Mathematische Annalen

The sequence of prime gaps is graphic

Péter L. Erdős[1] · Gergely Harcos[1] · Shubha R. Kharel[2] · Péter Maga[1] ·
Tamás Róbert Mezei[1] · Zoltán Toroczkai[2]

Abstract

Let us call a simple graph on $n \geq 2$ vertices a prime gap graph if its vertex degrees are
1 and the first $n-1$ prime gaps. We show that such a graph exists for every large n, and
in fact for every $n \geq 2$ if we assume the Riemann hypothesis. Moreover, an infinite
sequence of prime gap graphs can be generated by the so-called degree preserving
growth process. This is the first time a naturally occurring infinite sequence of positive
integers is identified as graphic. That is, we show the existence of an interesting, and
so far unique, infinite combinatorial object.

1 Introduction

1.1 The problem

This paper grew out from an empirical observation by one of us (Z.T.): there are large
graphs whose vertex degrees are consecutive members of the sequence of prime gaps.
Moreover, such graphs can be generated recursively by the so-called degree preserving
growth process [16]. To turn the observation into precise mathematical statements, we
introduce the following definition.

Definition Let p_n denote the n-th prime number, and let $p_0 = 1$. We call a simple graph
on $n \geq 2$ vertices a *prime gap graph* if its vertex degrees are $p_1 - p_0, \ldots, p_n - p_{n-1}$.

Dedicated to János Pintz on the occasion of his 71st and 73rd birthdays.

✉ Tamás Róbert Mezei
  mezei.tamas.robert@renyi.hu

Extended author information available on the last page of the article

Conjecture 1.1 *(Toroczkai, 2016) For every $n \geq 2$, there exists a prime gap graph on n vertices.*

Conjecture 1.2 *(Toroczkai, 2016) In every prime gap graph on n vertices, there exist $(p_{n+1} - p_n)/2$ independent edges.*

In fact, as will become clear in the next subsection, Conjecture 1.2 implies Conjecture 1.1. By combining techniques from analytic number theory and matching theory, we are able to almost fully settle these conjectures.

Theorem 1.1 *Conjectures 1.1 and 1.2 are true for every sufficiently large n. Assuming the Riemann hypothesis, they are true for every $n \geq 2$.*

The main input from analytic number theory is an upper bound on the sum of large prime gaps. The main input from matching theory is Vizing's theorem on edge colorings. See Subsection 1.3 for more details. In the next subsection, we discuss the background and motivation for Theorem 1.1.

1.2 Broader context

Networks are powerful, graph-based representations used in the study of complex systems. They appear in systems ranging from elementary particle interactions, through nucleosynthesis, chemistry, biology (gene interactions, protein interactions, metabolism, physiology), social sciences (human interactions), infrastructures (transportation, power grid, etc.), ecology (food webs) and climate, to the organization of visible and dark matter in the universe. In this paper we report on a novel family of networks, however, in number theory.

In this paper all graphs are *simple*: there are no parallel edges and loops. The most common characteristic of a graph is its *degree sequence*: we equip each of the n vertices with a unique label from $\{1, \ldots, n\}$, and an integer vector $\mathbf{D} = (d_1, \ldots, d_n)$ lists the *degrees* of the corresponding labeled vertices, that is, the number of edges incident on a given vertex. In chemistry, and in old-fashioned graph theory, this is called *valency*.

The inverse problem is the following: we are given a sequence $\mathbf{D}$ of nonnegative integers, and we want to know whether there exists a graph with this ensemble as a degree sequence. When the answer is affirmative, then we call the sequence $\mathbf{D}$ *graphic*. Clearly, if the sequence is graphic, then the sum of its members must be even. However, it is not self-evident whether a given sequence is graphic. The most well-known characterization of graphic degree sequences is the following theorem:

Theorem 1.2 (Erdős–Gallai [7]) *Let $d_1 \geq \cdots \geq d_n \geq 0$ be integers. Then the sequence $(d_1, \ldots, d_n)$ is graphic if and only if $d_1 + \cdots + d_n$ is even and for every $k \in \{1, \ldots, n\}$ we have*

$$\sum_{\ell=1}^{k} d_\ell \leq k(k-1) + \sum_{\ell=k+1}^{n} \min(k, d_\ell) . \tag{1}$$

It is well-understood that there are exponentially many different *realizations* for almost every graphic degree sequence. At the same time, the number of all graphic degree sequences is infinitesimal compared to the number of integer partitions of the sum of the degrees. More precisely, let $2m$ denote the sum of the degrees (as usual), so that the degree sequence is a *partition* of $2m$. By a difficult result of Pittel [23], as m tends to infinity, the probability of a random partition of $2m$ being graphic is zero in the limit.

Theorem 1.2 gives us a means to decide whether the degree sequence is graphic. It is, however, an entirely different problem to actually construct a realization of a graphic degree sequence. The simplest way to do that is via the Havel–Hakimi algorithm, which, in turn, is based on the following observation:

Theorem 1.3 (Havel [12] and Hakimi [11]) *Let $d_1 > 0$ and $d_2 \geq \cdots \geq d_n \geq 0$ be integers. Then $(d_1, \ldots, d_n)$ is graphic if and only if $(d_2 - 1, \ldots, d_{d_1+1} - 1, d_{d_1+2}, \ldots, d_n)$ is graphic.*

Assume we are given a long sequence of natural numbers $\mathbf{D}$ such that every *initial segment* $\mathbf{D}^n$ formed by the first n elements of $\mathbf{D}$ is graphic (we restrict this notion to $n \geq 2$). This means, in particular, that $d_n < n$ for $n \geq 2$. We would like to construct a realization G_n for any $n \geq 2$. Clearly, it can be done by separate applications of the Havel–Hakimi algorithm for every single $n \geq 2$. But this is rather uneconomical: in principle, for each new segment we have to restart the algorithm from scratch. Instead, we want to find a *graph growth dynamics (GGD)* such that G_{n+1} can be obtained from G_n rapidly.

There are only a few GGDs in the network science literature as most graph construction models are based on static algorithms. Arguably, the most popular GGDs is the preferential attachment algorithm of scale-free networks. However, in this and other GGDs, typically, some of the degrees of the vertices in G_{n+1} are bigger than in the degree sequence of G_n, and thus, they are unsuitable for our purposes.

Recently, a new network growth dynamics has been introduced: the *degree-preserving network growth (DPG)* model family (see [8] or [16]). The DPG-process can be described as follows: let G be a simple graph with degree sequence $\mathbf{D}$. In what follows, by a *matching* we mean a set of independent edges in the graph, that is, a set of pairwise non-adjacent edges. In a general step, a new vertex u joins G by removing a ν-element matching of G followed by connecting u to the vertices incident to the ν removed edges. The degree of the newly inserted vertex is $d = 2\nu$. This step does not join two vertices of G that are non-adjacent, furthermore, the degrees of vertices in G are not changed. The degree sequence of the newly generated graph is $\mathbf{D} \circ d$, that is, d is concatenated to the end of $\mathbf{D}$. This graph operation is called a *degree-preserving step (DP-step)*, and the *DPG-process* repeats such steps iteratively.

Returning to our long sequence $\mathbf{D}$ of natural numbers: if, for each step n, we can find a matching of size $d_{n+1}/2$ in G_n, then the application of a DP-step provides a realization G_{n+1} of $\mathbf{D}^{n+1}$. In this case, we say that the pair (G_n, d_{n+1}) is *DPG-graphic*. One can ask, why should a suitable matching be found in G_n? Actually, this is not inconceivable, as the following theorem shows:

Theorem 1.4 (Theorem 2.5 in [18]) *Given a graphic sequence* **D** *of length n and an even integer* $2 \leq d \leq n$, *the sequence* $\mathbf{D} \circ d$ *is graphic if and only if* **D** *has a realization with a matching of size* $d/2$.

Since every initial segment $\mathbf{D}^n$ is graphic, therefore, for each one there exists a "special" realization G'_n with the requested matching. However, it is not automatic that $G'_n = G_n$. A natural way to deal with this problem is to add the condition that every realization of $\mathbf{D}^n$ has a matching of size $d_{n+1}/2$.

We stress that it is not easy to find an infinite, naturally occurring sequence **D** whose initial segments are all graphic, or at least graphic beyond a certain point. As a matter of fact, until now we have only known one such example: when all elements in the sequence are equal. Such a GGD describes an ever growing regular graph sequence.

In this paper, we describe for the first time a nontrivial, naturally arising infinite sequence whose initial segments are all graphic. Furthermore, we show that any realization of the initial segments is admissible for the DPG-algorithm. This sequence is the sequence of prime gaps with a prefix 1:

$$\mathbf{PD} := (p_1 - p_0, p_2 - p_1, \dots) = (1, 1, 2, 2, 4, 2, \dots)$$

The prefix 1 was included to guarantee that the sum of the initial segments is even. The figure below is an illustration of the DPG-process on prime gap graphs. The independent edges used by the DP-steps are red zigzags.
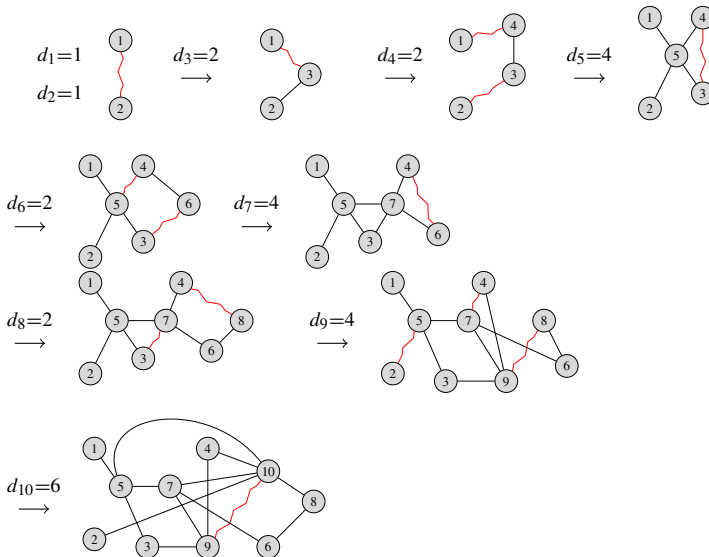
The proof that the DPG-process creates an infinite sequence of prime gap graphs incorporates two main ingredients. The first one is a symmetric inequality, which implies the Erdős–Gallai inequalities, and thus provides a practical sufficient condition for the graphicality of the underlying degree sequence. Moreover, another new

inequality implies the DPG-graphicality of a long sequence of natural numbers. The second ingredient combines classical L^2 and L^∞ bounds for prime gaps.

Theorem 1.5 *If n is sufficiently large, then the initial segment* $\mathbf{PD}^n$ *is graphic, and for any realization* G_n *of* $\mathbf{PD}^n$, *the pair* $(G_n, p_{n+1} - p_n)$ *is DPG-graphic.*

Ultimately, the proof relies on the rarity of zeros possibly violating the Riemann hypothesis. In principle, it allows one to deduce an effective constant beyond which Theorem 1.5 holds true, but this constant far exceeds the capabilities of computers. However, assuming the Riemann hypothesis, we can reduce the constant significantly and prove the result for all $n \geq 2$. Here the numerics are quite delicate, and for efficiency we depart from the symmetric treatment alluded to above. Instead, we go back to first principles and examine the contribution of large prime gaps more directly. We still need to rely on computational results, but they can either be obtained with very simple computer programs, or found in the literature (e.g. prime gap records until $2 \cdot 10^{18}$).

Theorem 1.6 *Assume the Riemann hypothesis. Then, for any* $n \geq 2$, *the initial segment* $\mathbf{PD}^n$ *is graphic, and for any realization* G_n *of* $\mathbf{PD}^n$, *the pair* $(G_n, p_{n+1} - p_n)$ *is DPG-graphic.*

Note that Theorems 1.5 and 1.6 are a reformulation of Theorem 1.1 in the terminology of combinatorics and network theory. They rely on the core theorems presented in the next subsection, which are of independent interest.

1.3 New results

Our first result provides, via two symmetric inequalities, sufficient conditions for a given sequence to be graphic and that in every graph realization of the sequence, there is a matching of a given size.

Theorem 1.7 *Let* $\mathbf{D} = (d_1, \ldots, d_n)$ *be a sequence of positive integers such that* $\|\mathbf{D}\|_1 = \sum_{\ell=1}^{n} d_\ell$ *is even. Let* $1 < p \leq \infty$ *be a parameter.*
    Part (a). *Assume that the following* L^p-*norm bound holds:*

$$\|2 + \mathbf{D}\|_p \leq n^{\frac{1}{2} + \frac{1}{2p}}. \tag{2}$$

Then there is a simple graph G with degree sequence $\mathbf{D}$.
    Part (b). *Let G be any simple graph with degree sequence* $\mathbf{D}$. *Assume that* $d \geq 2$ *is an even integer satisfying*

$$4d^{1 - \frac{1}{p}} \|\mathbf{D}\|_p \leq \|\mathbf{D}\|_1. \tag{3}$$

Then the pair (G, d) *is DPG-graphic, and consequently* $\mathbf{D} \circ d$ *is graphic.*

Our second result makes explicit a theorem of Selberg [25, Th. 2].

Theorem 1.8 *Assume the Riemann hypothesis. Then, for any $x \geq 2$ and $N > 0$, we have*

$$\sum_{\substack{x \leq p_\ell \leq 2x \\ p_{\ell+1} - p_\ell \geq N}} (p_{\ell+1} - p_\ell) < \frac{163x \log^2 x}{N}. \tag{4}$$

Remark The example $x = 2$ and $N = 2$ shows that this result would become false if we replaced the constant 163 by 4. On the other hand, Cramér's model predicts that it can be replaced by $o(1)$ for $x \to \infty$ (cf. [9, §1.1]).

In order to achieve the good numeric constant 163, we estimate carefully (among others) the error term in the truncated von Mangoldt formula for the Chebyshev psi function. This result, stated below, makes explicit a theorem of Goldston [10], and simultaneously extends and sharpens a theorem of Dudek [6, Th. 1.3] in the special case relevant for us. Here and later the notation $A = O^*(B)$ stands for $|A| \leq B$.

Theorem 1.9 *For any $z > x > 10^{18}$ we have*

$$\psi(x) = x - \sum_{|\Im\rho| < z} \frac{x^\rho}{\rho} + O^*\big(5 \log x \log \log x\big),$$

where the sum is over the nontrivial zeros of the Riemann zeta function (counted with multiplicity).

2 Preliminary results

2.1 An application of Vizing's theorem

Theorem 2.1 (Vizing [27]) *A simple graph with maximal degree Δ admits a proper edge coloring with $\Delta + 1$ colors.*

Lemma 2.1 *Let G be a simple graph on n vertices with degrees $d_1, \ldots, d_n$. Let $\delta \geq 1$ be an integer, and let $d \geq 2$ be an even integer satisfying*

$$\delta d \leq \sum_{d_\ell < \delta} d_\ell - \sum_{d_\ell \geq \delta} d_\ell. \tag{5}$$

Then G has a matching of size $d/2$.

Proof Let us delete all vertices of degree at least δ (and the incident edges) from G. The remaining subgraph H has maximal degree less than δ, and number of edges at least

$$\frac{1}{2} \sum_{\ell=1}^{n} d_\ell - \sum_{d_\ell \geq \delta} d_\ell = \frac{1}{2} \left(\sum_{d_\ell < \delta} d_\ell - \sum_{d_\ell \geq \delta} d_\ell \right) \geq \frac{\delta d}{2} \tag{6}$$

by (5). It follows from Theorem 2.1 that the edge set of H can be partitioned into δ matchings, and then (6) shows that the largest matching in this decomposition must be of size at least $d/2$. Since H is a subgraph of G, the proof is complete. □

2.2 Preliminaries about $\Gamma(z)$ and $\zeta(s)$

Lemma 2.2 *Assume that $\Re z > 0$. Then*

$$\Re \frac{\Gamma'(z)}{\Gamma(z)} + \frac{\Gamma'(\Re z)}{\Gamma(\Re z)} < \log|z| + \log \Re z - \Re \frac{1}{2z} - \frac{1}{2\Re z}.$$

Proof With the help of the well-known integral representation (cf. [28, §12.31])

$$\frac{\Gamma'(z)}{\Gamma(z)} = \log z - \frac{1}{2z} - \int_0^\infty \left(\frac{1}{2} - \frac{1}{t} + \frac{1}{e^t - 1} \right) e^{-tz}\, dt, \qquad \Re z > 0, \qquad (7)$$

the statement becomes

$$\int_0^\infty \left(\frac{1}{2} - \frac{1}{t} + \frac{1}{e^t - 1} \right) \left(\Re e^{-tz} + e^{-t\Re z} \right) dt > 0.$$

However, this one is clear, because the integrand is non-negative with a discrete set of zeros (there are no zeros when z is real). □

Lemma 2.3 (Delange [5]) *For any $\sigma > 1$ and $t \in \mathbb{R}$ we have*

$$\left| \frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} \right| < \frac{1}{\sigma - 1} - \frac{1}{2\sigma^2}.$$

Lemma 2.4 (Dudek [6]) *Let $\sigma \leq -1$ and $t \in \mathbb{R}$. Assume that either $\sigma \in 1 + 2\mathbb{Z}$ or $|t| \geq 1$. Then*

$$\left| \frac{\zeta'(\sigma + it)}{\zeta(\sigma + it)} \right| < 9 + \log|\sigma + it|.$$

Proof This is a variant of [6, Lem. 2.3], and can be proved in the same way. We note a small oversight in [6, p. 183]: instead of assuming that $U \geq 2$ is an even integer, one should assume that $U \geq 1$ is an odd integer, just as in [4, §17]. □

Lemma 2.5 (Dudek [6]) *Assume that $z > 100$. Then there exists $T \in (z - 2, z)$ such that*

$$\left| \frac{\zeta'(\sigma + iT)}{\zeta(\sigma + iT)} \right| < \log^2 z + 20 \log z, \qquad \sigma > -1. \qquad (8)$$

Proof The statement follows from [6, Lem. 2.8]. □

Definition For $T > 0$, we denote by $N(T)$ the number of zeros of $\zeta(s)$ with imaginary part in $(0, T)$, counted with multiplicity.

Lemma 2.6 *For any $T \geq \Delta + 2\pi > \Delta > 0$ we have*

$$N(T) - N(T - \Delta) < \left(\frac{\Delta}{2\pi} + 0.56\right) \log T.$$

Proof By [2, Cor. 1], we have

$$N(T) - N(T - \Delta) = \frac{1}{2\pi} \int_{T-\Delta}^{T} \log \frac{t}{2\pi} \, dt + O^*(0.56 \log T).$$

The integrand is less than $\log T$, and the result follows. □

2.3 Preliminaries about prime gaps

Theorem 2.2 (Ingham [15, Th. 4]) *Let $\varepsilon > 0$. For any $x \geq x_0(\varepsilon)$, there is a prime number in $[x, x + x^{5/8+\varepsilon}]$.*

Remark The exponent $5/8 + \varepsilon$ was improved multiple times, the current record $21/40$ being due to Baker–Harman–Pintz [1]. We have emphasized the classical result of Ingham [15] as it suffices for our purposes.

Theorem 2.3 (Carneiro–Milinovich–Soundararajan [3, Th. 1.5]) *Assume the Riemann hypothesis. Then, for any $x \geq 4$, there is a prime number in $[x, x + \frac{22}{25}\sqrt{x} \log x]$.*

In a restricted range, we have a stronger unconditional result thanks to explicit calculations.

Lemma 2.7 *For any $x \in [117, 10^{18}]$, there is a prime number in $[x, x + \sqrt{x}]$.*

Proof Assume that the conclusion fails for some $x \in [117, 10^{18}]$. Then there is a unique prime number p_ℓ such that $p_\ell < x < x + \sqrt{x} < p_{\ell+1}$. In particular, $\ell \geq 31$ and $p_\ell < x < (p_{\ell+1} - p_\ell)^2$. Hence the computations of Oliveira e Silva, Herzog, and Pardi [21, Table 8] show that the initial upper bound 10^{18} for p_ℓ successively improves to: $1442^2, 148^2, 52^2, 34^2, 22^2, 14^2$. This means that $31 \leq \ell \leq 44$, but then $x < (p_{\ell+1} - p_\ell)^2 \leq 10^2$ is a contradiction. □

Remark The conclusion of Lemma 2.7 is likely true for all $x \geq 117$. However, this statement is not known to follow from the Riemann hypothesis, and it is stronger than Oppermann's conjecture (which itself implies Legendre's conjecture, Andrica's conjecture, and Brocard's conjecture).

Theorem 2.4 (Heath-Brown [13]) *For any $x \geq 2$ we have*

$$\sum_{p_\ell \leq x} (p_{\ell+1} - p_\ell)^2 \ll x^{4/3}(\log x)^{10000}.$$

Remark The exponent 4/3 was improved to $23/18 + \varepsilon$ by Heath-Brown [14], and to $5/4 + \varepsilon$ independently by Peck [22] and Maynard [17]. The current record $123/100 + \varepsilon$ is due to Stadlmann [26]. We have emphasized the original breakthrough of Heath-Brown [13] as it suffices for our purposes.

3 Proof of the main theorem

In this section, we first prove Theorem 1.5 assuming Theorem 1.7, and then we prove Theorem 1.6 assuming Theorem 1.8. In other words, we deduce Theorem 1.1 from Theorems 1.7 and 1.8.

3.1 Proof of Theorem 1.5

Let n be sufficiently large. We shall verify the conditions of Theorem 1.7 for

$$p := 2, \qquad d_\ell := p_\ell - p_{\ell-1}, \qquad d := p_{n+1} - p_n.$$

Clearly, $\|\mathbf{D}\|_1 = p_n - 1$ is even. Condition (2) reads

$$\sum_{\ell=1}^{n} (2 + p_\ell - p_{\ell-1})^2 \le n^{3/2}, \tag{9}$$

while condition (3) reads

$$16(p_{n+1} - p_n) \sum_{\ell=1}^{n} (p_\ell - p_{\ell-1})^2 \le (p_n - 1)^2. \tag{10}$$

Now (9) and (10) follow from Theorems 2.2 and 2.4, hence we are done:

$$\sum_{\ell=1}^{n} (2 + p_\ell - p_{\ell-1})^2 \le 9 \sum_{\ell=1}^{n} (p_\ell - p_{\ell-1})^2 \le p_n^{4/3+o(1)} = n^{4/3+o(1)},$$

$$16(p_{n+1} - p_n) \sum_{\ell=1}^{n} (p_\ell - p_{\ell-1})^2 \le p_n^{5/8+o(1)} p_n^{4/3+o(1)} = p_n^{47/24+o(1)}.$$

3.2 Proof of Theorem 1.6

Assume the Riemann hypothesis, and let G be a prime gap graph on n vertices. It suffices to show that G has $(p_{n+1} - p_n)/2$ independent edges (cf. Conjecture 1.2), because then a straightforward induction argument based on Theorem 1.4 shows that every initial segment of $\mathbf{PD}$ is graphic (cf. Conjecture 1.1). The statement is clear for $2 \le n \le 4$, hence we shall restrict to $n \ge 5$.

By Lemma 2.1, it suffices to exhibit an integer $N \geq 1$ satisfying

$$N(p_{n+1} - p_n) + 2 \sum_{\substack{1 \leq \ell \leq n \\ p_\ell - p_{\ell-1} \geq N}} (p_\ell - p_{\ell-1}) < p_n. \tag{11}$$

For $p_n < 10^{18}$ we take

$$N := \max_{1 \leq \ell \leq n} (1 + p_\ell - p_{\ell-1}),$$

so that (11) simplifies to $N(p_{n+1} - p_n) < p_n$. In fact the proof of Lemma 2.1 reveals that the last condition can be relaxed to

$$\frac{p_{n+1} - p_n}{2} \leq \left\lceil \frac{p_n - 1}{2N} \right\rceil, \tag{12}$$

which works better for very small $n \geq 5$. For $5 \leq n \leq 44$ the condition (12) can be checked by a simple computer program (or by hand). For $n \geq 45$ and $p_n < 10^{18}$ we verify (12) as follows. Let k be the unique positive integer satisfying $(k-1)^2 < p_n < k^2$. Note that $k \geq 15$, because $p_n \geq p_{45} = 197$. From Lemma 2.7 it follows that $p_{n+1} - p_n \leq k - 1$ and

$$N = \max \left(15, \max_{32 \leq \ell \leq n} (1 + p_\ell - p_{\ell-1}) \right) \leq k,$$

hence also that

$$\frac{p_n - 1}{2N} > \frac{k^2 - 2k}{2N} \geq \frac{k^2 - 2k}{2k} = \frac{k}{2} - 1.$$

Therefore, (12) is clear by

$$\frac{p_{n+1} - p_n}{2} \leq \frac{k-1}{2} \leq \left\lceil \frac{p_n - 1}{2N} \right\rceil.$$

For $p_n > 10^{18}$ we take

$$N := \left\lceil \frac{\sqrt{p_n}}{3 \log p_n} \right\rceil.$$

Then, by Theorems 2.3 and 1.8, we have

$$N(p_{n+1} - p_n) < \frac{p_n}{3} \quad \text{and} \quad \sum_{\substack{1 \leq \ell \leq n \\ p_\ell - p_{\ell-1} \geq N}} (p_\ell - p_{\ell-1}) < 489 \sqrt{p_n} \log^3 p_n < \frac{p_n}{3}.$$

From here the bound (11) is immediate, hence we are done.

4 A symmetric criterion for graphicality

In this section, we prove Theorem 1.7.

Part (a). By symmetry, we can assume that $d_1 \geq \cdots \geq d_n$. By Theorem 1.2, it suffices to check that for any $1 \leq k \leq n$,

$$\sum_{\ell=1}^{k} d_\ell \leq k(k-1) + \sum_{\ell=k+1}^{n} \min(k, d_\ell).$$

Since $d_\ell \geq 1$ for any $1 \leq \ell \leq n$, it suffices to prove that

$$\sum_{\ell=1}^{k} d_\ell \leq k(k-1) + (n-k),$$

which is equivalent to

$$\sum_{\ell=1}^{k} (2 + d_\ell) \leq k^2 + n.$$

This last condition follows from (2) and Hölder's inequality, hence we are done:

$$\|2 + \mathbf{D}^k\|_1 \leq k^{1-\frac{1}{p}} \|2 + \mathbf{D}^k\|_p \leq k^{1-\frac{1}{p}} \|2 + \mathbf{D}\|_p \leq k^{1-\frac{1}{p}} n^{\frac{1}{2}+\frac{1}{2p}} \leq \max(k^2, n).$$

In the last step, we used that both k^2 and n are upper bounded by $\max(k^2, n)$.

Part (b). By Theorem 1.4 and Lemma 2.1, it suffices to verify that (5) holds for some integer $\delta \geq 1$. If $p = \infty$, then (3) says that $4d\|\mathbf{D}\|_\infty \leq \|\mathbf{D}\|_1$, hence (5) holds for $\delta := 1 + \|\mathbf{D}\|_\infty$. So let us focus on the case $1 < p < \infty$. For any integer $\delta \geq 1$, we have

$$\|\mathbf{D}\|_p^p \geq \sum_{d_\ell \geq \delta} d_\ell^p \geq \delta^{p-1} \sum_{d_\ell \geq \delta} d_\ell,$$

hence also

$$\sum_{d_\ell < \delta} d_\ell - \sum_{d_\ell \geq \delta} d_\ell \geq \|\mathbf{D}\|_1 - 2\delta^{1-p}\|\mathbf{D}\|_p^p.$$

So for the validity of (5), it suffices that

$$\delta^{1-p}\|\mathbf{D}\|_p^p \leq \frac{1}{4}\|\mathbf{D}\|_1 \quad \text{and} \quad \delta d \leq \frac{1}{2}\|\mathbf{D}\|_1.$$

In other words, it suffices to find an integer δ satisfying

$$\left(\frac{4\|\mathbf{D}\|_p^p}{\|\mathbf{D}\|_1}\right)^{\frac{1}{p-1}} \leq \delta \leq \frac{1}{2d}\|\mathbf{D}\|_1.$$

The left-hand side exceeds 1, hence δ exists as long as

$$2\left(\frac{4\|\mathbf{D}\|_p^p}{\|\mathbf{D}\|_1}\right)^{\frac{1}{p-1}} \leq \frac{1}{2d}\|\mathbf{D}\|_1.$$

This is equivalent to condition (3), hence the proof of Theorem 1.7 is complete.

5 The sum of large prime gaps

In this section, we prove Theorem 1.8 assuming Theorem 1.9. Throughout, we assume the Riemann hypothesis.

First we eliminate some simple cases. Let N^* denote the largest prime gap $p_{\ell+1} - p_\ell$ occurring in (4). Then we can clearly assume that

$$N \leq N^* \quad \text{and} \quad 2x + N^* > \frac{163x \log^2 x}{N}, \tag{13}$$

hence also that $N^*(2x + N^*) > 163x \log^2 x$. From Theorem 2.3 it follows that $N^* < 3\sqrt{x} \log x$, so our previous inequality yields $N^* > 77 \log^2 x$. By [21, Table 8], this forces $x > 10^{18}$. Indeed, for $x \in [2, 10^3]$ we have $N^* \leq 34$, while for $x \in [10^3, 10^{18}]$ we have $N^* \leq 1476$. On the other hand, for $x > 10^{18}$ we get from Therore 2.3 that $N^* < \frac{4}{3}\sqrt{x} \log x < 0.001x$, hence by (13) also that

$$81 \log^2 x < N < \frac{4}{3}\sqrt{x} \log x. \tag{14}$$

From now on we assume both $x > 10^{18}$ and (14). Following Heath-Brown [13], we write $N = 4\delta x$ with

$$\frac{81 \log^2 x}{4x} < \delta < \frac{\log x}{3\sqrt{x}},$$

and we set out to estimate the square mean of

$$E(y, \delta) := \psi(y + \delta y) - \psi(y) - \delta y, \quad x \leq y \leq 2x.$$

It follows from Theorem 1.9 and the crude bound $y + \delta y < 3x$ that

$$|E(y, \delta)| < \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right| + \log^2 x,$$

where

$$C(\rho) := \frac{1 - (1 + \delta)^\rho}{\rho}.$$

As a result,

$$\int_x^{2x} |E(y, \delta)|^2 \, dy < 2 \int_x^{2x} \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right|^2 dy + 2x \log^4 x. \tag{15}$$

Lemma 5.1 *We have*

$$|C(\rho)| < \min\left(\delta, \frac{\sqrt{4 + 4\delta}}{|\rho|} \right). \tag{16}$$

Proof The bound $|C(\rho)| < \delta$ is a consequence of the integral representation

$$C(\rho) = \int_{1+\delta}^1 x^{\rho-1} \, dx$$

and the triangle inequality for complex-valued Riemann integrals. In addition, the triangle inequality for complex numbers yields by the definition of $C(\rho)$ that

$$|C(\rho)| \leq \frac{1 + \sqrt{1 + \delta}}{|\rho|} < \frac{\sqrt{4 + 4\delta}}{|\rho|}.$$

$\square$

We estimate the integral on the right-hand side of (15) as in the proof of [24, Lem. 5]:

$$\int_x^{2x} \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right|^2 dy \leq \int_1^2 \int_{xv/2}^{2xv} \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right|^2 dy \, dv$$

$$= \sum_{|\Im\rho|,|\Im\rho'|<3x} x^{2+\rho-\rho'} C(\rho)\overline{C(\rho')} \cdot \frac{2^{2+\rho-\rho'} - 2^{-2-\rho+\rho'}}{2 + \rho - \rho'} \cdot \frac{2^{3+\rho-\rho'} - 1}{3 + \rho - \rho'}$$

$$\leq x^2 \sum_{|\Im\rho|,|\Im\rho'|<3x} |C(\rho)C(\rho')| \left| \frac{2^2 + 2^{-2}}{2 + \rho - \rho'} \right| \left| \frac{2^3 + 1}{3 + \rho - \rho'} \right|$$

$$\leq x^2 \sum_{|\Im\rho|,|\Im\rho'|<3x} |C(\rho)|^2 \left|\frac{2^2+2^{-2}}{2+\rho-\rho'}\right| \left|\frac{2^3+1}{3+\rho-\rho'}\right|.$$

Here the contribution of $\Im\rho < 0$ is the same as the contribution of $\Im\rho > 0$. Therefore, applying Lemma 5.1 along with the elementary inequality

$$|2+it| \cdot |3+it| \geq 6+t^2, \qquad t \in \mathbb{R},$$

we arrive at

$$\int_x^{2x} \left|\sum_{|\Im\rho|<3x} y^\rho C(\rho)\right|^2 dy < \frac{153}{2}x^2 \sum_{0<\Im\rho<3x} \min\left(\delta^2, \frac{4+4\delta}{|\rho|^2}\right) \sum_{|\Im\rho'|<3x} \frac{1}{6+|\rho-\rho'|^2}. \tag{17}$$

We need to estimate the inner sum in (17). The idea is to drop the condition $|\Im\rho'| < 3x$, and consider the full convergent series

$$\sum_{\rho'} \frac{1}{6+|\rho-\rho'|^2} = \frac{1}{\sqrt{6}} \sum_{\rho'} \Re\frac{1}{\sqrt{6}+\rho-\rho'} = \frac{1}{\sqrt{6}} \Re\frac{\xi'(\sqrt{6}+\rho)}{\xi(\sqrt{6}+\rho)}.$$

The first equality follows from the Riemann hypothesis, while the second equality follows from [20, Cor. 10.14]. Let us write $s := \sqrt{6} + \rho$ for simplicity. Then, as in the proof of [20, Cor. 10.14], we have

$$\frac{\xi'(s)}{\xi(s)} = -\frac{1}{2}\log\pi + \frac{1}{s-1} + \frac{\zeta'(s)}{\zeta(s)} + \frac{1}{2}\frac{\Gamma'(s/2+1)}{\Gamma(s/2+1)}.$$

We take the real part of both sides, and apply Lemma 2.2:

$$\Re\frac{\xi'(s)}{\xi(s)} < -\frac{1}{2}\log\pi + \Re\frac{1}{s-1} - \frac{\zeta'(\Re s)}{\zeta(\Re s)} + \frac{1}{2}\Re\frac{\Gamma'(s/2+1)}{\Gamma(s/2+1)}$$

$$< -\frac{1}{2}\log\pi + \Re\frac{1}{s-1} - \Re\frac{1}{2s+4} - \frac{\zeta'(\Re s)}{\zeta(\Re s)} - \frac{1}{2\Re s+4}$$

$$- \frac{1}{2}\frac{\Gamma'(\Re s/2+1)}{\Gamma(\Re s/2+1)} + \frac{1}{2}\log(\Re s/2+1) + \frac{1}{2}\log|s/2+1|.$$

Using that $\Re s = \sqrt{6} + 1/2$ and $\Im s > 14$, it is straightforward to check that $\Re\frac{1}{s-1} < \Re\frac{1}{2s+4}$, hence in fact

$$\Re\frac{\xi'(s)}{\xi(s)} < 0.181 - \frac{1}{2}\log\pi + \frac{1}{2}\log|s/2+1| < \frac{1}{4} + \frac{1}{2}\log\frac{\Im\rho}{2\pi}.$$

To sum up, we have proved that

$$\sum_{\rho'} \frac{1}{6 + |\rho - \rho'|^2} < \frac{1}{2\sqrt{6}} \left(\frac{1}{2} + \log \frac{\Im\rho}{2\pi} \right).$$

Going back to (17), we conclude that

$$\int_x^{2x} \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right|^2 dy < 15.616 x^2 \sum_{0<\Im\rho<3x} \min\left(\delta^2, \frac{4}{|\rho|^2} \right) \left(\frac{1}{2} + \log \frac{\Im\rho}{2\pi} \right).$$

In the sum on the right-hand side, we drop the condition $\Im\rho < 3x$ and replace $|\rho|$ by $\Im\rho$. By [2, Cor. 1, Lem. 5–6], the resulting bigger sum can be estimated as follows:

$$< \delta^2 N(2/\delta) \left(\frac{1}{2} + \log \frac{1}{\delta\pi} \right) + \sum_{\Im\rho \geq 2/\delta} \left(\frac{2}{(\Im\rho)^2} + \frac{4}{(\Im\rho)^2} \log \frac{\Im\rho}{2\pi} \right)$$

$$< \frac{\delta}{\pi} \left(\log \frac{1}{\delta\pi} \right) \left(\frac{1}{2} + \log \frac{1}{\delta\pi} \right) + \frac{\delta}{2\pi} \log \frac{2}{\delta} + \frac{\delta}{\pi} \left(\log^2 \frac{2}{\delta} - \log \frac{2}{\delta} \right)$$

$$< \frac{\delta}{\pi} \left(\log^2 \frac{1}{\delta\pi} + \log^2 \frac{2}{\delta} \right) < \frac{2\delta}{\pi} \log^2 x.$$

In the end, we get

$$\int_x^{2x} \left| \sum_{|\Im\rho|<3x} y^\rho C(\rho) \right|^2 dy < 9.942 \delta x^2 \log^2 x.$$

Plugging this bound into (15), we conclude that

$$\int_x^{2x} |E(y, \delta)|^2 dy < 19.884 \delta x^2 \log^2 x + 2x \log^4 x < 19.983 \delta x^2 \log^2 x. \quad (18)$$

Assume now that the prime $p_\ell \in [x, 2x]$ satisfies $p_{\ell+1} - p_\ell \geq N$. There is at most one p_ℓ such that $(p_\ell + p_{\ell+1})/2 > 2x$, so assume also that $(p_\ell + p_{\ell+1})/2 \leq 2x$. Then, for any

$$y \in (p_\ell, (p_\ell + p_{\ell+1})/2) \subset (x, 2x),$$

the interval

$$[y, y + \delta y] \subset [y, y + N/2] \subset (p_\ell, p_{\ell+1})$$

is free of primes, hence counting the possible higher prime powers in this interval, we get

$$\psi(y + \delta y) - \psi(y) \le (1 + \delta\sqrt{y}/2)\log_2(y + \delta y) < (2 + \delta\sqrt{y})\log x < 0.003\delta y.$$

That is, $|E(y, \delta)| > 0.997\delta y$ holds on $(p_\ell, (p_\ell + p_{\ell+1})/2)$. Squaring and integrating, we get

$$\int_{p_\ell}^{(p_\ell + p_{\ell+1})/2} |E(y, \delta)|^2 \, dy > 0.497\delta^2 x^2(p_{\ell+1} - p_\ell).$$

Summing over all such primes p_ℓ, and using (18) as well as Theorem 2.3 for the possible single exceptional p_ℓ, we obtain (4):

$$\sum_{\substack{x \le p_\ell \le 2x \\ p_{\ell+1} - p_\ell \ge N}} (p_{\ell+1} - p_\ell) < \frac{4}{3}\sqrt{x}\log x + \frac{1}{0.497\delta^2 x^2}\int_x^{2x} |E(y, \delta)|^2 \, dy < \frac{163x\log^2 x}{N}.$$

6 The error term in the truncated von Mangoldt formula

In this section, we prove Theorem 1.9. We follow Davenport [4, §17] and Goldston [10] with appropriate modifications.

We assume first that $x \notin \mathbb{Z}$. We choose $T \in (z - 2, z)$ according to Lemma 2.5, and we also fix $c := 1 + 1/\log x$. We record the following approximation to the characteristic function of $(1, \infty)$:

$$\mathbf{1}_{y>1} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds + O^*\left(y^c \min\left(0.501, \frac{1}{\pi T|\log y|}\right)\right), \qquad y \in (0, 1) \cup (1, \infty).$$

This formula follows by making explicit the calculation on [4, pp. 105–106]. The constant 0.501 follows by observing that the line $\Re s = c$ divides the circle $|s| = |c + iT|$ into two almost equal arcs, each of length less than $1.001\pi|c + iT|$. The constant $1/\pi$ arises as twice the size of the leading $1/(2\pi i)$. Applying the formula for $y = x/n$, multiplying by $\Lambda(n)$, and summing over $n \ge 1$, we get

$$\psi(x) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s}\, ds + O^*\left(\sum_{n=1}^{\infty} \left(\frac{x}{n}\right)^c \Lambda(n) \min\left(0.501, \frac{1}{\pi T\left|\log\frac{x}{n}\right|}\right)\right). \tag{19}$$

We shall abbreviate the integrand in (19) by $f(s)$, and estimate the error term by cutting the n-sum into four parts. Throughout, we keep in mind that $x/T < x/(x-2)$. As a preparation, we record the elementary inequalities

$$\left(\frac{x}{n}\right)^c \Lambda(n) \le \left(\frac{x}{n}\right)^c \log n = \frac{x}{n} \cdot \frac{e\log n}{e^{\log n/\log x}} \le \frac{x}{n}\log x, \tag{20}$$

$$\left| \log \frac{x}{y} \right| = \int_{\min(x,y)}^{\max(x,y)} \frac{du}{u} \geq \frac{|x - y|}{\max(x, y)}, \qquad y > 0.$$

(21)

We also observe that the function

$$y \mapsto \frac{x - y}{y \log \frac{x}{y}}, \qquad y > 0,$$

is positive and decreasing (the function has a removable discontinuity at $y = x$). Indeed, writing $v := \log \frac{x}{y}$, the claim is that $v \mapsto (e^v - 1)/v$ is positive and increasing on $\mathbb{R}$, which in turn follows from the fact that the exponential function is increasing and convex.

First we consider the n's satisfying $1 \leq |x - n| \leq \log x$. By (21) and the subsequent observation, in this range we have

$$0 < \frac{x - n}{n \log \frac{x}{n}} \leq \frac{\log x}{(x - \log x) \log \frac{x}{x - \log x}} \leq \frac{x}{x - \log x} < 1.001,$$

hence by (20) also

$$\left(\frac{x}{n} \right)^c \frac{\Lambda(n)}{\left| \log \frac{x}{n} \right|} \leq \frac{x \log x}{n \left| \log \frac{x}{n} \right|} < 1.001 \frac{x \log x}{|x - n|}.$$

(22)

So the corresponding n-subsum within (19) is at most

$$1.001 \frac{x \log x}{\pi T} \sum_{\substack{1 \leq |x-n| \leq \log x \\ \Lambda(n) \neq 0}} \frac{1}{|x - n|} < 0.638 \log x \sum_{1 \leq k \leq \log x} \frac{1}{k} < A(x),$$

(23)

where

$$A(x) := 0.638 \log x \cdot (\log \log x + 3/5).$$

Second, we consider the n's satisfying $\log x < |x - n| \leq x/5$. In this range we have the following variant of (22) proved in the same way:

$$\left(\frac{x}{n} \right)^c \frac{\Lambda(n)}{\left| \log \frac{x}{n} \right|} \leq \frac{x \log x}{n \left| \log \frac{x}{n} \right|} \leq \frac{1}{4 \log \frac{5}{4}} \cdot \frac{x \log x}{|x - n|}.$$

So the corresponding n-subsum within (19) is at most

$$\frac{1}{4 \log \frac{5}{4}} \cdot \frac{x \log x}{\pi T} \sum_{\substack{\log x < |x-n| \leq x \\ \Lambda(n) \neq 0}} \frac{1}{|x - n|},$$

(24)

where we relaxed the summation condition for convenience. If $p(u)$ denotes the number of prime powers in $[x - u, x + u]$, then the last sum can be written as

$$\sum_{\substack{\log x < |x-n| \leq x \\ \Lambda(n) \neq 0}} \frac{1}{|x-n|} = \int_{\log x}^{x} \frac{dp(u)}{u} = \left[\frac{p(u)}{u} \right]_{\log x}^{x} + \int_{\log x}^{x} \frac{p(u)}{u^2} \, du. \quad (25)$$

Since $0 \leq p(u) \leq 2u + 1$, the first term on the right-hand side is less than 2.001. By the Brun–Titchmarsh inequality in the form given by Montgomery and Vaughan [19, Th. 2], we also see that

$$p(u) < \frac{4u}{\log u} + \log_2(x) \left(1 + \frac{2u}{\sqrt{2x}} \right), \qquad u \in [2, x].$$

Indeed, on the right-hand side, the first term upper bounds the number of primes in $[x - u, x + u]$, while the second term upper bounds the number of higher prime powers in $[x - u, x + u]$. In particular,

$$\int_{\log x}^{x} \frac{p(u)}{u^2} \, du < \int_{\log x}^{x} \left(\frac{4.001}{u \log u} + \frac{\log_2 x}{u^2} \right) du < 4.001 \log \log x - 3.818. \quad (26)$$

We infer from (24)–(26) that the n's satisfying $\log x < |x - n| \leq x/5$ contribute to (19) less than

$$B(x) := 1.427 \log x \cdot (\log \log x - 2/5).$$

Now we turn to the n's satisfying $|x - n| > x/5$. In this range we have $|\log(x/n)| > \log(6/5)$, hence by Lemma 2.3 the corresponding n-subsum is at most

$$\sum_{|x-n| > x/5} \Lambda(n) \left(\frac{x}{n} \right)^c \frac{1}{\pi T \log(6/5)} < \frac{ex}{\pi T \log(6/5)} \left| \frac{\zeta'(c)}{\zeta(c)} \right| < 4.746 \log x.$$

Finally, there are two n's satisfying $|x - n| < 1$, and by (20) their contribution to (19) is less than

$$0.501 \left(\frac{x}{x-1} + \frac{x}{x} \right) \log x < 1.003 \log x.$$

Collecting everything, we arrive at (with room to spare)

$$\psi(x) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} f + O^*(A(x) + B(x) + 6 \log x). \quad (27)$$

We note here that $A(x) + B(x)$ is less than $2.1 \log x \log \log x$.

On the other hand, the residue theorem combined with (8) and Lemma 2.4 shows that

$$\frac{1}{2\pi i} \left(\int_{c-iT}^{c+iT} f + \int_{c+iT}^{-\infty+iT} f + \int_{-\infty-iT}^{c-iT} f \right)$$
$$= x - \sum_{|\Im\rho|<T} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2}\log(1 - x^{-2}),$$

where each integral is over a directed line segment or half-line. We estimate the second integral with the help of (8) and Lemma 2.4:

$$\left| \int_{c+iT}^{-\infty+iT} f \right| < \frac{\log^2 z + 20\log z}{z-2} \int_{-\infty}^{c} x^\sigma \, d\sigma < (\log x + 20)\frac{ex}{x-2}.$$

The third integral obeys the same bound, hence we infer that

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} f = x - \sum_{|\Im\rho|<T} \frac{x^\rho}{\rho} + O^*(\log x + 20).$$

By Lemma 2.6, we can extend the ρ-sum to $|\Im\rho| < z$ at the cost of an error of $O^*(2\log x)$. Therefore, going back to (27), we conclude for $x \notin \mathbb{Z}$ that

$$\psi(x) = x - \sum_{|\Im\rho|<z} \frac{x^\rho}{\rho} + O^*(A(x) + B(x) + 9\log x + 20).$$

Finally, if x is an integer, then we make use of the following simple observation. For a fixed $z > 10^{18}$, the ρ-sum on the right-hand side is continuous in $x \in (10^{18}, z)$, while the left-hand side equals $\psi(x-) + \Lambda(x)$. Therefore, the previous formula is valid at x with an extra error term of $O^*(\log x)$. We finish the proof of Theorem 1.9 by noting that

$$A(x) + B(x) + 10\log x + 20 < 5\log x \log\log x, \qquad x > 10^{18}.$$

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Baker, R.C., Harman, G., Pintz, J.: The difference between consecutive primes II. Proc. London Math. Soc. **83**, 532–562 (2001). https://doi.org/10.1112/plms/83.3.532
2. Brent, R.P., Platt, D.J., Trudgian, T.S.: The mean square of the error term in the prime number theorem. J. Number Theory **238**, 740–762 (2022). https://doi.org/10.1016/j.jnt.2021.09.016
3. Carneiro, E., Milinovich, M.B., Soundararajan, K.: Fourier optimization and prime gaps. Comment. Math. Helv. **94**, 533–568 (2019). https://doi.org/10.4171/CMH/467
4. Davenport, H.: Multiplicative number theory, 3rd ed., Revised and with a preface by H. L. Montgomery, Graduate Texts in Mathematics, Vol. 74, Springer-Verlag, New York, (2000). ISBN:978-0387950976
5. Delange, H.: Une remarque sur la dérivée logarithmique de la fonction zêta de Riemann [A remark on the logarithmic derivative of the Riemann zeta function]. Colloq. Math. (in French) **53**, 333–335 (1987). https://doi.org/10.4064/cm-53-2-333-335
6. Dudek, A.W.: An explicit result for primes between cubes. Funct. Approx. Comment. Math. **55**, 177–197 (2016). https://doi.org/10.7169/facm/2016.55.2.3
7. Erdős, P., Gallai, T.: Gráfok előírt fokú pontokkal (Graphs with vertices of prescribed degrees). Mat. Lapok **11**, 264–274 (1960). (**in Hungarian**)
8. Erdős, P. L., Kharel, S., Mezei, T. R., Toroczkai, Z.: Degree preserving graph dynamics – a versatile process to construct random networks, submitted, arXiv:2111.11994 (2021), 21 pages
9. Ford, K., Green, B., Konyagin, S., Maynard, J., Tao, T.: Long gaps between primes. J. Amer. Math. Soc. **31**, 65–105 (2017). https://doi.org/10.1090/jams/876
10. Goldston, D.A.: On a result of Littlewood concerning prime numbers II. Acta Arith. **43**, 49–51 (1983)
11. Hakimi, S.L.: On realizability of a set of integers as degrees of the vertices of a linear graph I. J. Soc. Indust. Appl. Math. **10**, 496–506 (1962)
12. Havel, V.: Poznámka o existenci konečných grafů (A remark on the existence of finite graphs). Časopis Pěst. Mat. **80**, 477–480 (1955). (**in Czech**)
13. Heath-Brown, D.R.: The differences between consecutive primes. J. London Math. Soc. **18**, 7–13 (1978). https://doi.org/10.1112/jlms/s2-18.1.7
14. Heath-Brown, D.R.: The differences between consecutive primes III. J. London Math. Soc. **20**, 177–178 (1979). https://doi.org/10.1112/jlms/s2-20.2.177
15. Ingham, A.E.: On the difference between consecutive primes. Quart. J. Math. Oxford Ser. **8**, 255–266 (1937). https://doi.org/10.1093/qmath/os-8.1.255
16. Kharel, S., Mezei, T.R., Chung, S., Erdős, P.L., Toroczkai, Z.: Degree-preserving network growth. Nat. Phys. **18**, 100–106 (2022). https://doi.org/10.1038/s41567-021-01417-7
17. Maynard, J.: On the difference between consecutive primes, arXiv:1201.1787 (2012), 41 pages
18. Mezei, T. R., Kharel, S., Erdős, P. L., Toroczkai, Z.: On graphic degree sequences and matching numbers, submitted, arXiv:2204.07423 (2022), 12 pages
19. Montgomery, H.L., Vaughan, R.C.: The large sieve. Mathematika **20**, 119–134 (1973). https://doi.org/10.1112/S0025579300004708
20. Montgomery, H. L., Vaughan, R. C.: Multiplicative number theory. I. Classical theory, Cambridge Studies in Advanced Mathematics, Vol. 97, Cambridge University Press, Cambridge, 2007. https://doi.org/10.1017/CBO9780511618314
21. Oliveira e Silva, T., Herzog, S., Pardi, S.: Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. Math. Comp. **83** (2014), 2033–2060. https://doi.org/10.1090/S0025-5718-2013-02787-1
22. Peck, A. S.: On the differences between consecutive primes, PhD thesis, University of Oxford, (1996)

23. Pittel, B.: Confirming two conjectures about the integer partitions. J. Combin. Theory Ser. A **88**, 123–135 (1999). https://doi.org/10.1006/jcta.1999.2986
24. Saffari, B., Vaughan, R.C.: On the fractional parts of x/n and related sequences II. Ann. Inst. Fourier (Grenoble) **27**, 1–30 (1977). https://doi.org/10.5802/aif.649
25. Selberg, A.: On the normal density of primes in small intervals, and the difference between consecutive primes. Arch. Math. Naturvid. **47**, 87–105 (1943)
26. Stadlmann, J.: On the mean square gap between primes, arXiv:2212.10867 (2022), 71 pages
27. Vizing, V.G.: On an estimate of the chromatic class of a p-graph. Diskret. Analiz. **3**, 25–30 (1964)
28. Whittaker, E.T., Watson, G.N.: A course of modern analysis, 5th edn. Cambridge University Press, Cambridge (2021). https://doi.org/10.1017/9781009004091

Authors and Affiliations

Péter L. Erdős[1] · **Gergely Harcos[1]** · **Shubha R. Kharel[2]** · **Péter Maga[1]** · **Tamás Róbert Mezei[1]** · **Zoltán Toroczkai[2]**

Péter L. Erdős
erdos.peter@renyi.hu

Gergely Harcos
harcos.gergely@renyi.hu

Shubha R. Kharel
skharel@nd.edu

Péter Maga
maga.peter@renyi.hu

Zoltán Toroczkai
toro@nd.edu

[1] Alfréd Rényi Institute of Mathematics (LERN), Reáltanoda utca 13–15, H-1053 Budapest, Hungary

[2] Department of Physics, University of Notre Dame, 225 Nieuwland Science Hall, Notre Dame IN 46556, USA