{\log(x+h)} e^{\alpha} K^2(x-e^\alpha; h) \, d\alpha \right]^{1/2} \left[ \int_{\log(x-h)}^{\log(x+h)} |S_\alpha(T)|^2 \, d\alpha \right]^{1/2},\]

and so we have that

\[ |\Sigma_1| \leq \left[ \int_{-h}^{h} K^2(u; h) \, du \right]^{1/2} \left[ \int_{\log(x-h)}^{\log(x+h)} |S_\alpha(T)|^2 \, d\alpha \right]^{1/2} \]

\[ = \sqrt{\frac{2}{3}} h^{3/2} \left[ \int_{\log(x-h)}^{\log(x+h)} |S_\alpha(T)|^2 \, d\alpha \right]^{1/2}.\]

We can now apply Proposition \( \square \) to get the estimate

\[ |\Sigma_1| = \sqrt{\frac{2}{3}} h^{3/2} \left( \frac{1}{\pi^2} F(A T) T \log^2 \left( \frac{T}{2\pi} \right) + H(A, T) \right)^{1/2}.\]
\[ = \frac{1}{\pi^3} \left( \frac{2 \beta}{3} F(AT) + \frac{2 \beta^2}{3} \frac{H(A,T)}{T \log^2(T/2\pi)} \right)^{1/2} h \sqrt{T \log \left( \frac{T}{2\pi} \right)} \]

(A.1)

\[ = \frac{1}{\pi} \left( \frac{2 \beta}{3} F(AT) + \frac{2 \beta^2}{3} \frac{H(A,T)}{T \log^2(T/2\pi)} \right)^{1/2} h \sqrt{T \log \left( \frac{T}{2\pi} \right)}. \]

We now proceed to prove the analog of Theorem 1.1.

First claim. From (1.1) and the bound (A.1) for \( \Sigma \) we get

\[
h^{-1} \sum_{|p-x|<h} \log p \geq 1 - \left( \frac{2 \beta}{3} F(AT) + \frac{2 \beta^2}{3} \frac{H(A,T)}{T \log^2(T/2\pi)} \right)^{1/2} \frac{\sqrt{x}}{\pi h} \log \left( \frac{T}{2\pi} \right) - 4(x+h)^{3/2} \frac{\log T}{\pi h^2 T}
- 0.002 \frac{\sqrt{x}}{h} - 3 \frac{\sqrt{x}}{h} \frac{2}{\sqrt{x}} \frac{3}{x^2} \]

when \( x \geq 121 \). When \( x \) diverges to infinity, Inequality (A.2) becomes

\[
h^{-1} \sum_{|p-x|<h} \log p \geq 1 - \frac{\alpha}{c} \left( \frac{2 \alpha}{c} + o(1) \right) \frac{\log x}{x}, \]

where \( \alpha := \frac{1}{3} \left( \sqrt{\frac{2 \pi}{A(T)}} \right) \), uniformly for \( \beta \) and \( c \) in any compact set of \((0, +\infty)\). The minimum of \( \alpha \) is attained for \( \beta = \beta_{\text{min}} := 2.4934 \ldots \) and is \( \alpha_{\text{min}} := 0.61019 \ldots \). Thus, setting \( \beta = \beta_{\text{min}} \) and \( c = \alpha_{\text{min}} \), the right hand side of (A.3) is positive when \( x \) is large enough. Inserting \( \beta = 2.493 \) and \( c = 0.6102 \) directly into (A.2) one obtains an inequality where the function appearing on the right hand side is positive whenever \( x \geq 16000 \), so that the theorem is proved in this range. Lastly, for \( x \in [2, 16000] \) it is sufficient to check that \( p_{k+1} - p_k \leq 2 \sqrt{k \log p_k} \) (which gives the claim for \( x \in [p_k, p_{k+1}] \)) when \( k \leq 2000 \).

Second claim. Since

\[ h^{-1} \sum_{|p-x|<h} \log p \leq \frac{1+O\left( \frac{\log x}{\sqrt{x}} \right)}{c \sqrt{x}} \sum_{|p-x|<h} 1, \]

from (A.3) we also get that

\[ \sum_{|p-x|<h} 1 \geq \left( c - \alpha + (2 \alpha + o(1)) \frac{\log x}{x} \right) \sqrt{x}. \]

In particular, setting \( \alpha = \alpha_{\text{min}} \) and \( c = \alpha_{\text{min}} + 1 \) this shows that

\[ \sum_{|p-x|<\left( \alpha_{\text{min}} + 1 \right) \sqrt{x} \log x} 1 \geq \sqrt{x}, \]

when \( x \) is large enough. Once again, choosing these values directly in (A.3) one gets an explicit inequality which can be proved for \( x \geq 1500 \), proving the statement in this range. The claim for \( x \in [2, 16000] \) may be checked directly by noticing that \( p_{n+\left\lceil \sqrt{p_n} \right\rceil} - p_n \leq 2(c+1) \sqrt{p_n} \log p_n \) (giving the claim for \( x \in [p_n, p_{n+\left\lceil \sqrt{p_n} \right\rceil}] \) for \( n = 1, \ldots, 251 \).

Acknowledgements. All computations have been done using PARI/GP [9].

References

[1] D. Bazzanella, Primes between consecutive squares, Arch. Math. (Basel) 75 (2000), no. 1, 29–34.
[2] H. Cramér, Some theorems concerning prime numbers, Arkiv för Math. Astr. Fys. 15 (1920), no. 5, 1–17;
Collected works I, 138–170, Springer, Berlin–Heidelberg, 1994.
[3] A. W. Dudek, On the Riemann Hypothesis and the difference between primes, Int. J. Number Theory (2014), 1–8, [10.1142/S1793042115500426]
[4] L. Grenié and G. Molteni, Explicit smoothed prime ideals theorems under GRH, to appear in Math. Comp.,
http://arxiv.org/abs/1312.7465
[5] D. R. Heath-Brown and D. A. Goldston, A note on the differences between consecutive primes, Math. Ann. 266 (1984), no. 3, 317–320.
[6] A. E. Ingham, *The distribution of prime numbers*, Cambridge University Press, Cambridge, 1990.

[7] A. Languasco, A. Perelli, and Zaccagnini A., *An extension of the pair-correlation conjecture and applications*, to appear in Math. Res. Lett., http://arxiv.org/abs/1308.3934, 2014.

[8] T. Oliveira e Silva, S. Herzog and S. Pardi, *Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$*, Math. Comp. 83 (288) (2014), 2033–2060.

[9] The PARI Group, Bordeaux, *PARI/GP*, version 2.6.0, 2013, available from http://pari.math.u-bordeaux.fr/.

[10] D. J. Platt and T. S. Trudgian, *On the first sign change of $\theta(x) - x$*, to appear in Math. Comp., 2015.

[11] P. Ribenboim, *The new book of prime number records*, Springer-Verlag, New York, 1996.

[12] J. B. Rosser and L. Schoenfeld, *Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$*, Math. Comp. 29 (1975), 243–269.

[13] S. Skewes, *On the difference $\pi(x) - li(x)$, II*, Proc. London Math. Soc. (3) 5 (1955), 48–70.

[14] T. S. Trudgian, *An improved upper bound for the argument of the Riemann zeta-function on the critical line II*, J. Number Theory 134 (2014), 280–292.

(A. Dudek) Mathematical Sciences Institute, The Australian National University
E-mail address: adrian.dudek@anu.edu.au

(L. Grenié) Dipartimento di Ingegneria gestionale, dell’informazione e della produzione, Università di Bergamo, viale Marconi 5, 24044 Dalmine (BG) Italy
E-mail address: loic.grenie@gmail.com

(G. Molteni) Dipartimento di Matematica, Università di Milano, via Saldini 50, 20133 Milano, Italy
E-mail address: giuseppe.molteni1@unimi.it