Primes between consecutive squares and the Lindelof hypothesis

(Article begins on next page)

PRIMES BETWEEN CONSECUTIVE SQUARES AND THE LINDELÖF HYPOTHESIS

DANILO BAZZANELLA

ABSTRACT. At present it is not known an unconditional proof that between two consecutive squares there is always a prime number. In a previous paper the author proved that, under the assumption of the Lindelöf hypothesis, each of the intervals $[n^2, (n+1)^2] \subset [1, N]$, with at most $O(N^\varepsilon)$ exceptions, contains the expected number of primes, for every constant $\varepsilon > 0$. In this paper we improve the result by weakening the hypothesis in two different ways.

This is the authors' post-print version of an article published on *Period. Math. Hungar.* 66, n. 1 (2013), 111–117, DOI:10.1007/s10998-013-1457-y.[1]

1. INTRODUCTION

A well known conjecture about the distribution of primes asserts that for every positive integer n, the interval $[n^2, (n+1)^2]$ contains at least one prime. The proof of this conjecture is quite out of reach at present, even under the assumption of the Riemann Hypothesis. Anyway it is not difficult to prove unconditionally that the conjecture holds for almost all positive integers n. More precisely, we can prove immediately that almost all intervals of the type $[n^2, (n+1)^2]$ contain the expected number of primes.

In a previous paper the author proved that, under the assumption of the Lindelöf hypothesis, each of the intervals $[n^2, (n+1)^2] \subset [1, N]$, with at most $O(N^\varepsilon)$ exceptions, contains the expected number of primes, for every constant $\varepsilon > 0$, see Theorem 2.1 of D. Bazzanella [3].

1991 *Mathematics Subject Classification.* 11NO5.

Key words and phrases. distribution of prime numbers, primes between squares.

In this paper we prove the same result assuming in turn two different heuristic hypotheses. It must be stressed that both hypotheses are implied by the Lindelöf hypothesis.

The first new hypothesis is a weakened version of the hypothesis stated in D. Bazzanella [2].

Hypothesis 1. *There exist a constant X_0 and a function $\Delta(y,T)$ such that, for every $5/12 < \beta < 1/2$ and $\varepsilon > 0$, we have*

$$\int_X^{2X} |\psi(y+y/T) - \psi(y) - y/T + \Delta(y,T)|^{2k} \mathrm{d}y \ll X^{2k+\varepsilon} T^{1-2k}$$

and

$$\Delta(y,T) \ll y/(T\log y)$$

for at least one integer $k \geq 1$, uniformly for $X \geq X_0$, $X^{5/12} \leq T \leq X^{\beta}$ and $X \leq y \leq 2X$.

To state the second new hypothesis we need to use the counting functions $N(\sigma,T)$ and $N^{(k)}(\sigma,T)$. The former is defined as the number of zeros $\rho = \beta + i\gamma$ of the Riemann zeta function which satisfy $\sigma \leq \beta \leq 1$ and $|\gamma| \leq T$, while $N^{(k)}(\sigma,T)$ is defined as the number of ordered sets of zeros $\rho_j = \beta_j + i\gamma_j$ $(1 \leq j \leq 2k)$, each counted by $N(\sigma,T)$, for which $|\gamma_1 + \cdots + \gamma_k - \gamma_{k+1} - \cdots - \gamma_{2k}| \leq 1$.

We start to observe that D. Bazzanella and A. Perelli [4] made the heuristic assumption that there exists a constant T_0 such that

$$N^{(2)}(\sigma,T) \ll \frac{N(\sigma,T)^4}{T} T^{\varepsilon} \tag{1}$$

for every $T \geq T_0$ and arbitrarily small $\varepsilon > 0$, which is close to being the best possible, in view of the trivial estimate

$$N^{(2)}(\sigma,T) \gg \frac{N(\sigma,T)^4}{T}.$$

The above may be generalized and weakened to

$$N^{(k)}(\sigma,T) \ll \frac{N(\sigma,T)^{2k}}{T} T^{\varepsilon} \qquad (1/2 \leq \sigma \leq \overline{\sigma}),$$

with suitable $\overline{\sigma} < 1$ and arbitrarily small $\varepsilon > 0$. We now observe that the Lindelöf hypothesis implies that for every $\eta > 0$ we have

$$N(\sigma,T) \ll T^{2(1-\sigma)+\eta} \quad (1/2 \leq \sigma \leq 1),$$

see A. E. Ingham [8], and then we are led to claim the following.

Hypothesis 2. *For every $0 \leq \eta < 1/6$ there exists an integer $k \geq 2$ such that*

$$N^{(k)}(\sigma,T) \ll T^{4k(1-\sigma)-1+\eta} \qquad (1/2 \leq \sigma \leq 5/6+\eta).$$

We note that Hypothesis 1 and 2 are weaker than the Lindelöf hypothesis, see G. Yu [11, Lemma B] and D. R. Heath-Brown [7, Lemma 1] respectively.

We are now able to state our main theorems.

Theorem 1. *Let $\varepsilon > 0$ be arbitrarily small and assume Hypothesis 2. Then each of the intervals $[n^2, (n+1)^2] \subset [1, N]$, with at most $O(N^\varepsilon)$ exceptions, contains the expected number of primes.*

Theorem 2. *Let $\varepsilon > 0$ be arbitrarily small and assume Hypothesis 3. Then each of the intervals $[n^2, (n+1)^2] \subset [1, N]$, with at most $O(N^\varepsilon)$ exceptions, contains the expected number of primes.*

Note that despite Hypotheses 1 and 2 are implied by the Lindelöf hypothesis, see G. Yu [11, Lemma B] and D. R. Heath-Brown [7, Lemma 1] respectively, we obtain the same expected distribution of primes between consecutive squares and consequently two theorems are stronger than Theorem 2.1 of [3].

2. Definitions and fundamental lemma

We will always assume that n, x and N are sufficiently large as prescribed by the various statements, and $\varepsilon > 0$ is arbitrarily small and not necessarily the same at each occurrence. The constants implied by the "O" and "$\ll$" symbols may depend on k. As in [4] we define a set related to the asymptotic formula

$$\psi(x + h(x)) - \psi(x) \sim h(x) \qquad (x \to \infty) \tag{2}$$

as

$$E_\delta(N, h) = \{N \le x \le 2N : |\psi(x + h(x)) - \psi(x) - h(x)| \ge \delta h(x)\},$$

where $h(x)$ is an increasing function such that $x^\varepsilon \le h(x) \le x$ for some $\varepsilon > 0$. It is clear that (2) holds if and only if for every $\delta > 0$ there exists $N_0(\delta)$ such that $E_\delta(N, h) = \emptyset$ for every $N \ge N_0(\delta)$. Hence for small $\delta > 0$, N tending to ∞ and with a function $h(x)$ which is suitably small with respect to x, the set $E_\delta(N, h)$ contains the exceptions, if any, to the expected asymptotic formula for the number of primes in short intervals.

Moreover we define a set related to the asymptotic formula

$$\psi((n+1)^2) - \psi(n^2) \sim 2n \qquad (n \to \infty) \tag{3}$$

as

$$A_\delta(N) = \{\sqrt{N} \le n \le \sqrt{2N} : |\psi((n+1)^2) - \psi(n^2) - (2n+1)| \ge \delta(2n+1)\},$$

that contains the exceptions, if any, to the expected asymptotic formula for the number of primes in intervals of the type $[n^2,(n+1)^2]\subset[N,2N]$. The main tool of the proofs is the following lemma.

Lemma. *For $h(x)=2\sqrt{x}+1$ and every $\delta>0$ we have*

$$|A_\delta(N)|\ll_\delta \frac{|E_{\delta/2}(N,h)|}{\sqrt{N}}+1.$$

The fundamental lemma is due to the author, see [3, Lemma 2].

3. Proof of the theorems

Let $h(x)=2\sqrt{x}+1$ and let $y\in E_\delta(N,h)$. Then we get

$$|\psi(y+h(y))-\psi(y)-h(y)|\gg\sqrt{N}. \tag{4}$$

We divide the interval $[N,2N]$ into $O(\ln^2 N)$ subintervals $J_i=[a_i,a_{i+1}]$, with

$$a_i=N+\frac{iN}{\log^2 N} \tag{5}$$

and define

$$E^i_\delta(N,h)=E_\delta(N,h)\cap J_i.$$

We let

$$T_i=\sqrt{a_i}/2. \tag{6}$$

Hypothesis 1 implies that there exist an integer $k\ge 1$, a constant X_0 and a function $\Delta(y,T)$ such that, for every i, we have

$$\int_N^{2N}|\psi(y+y/T_i)-\psi(y)-y/T_i+\Delta(y,T_i)|^{2k}\mathrm{d}y\ll N^{2k+\varepsilon}T_i^{1-2k} \tag{7}$$

and

$$\Delta(y,T_i)\ll y/(T_i\log y), \tag{8}$$

uniformly for $N\ge X_0$ and $N\le y\le 2N$. From the Brun–Titchmarsh theorem, see H. L. Montgomery and R. C. Vaughan [10], we can deduce that for every i we have

$$\psi(y+h(y))-\psi(y)-h(y)=\psi(y+y/T_i)-\psi(y)-y/T_i+\Delta(y,T_i)+O\left(\frac{\sqrt{N}}{\log N}\right),$$

for every $y\in J_i$. The above bound and (4) imply that

$$|\psi(y+y/T_i)-\psi(y)-y/T_i+\Delta(y,T_i)|\gg\sqrt{N},$$

for every $y \in E^i_\delta(N,h)$. Thus we obtain

$$|E_\delta(N,h)| \ll N^{-k} \sum_i \int\limits_{E^i_\delta(N,h)} |\psi(y+y/T_i) - \psi(y) - y/T_i + \Delta(y,T_i)|^{2k}\,\mathrm{d}y$$
$$\ll N^{-k} \sum_i \int_N^{2N} |\psi(y+y/T_i) - \psi(y) - y/T_i + \Delta(y,T_i)|^{2k}\,\mathrm{d}y.$$

By (7) we conclude that

$$|E_\delta(N,h)| \ll N^{-k} \sum_i N^{2k+\varepsilon} T_i^{1-2k} \ll N^{1/2+\varepsilon}. \tag{9}$$

By the lemma and (9), we can conclude that

$$|A_\delta(N)| \ll_\delta \frac{|E_{\delta/2}(N,h)|}{\sqrt{N}} + 1 \ll N^\varepsilon,$$

for every $\delta > 0$, and this complete the proof of Theorem 1.

To prove Theorem 2 we use the classical explicit formula, see H. Davenport [5, Chapter 17], to write

$$\psi(y+y/T_i) - \psi(y) - y/T_i = -\sum_{|\gamma|\le R_i} y^\rho \, \frac{e^{\delta_i\rho}-1}{\rho} + O\left(\frac{N\log^2 N}{R_i}\right), \tag{10}$$

uniformly for $N \le y \le 2N$, where $\delta_i = \log(1+T_i^{-1})$, $10 \le R_i \le N$ and $\rho = \beta + i\gamma$ runs over the non-trivial zeros of $\zeta(s)$. If we choose $R_i = T_i \log^3 N$ and recall (5) and (6) then we have

$$\sqrt{N}\log^3 N \ll R_i \ll \sqrt{N}\log^3 N$$

and

$$\psi(y+y/T_i) - \psi(y) - y/T_i = -\sum_{|\gamma|\le R_i} y^\rho \, \frac{e^{\delta_i\rho}-1}{\rho} + O\left(\frac{\sqrt{N}}{\log N}\right)$$

for every i and $y \in J_i$. As before we observe that for every $y \in J_i$ we have

$$\psi(y+h(y)) - \psi(y) - h(y) = \psi(y+y/T_i) - \psi(y) - y/T_i + O\left(\frac{\sqrt{N}}{\log N}\right)$$

and then

$$\psi(y+h(y)) - \psi(y) - h(y) = -\sum_{|\gamma|\le R_i} y^\rho \, \frac{e^{\delta_i\rho}-1}{\rho} + O\left(\frac{\sqrt{N}}{\log N}\right),$$

for every i and $y \in J_i$. This implies that

$$|E_\delta(N,h)|N^k \ll \sum_i \int_N^{2N} \left| \sum_{|\gamma|\le R_i} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k} \mathrm{d}x. \tag{11}$$

To estimate the $2k$-power integral we divide the interval $[0,1]$ into $O(\ln N)$ subintervals I_j of the form

$$I_j = \left[\frac{j}{\log N}, \frac{j+1}{\log N} \right].$$

By Hölder inequality we obtain

$$\left| \sum_{|\gamma|\le R_i} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k} = \left| \sum_j \sum_{\substack{|\gamma|\le R_i \\ \beta\in I_j}} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k}$$

$$\ll (\ln N)^{2k-1} \sum_j \left| \sum_{\substack{|\gamma|\le R_i \\ \beta\in I_j}} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k}.$$

Following the method of D. R. Heath-Brown, we write

$$(\ln N)^{1-2k} \int_N^{2N} \left| \sum_{|\gamma|\le R_i} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k} \mathrm{d}x \ll$$

$$\sum_j \sum_{\substack{\beta_1,\dots,\beta_{2k}\in I_j \\ |\gamma_1|\le R_i,\dots,|\gamma_{2k}|\le R_i}} \frac{(2N)^{\rho_1+\dots+\rho_k+\overline{\rho_{k+1}}+\dots+\overline{\rho_{2k}}+1} - N^{\rho_1+\rho_2+\dots+\rho_k+\overline{\rho_{k+1}}+\dots+\overline{\rho_{2k}}+1}}{\rho_1\dots\rho_{2k}\,(\rho_1+\dots+\rho_k+\overline{\rho_{k+1}}+\dots+\overline{\rho_{2k}}+1)}$$

$$\times (e^{\delta_i\rho_1}-1)\cdots(e^{\delta_i\rho_k}-1)(e^{\delta_i\overline{\rho_{k+1}}}-1)\dots(e^{\delta_i\overline{\rho_{2k}}}-1)$$

$$\ll \sum_j \frac{1}{T_i^{2k}} N^{1+2kj/\log N} \sum_{\substack{\beta_1,\dots,\beta_{2k}\ge j/\log N \\ |\gamma_1|\le R_i,\dots,|\gamma_{2k}|\le R_i}} \frac{1}{|\rho_1+\dots+\rho_k+\overline{\rho_{k+1}}+\dots+\overline{\rho_{2k}}+1|}.$$

This implies

$$\int_N^{2N} \left| \sum_{|\gamma|\le R_i} x^\rho \, \frac{e^{\delta_i \rho}-1}{\rho} \right|^{2k} \mathrm{d}x \ll \frac{1}{T_i^{2k}} \max_\sigma N^{2k\sigma+1+\varepsilon} M_k(\sigma, R_i), \tag{12}$$

where

$$M_k(\sigma, R_i) = \sum_{\substack{\beta_1,\dots,\beta_{2k}\geq\sigma \\ |\gamma_1|\leq R_i,\dots,|\gamma_{2k}|\leq R_i}} \frac{1}{1+|\gamma_1+\cdots+\gamma_k-\gamma_{k+1}-\cdots-\gamma_{2k}|}$$

and

$$M_k(\sigma, R_i) \ll N^{(k)}(\sigma, R_i)\log N, \tag{13}$$

see [9, p. 336]. From (11), (12) and (13) we have

$$|E_\delta(N,h)| \ll N^{1-2k+\varepsilon}\sum_i \max_\sigma N^{2k\sigma}N^{(k)}(\sigma, R_i). \tag{14}$$

By Hypothesis 2 and (14) we get

$$|E_\delta(N,h)| \ll N^{1-2k+\varepsilon}\sum_i \max_\sigma N^{2k\sigma}R_i^{4k(1-\sigma)-1} \ll N^{1/2+\varepsilon}.$$

Again by the lemma we can conclude that

$$|A_\delta(N)| \ll_\delta \frac{|E_{\delta/2}(N,h)|}{\sqrt{N}} + 1 \ll N^\varepsilon,$$

for every $\delta > 0$, and this completes the proof of Theorem 2.

References

[1] D. Bazzanella, *Primes between consecutive squares*, Arch. Math. **75** (2000), 29–34.

[2] D. Bazzanella, *A note on primes in short intervals*, Arch. Math. **91** (2008), 131–135.

[3] D. Bazzanella, *Some conditional results on primes between consecutive squares*, to appear in Funct. Approx. Comment. Math.

[4] D. Bazzanella and A. Perelli, *The exceptional set for the number of primes in short intervals*, J. Number Theory **80** (2000), 109–124 .

[5] H. Davenport, *Multiplicative Number Theory*, Second edition, Graduate Texts in Mathematics 74. Springer–Verlag, New York (1980).

[6] D. R. Heath-Brown. *The difference between consecutive primes II*, J. London Math. Soc. (2), **19** (1979), 207–220.

[7] D. R. Heath-Brown, *The difference between consecutive primes IV*, A tribute to Paul Erdős, 277–287, Cambridge Univ. Press, Cambridge (1990).

[8] A. E. Ingham, *On the difference between consecutive primes*, Quart. J. of Math. (Oxford) **8** (1937), 255–266.

[9] A. Ivić, *The Riemann Zeta-Function* , John Wiley & Sons, New York (1985).

[10] H. L. Montgomery and R. C. Vaughan, *The large sieve* , Mathematika **20** (1973), 119–134.

[11] G. Yu, *The differences between consecutive primes*, Bull. London Math. Soc. **28** (1996), no. 3, 242–248.

Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino - Italy
E-mail address: danilo.bazzanella@polito.it