

THE CONVERGENCE OF AN ALTERNATING SERIES OF ERDŐS, ASSUMING THE HARDY–LITTLEWOOD PRIME TUPLES CONJECTURE

TERENCE TAO

ABSTRACT. It is an open question of Erdős as to whether the alternating series $\sum_{n=1}^{\infty} \frac{(-1)^n n}{p_n}$ is (conditionally) convergent, where p_n denotes the n th prime. By using a random sifted model of the primes recently introduced by Banks, Ford, and the author, as well as variants of a well known calculation of Gallagher, we show that the answer to this question is affirmative assuming a suitably strong version of the Hardy–Littlewood prime tuples conjecture.

1. INTRODUCTION

Let p_n denote the n th prime. This paper concerns the following question of Erdős [13, p. 203, E7] (see also [8, p. 96], [6], [7]):

Question 1.1. Does the series $\sum_{n=1}^{\infty} \frac{(-1)^n n}{p_n}$ converge?

Numerically¹, the series appears to converge (rather slowly) to about -0.052161 ; we do not expect the exact value of this sum to have any additional notable properties. As recorded in [13, p. 203, E7], Erdős noted that the variant series $\sum_{n=1}^{\infty} \frac{(-1)^n n \log n}{p_n}$ is divergent (the summands fail to converge to zero).

The difference between consecutive magnitudes $\frac{n}{p_n}, \frac{n+1}{p_{n+1}}$ of this alternating series is largely controlled by the prime gap $p_{n+1} - p_n$, so the answer to Question 1.1 revolves around how this prime gap correlates with the parity of n . If such a correlation existed, it would create a bias in the parity of the prime counting function² $\pi(x) := \sum_{p \leq x} 1$. Indeed, it is an unpublished observation of Said³ that Question 1.1 is equivalent to the following.

Question 1.2. Does the series $\sum_{n=2}^{\infty} \frac{(-1)^{\pi(n)}}{n \log n}$ converge?

For the convenience of the reader, we establish the equivalence of Question 1.1 and Question 1.2 in Section 2.

In this paper we show that the answer to both of these questions is affirmative assuming a suitably strong version of the Hardy–Littlewood prime tuples conjecture [14]. We use a (very slightly strengthened) formulation from the recent paper [15]:

Received by the editors August 14, 2023, and, in revised form, August 24, 2023, and October 24, 2023.
2020 *Mathematics Subject Classification.* Primary 11N05, 40A05.

The author was supported by NSF grant DMS-1764034 and by a Simons Investigator Award.

¹Tomás Oliveira da Silva, private communication.

²The variable p is always understood to range over primes in this paper.

³<https://mathoverflow.net/questions/313999>

Conjecture 1.3 (Quantitative Hardy–Littlewood prime tuples conjecture [15, Conjecture 1.3]). *There exist two absolute constants $\varepsilon > 0$ and $C > 0$ such that for all $x \geq 10$, all $k \leq (\log \log x)^5$, and all tuples $\mathcal{H} = \{h_1, \dots, h_k\} \subset [0, \log^2 x]$ of distinct integers h_1, \dots, h_k , one has*

$$(1.1) \quad \left| \sum_{n \leq x} 1_{\mathcal{P}}(n + h_1) \dots 1_{\mathcal{P}}(n + h_k) - \mathfrak{S}(\mathcal{H}) \int_2^x \frac{dy}{\log^k y} \right| \leq Cx^{1-\varepsilon}.$$

Here $1_{\mathcal{P}}$ is the indicator function of the primes \mathcal{P} , $\mathfrak{S}(\mathcal{H})$ is the singular series

$$\mathfrak{S}(\mathcal{H}) := \prod_p \frac{1 - \nu_{\mathcal{H}}(p)/p}{(1 - 1/p)^k},$$

and $\nu_{\mathcal{H}}(p)$ denotes the number of distinct residue classes modulo p occupied by \mathcal{H} . It is common to restrict (1.1) to the admissible case $\mathfrak{S}(\mathcal{H}) > 0$, but this is not necessary since (1.1) is very easy to establish in the non-admissible case $\mathfrak{S}(\mathcal{H}) = 0$. In [15], k was restricted to the slightly smaller range $k \leq (\log \log x)^3$, but we will need to increase the exponent 3 to 5 for technical reasons (actually any exponent larger than 4 would suffice for our arguments). On the other hand, the power savings $x^{1-\varepsilon}$ in (1.1) is significantly stronger than is needed for our arguments; a weaker bound such as $x \exp(-c \log^c x)$, or even $x \exp(-c(\log \log x)^5)$, for some constant $c > 0$, would in fact suffice for us. We remark that Conjecture 1.3 has been verified almost surely (even with the enlarged range of parameters, and with power savings in the bound) for a certain random model of the primes; see [2, Theorem 1.3]. A version of this random model will also play a key role in the arguments of this paper.

Our main result is then

Theorem 1.4. *Assume Conjecture 1.3. Then the answer to Question 1.2 (and hence Question 1.1) is affirmative.*

We prove this theorem in Section 3. After an application of the van der Corput A -process, the problem reduces to showing that the parity $(-1)^{\pi(x+\lambda \log x) - \pi(x)}$ of the number of primes in intervals of the form $(x, x + \lambda \log x]$ is fairly well equidistributed, for moderately large λ ; a typical choice of parameters is $\lambda \asymp (\log \log x)^{4.4}$. Using Conjecture 1.3, this then reduces to controlling a certain alternating sum involving the singular series $\mathfrak{S}(\mathcal{H})$ over various k -tuples $\mathcal{H} = \{h_1, \dots, h_k\}$. It is tempting to then control the contribution of each k separately and treat the error terms by the triangle inequality, in the spirit of a well known calculation of Gallagher [9] which predicts Poissonian statistics for the number of primes in $(x, x + \lambda \log x]$ (which implies in particular that the mean value of $(-1)^{\pi(x+\lambda \log x) - \pi(x)}$ converges to $e^{-2\lambda}$ in the limit $x \rightarrow \infty$ for λ fixed). Unfortunately, the range of k needed is a little bit too large to use the strongest available estimates (due to V. Kuperberg [15]) for mean values of the singular series in this fashion; we need k to be as large as $(\log \log x)^{4.5}$ or so, but quantitative estimates only currently have good error terms in the regime $k \leq (\log \log x)^{0.5}$. However, we can proceed instead by expressing these sums (up to acceptable errors) using the random sieve model \mathfrak{S}_z introduced in [2], effectively replacing the primes by this model. One then performs some probabilistic calculations on that model to ensure the required equidistribution, taking advantage of cancellations between the contributions of different values of k . The key point is that many of the sifting steps used to build the

random sieve model act to reduce any irregularities of distribution in the parity of the number of sifted elements, rather than to increase it.

We emphasize that the random sieve model is not used merely to give heuristic support to various statements about the primes, but to actually prove rigorous (albeit conditional) results about the primes themselves. With this perspective, one can interpret the prime tuples conjecture as an assertion that various statistics concerning the primes are very well approximated by the corresponding statistics of the random sieve model, so that calculations involving the latter can be used to also control the former.

We use the usual asymptotic notation $X = O(Y)$, $X \ll Y$, or $Y \gg X$ to denote the estimate $|X| \leq CY$ for an absolute constant C , and $X \asymp Y$ to denote the estimate $X \ll Y \ll X$.

2. EQUIVALENCE OF QUESTION 1.1 AND QUESTION 1.2

We now show the equivalence of Question 1.1 and Question 1.2. We will in fact show the more precise relation

$$(2.1) \quad \sum_{n \leq x} \frac{(-1)^n n}{p_n} = \frac{1}{2} \sum_{2 \leq m \leq x \log x} \frac{(-1)^{\pi(m)}}{m \log m} + C + o(1)$$

for some absolute constant C , where $o(1)$ denotes a quantity that goes to zero as $x \rightarrow \infty$, from which the equivalence of the two questions clearly follows.

Firstly, by shifting the summation variable n by 1, we have

$$\sum_{n \leq x} \frac{(-1)^n n}{p_n} = -\frac{1}{2} + \sum_{n \leq x} \frac{(-1)^{n+1}(n+1)}{p_{n+1}} + o(1).$$

Averaging the left- and right-hand sides, we conclude that

$$\sum_{n \leq x} \frac{(-1)^n n}{p_n} = -\frac{1}{4} + \frac{1}{2} \sum_{n \leq x} \left(\frac{(-1)^n n}{p_n} + \frac{(-1)^{n+1}(n+1)}{p_{n+1}} \right) + o(1).$$

One could also obtain this identity by writing $(-1)^n = \frac{1}{2}((-1)^n - (-1)^{n-1})$ and then summing by parts. We may rewrite the summand on the right-hand side using the identity

$$(2.2) \quad \frac{(-1)^n n}{p_n} + \frac{(-1)^{n+1}(n+1)}{p_{n+1}} = \frac{(-1)^n n(p_{n+1} - p_n)}{p_n p_{n+1}} - \frac{(-1)^n}{p_{n+1}}.$$

By the alternating series test we have

$$\sum_{n \leq x} \frac{(-1)^n}{p_{n+1}} = C_0 + o(1)$$

for some absolute constant C_0 . Also, from the prime number theorem and subdivision of the m variable we have

$$\sum_{2 \leq m \leq x \log x} \frac{(-1)^{\pi(m)}}{m \log m} = \sum_{n \leq x} \sum_{p_n \leq m < p_{n+1}} \frac{(-1)^{\pi(m)}}{m \log m} + o(1).$$

Thus it will suffice to show that the series

$$\sum_{n=1}^{\infty} \left(\sum_{p_n \leq m < p_{n+1}} \frac{(-1)^{\pi(m)}}{m \log m} - \frac{(-1)^n n(p_{n+1} - p_n)}{p_n p_{n+1}} \right)$$

is absolutely convergent. Since $(-1)^{\pi(m)} = (-1)^n$ for $p_n \leq m < p_{n+1}$, this is equivalent to showing that⁴

$$(2.3) \quad \sum_{n=10}^{\infty} \left| \sum_{p_n \leq m < p_{n+1}} \frac{1}{m \log m} - \frac{n(p_{n+1} - p_n)}{p_n p_{n+1}} \right| < \infty.$$

By the intermediate value theorem we have

$$\sum_{p_n \leq m < p_{n+1}} \frac{1}{m \log m} = \frac{p_{n+1} - p_n}{x_n \log x_n}$$

for some $p_n \leq x_n < p_{n+1}$. From the prime number theorem we have $p_n, p_{n+1} = n \log n \left(1 + O\left(\frac{\log \log n}{\log n}\right)\right)$ for $n \geq 10$, and thus after some calculation

$$\frac{1}{x_n \log x_n} = \frac{n}{p_n p_{n+1}} + O\left(\frac{\log \log n}{n \log^3 n}\right)$$

for $n \geq 10$. On the other hand, from summation by parts and the prime number theorem we have

$$\sum_{n=10}^{\infty} \frac{\log \log n}{n \log^3 n} (p_{n+1} - p_n) \ll 1 + \sum_{n=10}^{\infty} \frac{\log \log n}{n^2 \log^3 n} p_n < \infty,$$

and the claim follows.

Remark 2.1. A more careful analysis of the above argument shows that the $o(1)$ error in (2.1) can be taken to be $O\left(\frac{\log \log x}{\log x}\right)$ for $x \geq 10$.

3. PROOF OF MAIN THEOREM

We now establish Theorem 1.4. By summation by parts, it will suffice to establish the bound

$$(3.1) \quad \sum_{n \leq x} (-1)^{\pi(n)} \ll \frac{x}{(\log \log x)^{1.1}}$$

for all sufficiently large x , basically because the series $\sum_{n=10}^{\infty} \frac{1}{n \log n (\log \log n)^{1.1}}$ is (barely) convergent. It is convenient to work with slightly shorter intervals than $[1, x]$. By subdivision, it will suffice to show that

$$\sum_{x \leq n < x + x^{1-\varepsilon/2}} (-1)^{\pi(n)} \ll \frac{x^{1-\varepsilon/2}}{(\log \log x)^{1.1}}$$

for all sufficiently large x , where ε is the constant in Conjecture 1.3 (which we can assume to be small).

Fix a sufficiently large x . We express the above estimate probabilistically. Let \mathbf{n} ⁵ be an integer drawn uniformly at random from the interval $[x, x + x^{1-\varepsilon/2}]$, then the above estimate is equivalent to the assertion that

$$\mathbf{E}(-1)^{\pi(\mathbf{n})} \ll \frac{1}{(\log \log x)^{1.1}},$$

where \mathbf{E} denotes expectation.

⁴We choose the initial index $n = 10$ of this sum rather arbitrarily; any index for which $\log \log n$ is well-defined and positive would suffice here.

⁵We will denote random variables in boldface.

The next step is to apply the van der Corput A -process (or Weyl differencing), so that we work with the parity of the number of primes in short intervals rather than long intervals. We introduce the length scale

$$(3.2) \quad H := \lfloor (\log \log x)^{4.4} \log x \rfloor,$$

then for any $1 \leq h \leq H$ we have the approximate shift invariance

$$\sum_{x \leq n \leq x+x^{1-\varepsilon/2}} (-1)^{\pi(n+h)} = \sum_{x \leq n \leq x+x^{1-\varepsilon/2}} (-1)^{\pi(n)} + O(H)$$

which on dividing through by $x^{1-\varepsilon/2}$ soon gives

$$\mathbf{E}(-1)^{\pi(\mathbf{n})} = \mathbf{E}(-1)^{\pi(\mathbf{n}+h)} + O\left(\frac{H}{x^{1-\varepsilon/2}}\right)$$

hence on averaging in h

$$\mathbf{E}(-1)^{\pi(\mathbf{n})} = \mathbf{E} \frac{1}{H} \sum_{h=1}^H (-1)^{\pi(\mathbf{n}+h)} + O\left(\frac{H}{x^{1-\varepsilon/2}}\right).$$

Applying Cauchy–Schwarz, it thus suffices to show that

$$\mathbf{E} \left| \frac{1}{H} \sum_{h=1}^H (-1)^{\pi(\mathbf{n}+h)} \right|^2 \ll \frac{1}{(\log \log x)^{2.2}}.$$

Expanding the square and using the triangle inequality we have

$$\mathbf{E} \left| \frac{1}{H} \sum_{h=1}^H (-1)^{\pi(\mathbf{n}+h)} \right|^2 \ll \frac{1}{H^2} \sum_{1 \leq h, h' \leq H} |\mathbf{E}(-1)^{\pi(\mathbf{n}+h)} - (-1)^{\pi(\mathbf{n}+h')}|.$$

By symmetry we may restrict the supremum to the regime $h \geq h'$, and then by approximate shift invariance we have

$$\mathbf{E}(-1)^{\pi(\mathbf{n}+h)} - (-1)^{\pi(\mathbf{n}+h')} = \mathbf{E}(-1)^{\pi(\mathbf{n}+h-h')} - (-1)^{\pi(\mathbf{n})} + O\left(\frac{H}{x^{1-\varepsilon/2}}\right).$$

The claim would therefore follow from the estimate

$$\sum_{0 \leq h \leq H} |\mathbf{E}(-1)^{\pi(\mathbf{n}+h)-\pi(\mathbf{n})}| \ll \frac{H}{(\log \log x)^{2.2}}.$$

To show this, it suffices from the choice (3.2) of H to establish the estimate

$$(3.3) \quad \mathbf{E}(-1)^{\pi(\mathbf{n}+\lambda \log x)-\pi(\mathbf{n})} \ll \frac{1}{\sqrt{\lambda}}$$

for any $1 \ll \lambda \ll (\log \log x)^{4.4}$ (using the trivial bound $|\mathbf{E}(-1)^{\pi(\mathbf{n}+h)-\pi(\mathbf{n})}| \leq 1$ to treat the cases where $h \leq \log x$). The $\frac{1}{\sqrt{\lambda}}$ bound here is probably not optimal (heuristically this expression should be more like $e^{-2\lambda}$, as predicted both by the Cramér random model and the calculations of Gallagher [9]), but is what arises from our (slightly crude) methods, and suffices for our purposes (at least with the strength of Hardy–Littlewood conjecture assumed in Conjecture 1.3). Informally, the estimate (3.3) asserts that the parity of the prime counting function in an interval $(\mathbf{n}, \mathbf{n} + \lambda \log x]$ is approximately equidistributed for λ moderately large (but not too large).

Fix λ in the indicated range. By rounding if necessary, we may take $\lambda \log x$ to be an integer. We now invoke a Bonferroni type inequality (in the spirit of the Brun pure

sieve) to provide tractable upper and lower bounds on the random variable $(-1)^{\pi(\mathbf{n}+\lambda \log x)-\pi(\mathbf{n})}$:

Lemma 3.1 (Bonferroni type inequality). *For any nonnegative integers N, r , we have*

$$(-1)^N \leq \sum_{k=0}^r (-2)^k \binom{N}{k}$$

when r is even and

$$(-1)^N \geq \sum_{k=0}^r (-2)^k \binom{N}{k}$$

when r is odd, using the convention that $\binom{N}{k}$ vanishes for $k > N$ (or $k < 0$).

For instance, we have

$$1 - 2\binom{N}{1} \leq (-1)^N \leq 1 - 2\binom{N}{1} + 4\binom{N}{2}$$

for all $N \geq 0$.

Proof. The bounds are obvious when $N = 0$ or $r = 0$. In the case $N, r > 0$, one can proceed inductively, noting that

$$(-1)^N = (-1)^{N-1} + (-2) \times (-1)^{N-1}$$

and (by the Pascal identity $\binom{N}{k} = \binom{N-1}{k} + \binom{N-1}{k-1}$)

$$\sum_{k=0}^r (-2)^k \binom{N}{k} = \sum_{k=0}^r (-2)^k \binom{N-1}{k} + (-2) \times \sum_{k=0}^{r-1} (-2)^k \binom{N-1}{k}.$$

□

In view of this lemma and linearity of expectation, it suffices to show that

$$(3.4) \quad \sum_{k=0}^r (-2)^k \mathbf{E} \left(\binom{\pi(\mathbf{n} + \lambda \log x) - \pi(\mathbf{n})}{k} \right) \ll \frac{1}{\sqrt{\lambda}}$$

whenever $r = (\log \log x)^{4.5} + O(1)$ (the significance of this choice being that r is slightly larger than λ , while still within the range for which Conjecture 1.3 can be applied). It will be necessary to exploit the oscillation in k , and so we will largely refrain from using the triangle inequality in k when proving (3.4), except with regards to error terms that enjoy a power savings in x , which will always be negligible in practice.

Fix r in the indicated range. Observe from the combinatorial interpretation of the binomial coefficient that

$$\binom{\pi(\mathbf{n} + \lambda \log x) - \pi(\mathbf{n})}{k} = \sum_{0 < h_1 < \dots < h_k \leq \lambda \log x} \mathbf{1}_{\mathbf{n}+h_1, \dots, \mathbf{n}+h_k \in \mathcal{P}}$$

and thus by linearity of expectation

$$\mathbf{E} \left(\binom{\pi(\mathbf{n} + \lambda \log x) - \pi(\mathbf{n})}{k} \right) = \sum_{0 < h_1 < \dots < h_k \leq \lambda \log x} \mathbf{P}(\mathbf{n} + h_1, \dots, \mathbf{n} + h_k \in \mathcal{P}).$$

For $k \leq r \ll (\log \log x)^{4.5}$, we see from Conjecture 1.3 and routine manipulations that

$$\mathbf{P}(\mathbf{n} + h_1, \dots, \mathbf{n} + h_k \in \mathcal{P}) = \frac{\mathfrak{S}(\mathcal{H})}{\log^k x} + O(x^{-\varepsilon/2})$$

and thus the left-hand side of (3.4) can be written as

$$\sum_{k=0}^r \frac{(-2)^k}{\log^k x} \sum_{0 < h_1 < \dots < h_k \leq \lambda \log x} \mathfrak{S}(\mathcal{H}) + O\left(\sum_{k=0}^r \frac{(\lambda \log x)^k}{k!} 2^k x^{-\varepsilon/2}\right)$$

where $\mathcal{H} := \{h_1, \dots, h_k\}$. The error term can be crudely bounded by $O(r(2\lambda \log x)^r x^{-\varepsilon/2}) = O(x^{-\varepsilon/4})$, which is acceptable. Thus it will suffice to show that

$$(3.5) \quad \sum_{k=0}^r \frac{(-2)^k}{\log^k x} \sum_{0 < h_1 < \dots < h_k \leq \lambda \log x} \mathfrak{S}(\mathcal{H}) \ll \frac{1}{\sqrt{\lambda}}.$$

It is tempting to use Gallagher type estimates [9] on the mean value of singular series to control the left-hand side (which would morally obtain the sharper bound $O(e^{-2\lambda})$), but unfortunately the best known estimates of this type, see [15, Theorem 1.1], are only available for k a little bit smaller than $(\log \log x)^{0.5}$, which is not quite good enough⁶ in our situation. To get around this, we will reformulate the above estimate using a version of the random sieve model from [2], allowing one to exploit cancellations between the different values of k . As in [2, §1.3] (see also [21]), we define the sieve cutoff z to be the largest prime number for which

$$\prod_{p \leq z} \left(1 - \frac{1}{p}\right) \leq \frac{1}{\log x},$$

then from Mertens' theorem we have $z \asymp x^{1/e^\gamma}$ and

$$(3.6) \quad \prod_{p \leq z} \left(1 - \frac{1}{p}\right) = \frac{1}{\log x} + O(x^{-1/e^\gamma})$$

where $\gamma = 0.5772\dots$ is the Euler–Mascheroni constant. We select a residue class $\mathbf{a}_p \pmod{p}$ for each prime $p \leq z$ uniformly at random, independently in p , and then define the random sifted sets \mathcal{S}_w for every $w \leq z$ by removing these residue classes for $p \leq w$ from the interval $(0, \lambda \log x]$, thus

$$\mathcal{S}_w := \{0 < h \leq \lambda \log x : h \not\equiv \mathbf{a}_p \pmod{p} \text{ for all } p \leq w\}.$$

As in [2], one should think of \mathcal{S}_z as a simplified model for the random set $\{0 < h \leq \lambda \log x : \mathbf{n} + h \in \mathcal{P}\}$.

⁶More specifically, we were only able to use the estimates in [15] to prove a weaker version of (3.1) in which the exponent 1.1 is replaced by any exponent less than 1/4. On the other hand, the *upper bound* estimates in [15] are valid for all k , and will be useful to us to control an error term later in the argument. It may also be possible to use the random sieve model to strengthen some of the estimates in [15], but we have not pursued this question further.

For any tuple $0 < h_1 < \dots < h_k \leq \lambda \log x$ and any $\lambda \log x \leq w \leq z$, we have from construction that

$$\begin{aligned}
 \mathbf{P}(h_1, \dots, h_k \in \mathcal{S}_w) &= \prod_{p \leq w} \mathbf{P}(h_1, \dots, h_k \neq \mathbf{a}_p \pmod{p}) \\
 &= \prod_{p \leq w} \left(1 - \frac{\nu_p(\{h_1, \dots, h_k\})}{p} \right) \\
 &= \mathfrak{C}(\{h_1, \dots, h_k\}) \left(\prod_{p \leq w} \left(1 - \frac{1}{p} \right)^k \right) \prod_{p > w} \frac{\left(1 - \frac{1}{p} \right)^k}{1 - \frac{k}{p}}
 \end{aligned}
 \tag{3.7}$$

since $\nu_p(\{h_1, \dots, h_k\}) = k$ for $p > w$. If we restrict to the regime $k \leq r$, then we can write

$$1 - \frac{k}{p} = \exp\left(-\frac{k}{p} + O\left(\frac{k^2}{p^2}\right)\right)$$

and

$$1 - \frac{1}{p} = \exp\left(-\frac{1}{p} + O\left(\frac{1}{p^2}\right)\right)$$

for $p > w$, and hence

$$\prod_{p > w} \frac{\left(1 - \frac{1}{p} \right)^k}{1 - \frac{k}{p}} = \exp\left(O\left(\frac{k^2}{w}\right)\right)$$

and thus

$$\mathbf{P}(h_1, \dots, h_k \in \mathcal{S}_w) = \mathfrak{C}(\{h_1, \dots, h_k\}) \left(\prod_{p \leq w} \left(1 - \frac{1}{p} \right)^k \right) \left(1 + O\left(\frac{k^2}{w}\right) \right).$$

In particular, from (3.6) we have

$$\mathbf{P}(h_1, \dots, h_k \in \mathcal{S}_z) = \frac{\mathfrak{C}(\{h_1, \dots, h_k\})}{\log^k x} + O(x^{-\varepsilon})$$

for any $0 \leq k \leq r$ (shrinking ε if necessary). We can then rewrite (3.5) as

$$\sum_{k=0}^r (-2)^k \sum_{0 < h_1 < \dots < h_k \leq \lambda \log x} \mathbf{P}(h_1, \dots, h_k \in \mathcal{S}_z) \ll \frac{1}{\sqrt{\lambda}},$$

since the contribution of the $O(x^{-\varepsilon})$ error is negligible as before. We can simplify the left-hand side to

$$\sum_{k=0}^r (-2)^k \mathbf{E} \binom{\mathbf{S}_z}{k},$$

where $\mathbf{S}_w := |\mathcal{S}_w|$ denotes the size of the sifted set at level w . From Lemma 3.1 we have

$$\sum_{k=0}^r (-2)^k \binom{\mathbf{S}_z}{k} = (-1)^{\mathbf{S}_z} + O\left(2^r \binom{\mathbf{S}_z}{r}\right),$$

so it will suffice to establish the bounds

$$\mathbf{E}(-1)^{\mathbf{S}_z} \ll \frac{1}{\sqrt{\lambda}}$$

and

$$(3.11) \quad 2^r \mathbf{E} \binom{\mathbf{S}_z}{r} \ll \frac{1}{\sqrt{\lambda}}.$$

We begin with (3.11). We can expand the left-hand side as

$$2^r \sum_{0 < h_1 < \dots < h_r \leq \lambda \log x} \mathbf{P}(h_1, \dots, h_r \in \mathcal{S}_z).$$

Using (3.9) and discarding the error term using crude estimates, it suffices to show that

$$\frac{2^r}{\log^r x} \sum_{0 < h_1 < \dots < h_r \leq \lambda \log x} \mathfrak{E}(\mathcal{H}) \ll \frac{1}{\sqrt{\lambda}}.$$

Using [15, Theorem 1.2], we may bound the left-hand side by

$$\ll \frac{2^r \lambda^r}{r!} (3 \log r)^r$$

which by Stirling's formula can be estimated by

$$\ll \left(\frac{O(\lambda \log r)}{r} \right)^r.$$

Since $\lambda \leq (\log \log x)^{4.4}$ and $r = (\log \log x)^{4.5} + O(1)$, one easily verifies that this quantity is $O(1/\sqrt{\lambda})$ as required.

It remains to establish (3.10), which is an analogue of (3.3) for the random sieve model \mathcal{S}_z of the primes. We first control the first two moments of \mathbf{S}_w :

Lemma 3.2 (Mean and variance). *For $\lambda \log x \leq w \leq z$, we have*

$$(3.12) \quad \mathbf{E} \mathbf{S}_w = \lambda \log x \prod_{p \leq w} \left(1 - \frac{1}{p} \right) = \frac{\lambda \log x}{e^{\gamma} \log w} \left(1 + O \left(\frac{1}{\log w} \right) \right)$$

and

$$(3.13) \quad \mathbf{Var}(\mathbf{S}_w) \ll \frac{\lambda \log x}{\log w}.$$

Proof. From linearity of expectation we have

$$\mathbf{E} \mathbf{S}_w = \sum_{0 < h \leq \lambda \log x} \mathbf{P}(h \in \mathcal{S}_w)$$

and similarly

$$\mathbf{E}(\mathbf{S}_w^2 - \mathbf{S}_w) = 2 \mathbf{E} \binom{\mathbf{S}_w}{2} = 2 \sum_{0 < h_1 < h_2 \leq \lambda \log x} \mathbf{P}(h_1, h_2 \in \mathcal{S}_w).$$

From (3.7) and Mertens' theorem we have

$$\mathbf{P}(h \in \mathcal{S}_w) = \prod_{p \leq w} \left(1 - \frac{1}{p} \right) = \frac{1}{e^{\gamma} \log w} \left(1 + O \left(\frac{1}{\log w} \right) \right)$$

for all $0 < h \leq w$, which gives (3.12). To establish (3.13), we see from (3.8) that

$$\mathbf{P}(h_1, h_2 \in \mathcal{S}_w) = \mathfrak{E}(\{h_1, h_2\}) \prod_{p \leq w} \left(1 - \frac{1}{p} \right)^2 \left(1 + O \left(\frac{1}{w} \right) \right)$$

for $0 < h_1 < h_2 \leq \lambda \log x \leq w$. On the other hand, the bound

$$(3.14) \quad 2 \sum_{0 < h_1 < h_2 \leq H} \mathfrak{S}(\{h_1, h_2\}) \leq H^2$$

is known for H large enough. In fact there is a long history⁷ of results of this type:

- (i) The left-hand side of (3.14) was shown to be $H^2 + O(H \log H)$ in [18, Lemma 2].
- (ii) This was improved to $H^2 - H \log H + O(H)$ in unpublished work of Montgomery, with a full proof appearing in [5]. Note that this is already enough to establish (3.14).
- (iii) The sharper asymptotic $H^2 - H \log H + (1 - \gamma - \log 2\pi)H + O_\varepsilon(H^{1/2+\varepsilon})$ for any $\varepsilon > 0$ was stated without proof in [10], but a proof can be found in [20, (16), (17)] or [19].
- (iv) Assuming the Riemann hypothesis, an even sharper asymptotic of $H^2 - H \log H + (1 - \gamma - \log 2\pi)H + O_\varepsilon(H^{5/12+\varepsilon})$ was established in [23], who also obtained a more modest sharpening (of Vinogradov–Korobov type) to (iii) without assuming the Riemann hypothesis.
- (v) In [12], it is shown the error term in (iii) or (iv) is of magnitude $\gg H^{1/4}$ (with either choice of sign) for infinitely many H .

We also remark that the fact that the left-hand side of (3.14) is somewhat less than H^2 is closely related to the pair correlation conjecture, as well as the expected belief that the error term in the prime number theorem in medium length intervals is somewhat smaller than what is predicted from the Cramér random model. See [11], [20], [19], [3] for further discussion. These asymptotics are also related to unexpected biases in the distribution of consecutive primes [16].

From (3.14) we have

$$2 \sum_{0 < h_1 < h_2 \leq \lambda \log x} \mathfrak{S}(\{h_1, h_2\}) \leq (\lambda \log x)^2.$$

We conclude that

$$\mathbf{E}(\mathbf{S}_w^2 - \mathbf{S}_w) \leq (\lambda \log x)^2 \prod_{p \leq w} \left(1 - \frac{1}{p}\right)^2 + O\left(\frac{(\lambda \log x)^2}{w} \prod_{p \leq w} \left(1 - \frac{1}{p}\right)^2\right).$$

Combining this with (3.12) and rearranging, we conclude that

$$\mathbf{Var}(\mathbf{S}_w) \ll \mathbf{E}\mathbf{S}_w + \frac{(\lambda \log x)^2}{w} \prod_{p \leq w} \left(1 - \frac{1}{p}\right)^2$$

and the claim now follows from (3.12), Mertens’ theorem, and the hypothesis $w \geq \lambda \log x$. □

To establish (3.10), we will obtain a recursive inequality for the bias $\mathbf{E}(-1)^{\mathbf{S}_{p_n}}$. If $\lambda \log x \leq p_n < p_{n+1} \leq z$ are consecutive primes, we observe that $\mathfrak{S}_{p_{n+1}}$ is obtained from \mathfrak{S}_{p_n} by removing one residue class modulo p_{n+1} , chosen uniformly and independently of \mathfrak{S}_{p_n} . Since $p_{n+1} \geq \lambda \log x$ exceeds the diameter of \mathfrak{S}_{p_n} , we conclude that after conditioning on \mathbf{S}_{p_n} , we have $\mathbf{S}_{p_{n+1}} = \mathbf{S}_{p_n}$ with conditional probability $1 - \mathbf{S}_{p_n}/p_{n+1}$,

⁷We are indebted to Ofir Gorodetsky and Dan Goldston for these references.

and $\mathbf{S}_{p_{n+1}} = \mathbf{S}_{p_n} - 1$ with conditional probability \mathbf{S}_{p_n}/p_{n+1} . From the law of total expectation, we conclude the identity

$$\begin{aligned} \mathbf{E}(-1)^{\mathbf{S}_{p_{n+1}}} &= \mathbf{E}\left(1 - \frac{\mathbf{S}_{p_n}}{p_{n+1}}\right)(-1)^{\mathbf{S}_{p_n}} + \frac{\mathbf{S}_{p_n}}{p_{n+1}}(-1)^{\mathbf{S}_{p_n}-1} \\ &= \mathbf{E}\left(1 - \frac{2\mathbf{S}_{p_n}}{p_{n+1}}\right)(-1)^{\mathbf{S}_{p_n}}. \end{aligned}$$

We now exploit the concentration of \mathbf{S}_{p_n} around the mean $\mathbf{E}\mathbf{S}_{p_n}$ provided by Lemma 3.2. By the triangle inequality, we have

$$|\mathbf{E}(-1)^{\mathbf{S}_{p_{n+1}}}| \leq \left(1 - \frac{2\mathbf{E}\mathbf{S}_{p_n}}{p_{n+1}}\right)|\mathbf{E}(-1)^{\mathbf{S}_{p_n}}| + O\left(\frac{|\mathbf{E}\mathbf{S}_{p_n} - \mathbf{E}\mathbf{S}_{p_n}|}{p_{n+1}}\right).$$

The crucial point here is that the factor $1 - \frac{2\mathbf{E}\mathbf{S}_{p_n}}{p_{n+1}}$, while positive, is slightly less than one, indicating a tendency of the sieving process to reduce the bias as n increases. This bias reduction effect is difficult to discern directly from the alternating series appearing in (3.5), and is the main advantage of working with the random sieve model in this problem.

By Cauchy–Schwarz and (3.13), we have

$$\mathbf{E}|\mathbf{S}_{p_n} - \mathbf{E}\mathbf{S}_{p_n}| \ll \left(\frac{\lambda \log x}{\log p_n}\right)^{1/2}$$

while from (3.12) we have

$$\mathbf{E}\mathbf{S}_{p_n} = \frac{\lambda \log x}{e^\gamma \log p_n} \left(1 + O\left(\frac{1}{\log p_n}\right)\right).$$

Bounding $1 - \frac{2\mathbf{E}\mathbf{S}_{p_n}}{p_{n+1}}$ by $\exp\left(-\frac{2\mathbf{E}\mathbf{S}_{p_n}}{p_{n+1}}\right)$, we conclude the recursive inequality

$$|\mathbf{E}(-1)^{\mathbf{S}_{p_{n+1}}}| \leq \exp\left(-\frac{2\lambda \log x}{e^\gamma p_n \log p_n} + O\left(\frac{\lambda \log x}{p_n \log^2 p_n}\right)\right)|\mathbf{E}(-1)^{\mathbf{S}_{p_n}}| + O\left(\frac{1}{p_n} \left(\frac{\lambda \log x}{\log p_n}\right)^{1/2}\right).$$

Iterating⁸ this inequality starting from the trivial bound $|\mathbf{E}(-1)^{\mathbf{S}_p}| \leq 1$ for p the first prime greater than $\lambda \log x$, we conclude that

$$\mathbf{E}(-1)^{\mathbf{S}_z} \ll \alpha_{\lambda \log x} + \sum_{\lambda \log x \leq p \leq z} \frac{\alpha_p}{p} \left(\frac{\lambda \log x}{\log p}\right)^{1/2},$$

where α_w is a weight of the form

$$\alpha_w := \exp\left(-\sum_{w \leq p < z} \left(\frac{2\lambda \log x}{e^\gamma p \log p} + O\left(\frac{\lambda \log x}{p \log^2 p}\right)\right)\right).$$

From the prime number theorem and summation by parts, we have

$$\sum_{w \leq p < z} \frac{1}{p \log p} = \frac{1}{\log w} - \frac{1}{\log z} + O\left(\frac{1}{\log^2 w}\right)$$

and

$$\sum_{w \leq p < z} \frac{1}{p \log^2 p} \ll \frac{1}{\log^2 w}$$

⁸Here we are basically using the discrete version of Gronwall's inequality.

for any $\lambda \log x \leq w \leq z$, thus

$$(3.15) \quad \alpha_w = \exp\left(-\frac{2\lambda \log x}{e^\gamma \log w} + \frac{2\lambda \log x}{e^\gamma \log z} + O\left(\frac{\lambda \log x}{\log^2 w}\right)\right).$$

In particular we have

$$\alpha_{\lambda \log x} \leq \exp\left(-c \frac{\log x}{\log \log x}\right)$$

for some absolute constant $c > 0$, which is certainly $O(1/\sqrt{\lambda})$. It thus remains to establish the bound

$$\sum_{\lambda \log x \leq p \leq z} \frac{\alpha_p}{p} \left(\frac{\lambda \log x}{\log p}\right)^{1/2} \ll \frac{1}{\sqrt{\lambda}}.$$

We first dispose of the contribution of the small primes $p \leq x^{\frac{1}{100 \log \log x}}$, which will turn out to be quite negligible. From (3.15) we have

$$\alpha_p \leq \exp(-10 \log \log x) = \log^{-10} x$$

(say), while we may crudely bound

$$\left(\frac{\lambda \log x}{\log p}\right)^{1/2} \ll \log x$$

in this case. From Mertens' theorem, the contribution of this case is thus $O(\log^{-8} x)$, which is certainly $O(1/\sqrt{\lambda})$. Thus it remains to show that

$$(3.16) \quad \sum_{x^{\frac{1}{100 \log \log x}} \leq p \leq z} \frac{\alpha_p}{p} \left(\frac{\lambda \log x}{\log p}\right)^{1/2} \ll \frac{1}{\sqrt{\lambda}}.$$

For each prime p in this range we have

$$(3.17) \quad m \leq \frac{\lambda \log x}{\log p} < m + 1$$

for some natural number m with

$$\frac{\lambda \log x}{\log z} - 1 \leq m \ll (\log \log x)^{O(1)}.$$

For p in such a range, the $\frac{\lambda \log x}{\log^2 p}$ term in (3.15) is bounded, hence

$$\alpha_p \ll \exp\left(-\frac{2}{e^\gamma} \left(m - \frac{\lambda \log x}{\log z}\right)\right)$$

and so the contribution of a given m to (3.16) is

$$\ll \exp\left(-\frac{2}{e^\gamma} \left(m - \frac{\lambda \log x}{\log z}\right)\right) m^{1/2} \sum_{p: m \leq \frac{\lambda \log x}{\log p} < m+1} \frac{1}{p}.$$

The restriction (3.17) restricts $\log \log p$ to an interval of size $O(1/m)$, but also (due to the upper bound on m) encompasses at least one dyadic range $[X, 2X]$ of p . By Mertens' theorem (or the Brun–Titchmarsh inequality), we thus have

$$\sum_{p: m \leq \frac{\lambda \log x}{\log p} < m+1} \frac{1}{p} \ll \frac{1}{m},$$

and so the total contribution of (3.16) is

$$\ll \sum_{m \geq \frac{\lambda \log x}{\log z} - 1} \frac{1}{m^{1/2}} \exp\left(-\frac{2}{e^\gamma} \left(m - \frac{\lambda \log x}{\log z}\right)\right).$$

This is an exponentially decaying sum in m , and so can be controlled by its first term

$$\ll \left(\frac{\lambda \log x}{\log z}\right)^{-1/2} \asymp \frac{1}{\lambda^{1/2}},$$

giving the desired claim.

Remark 3.3. The above analysis in fact shows that the convergence rates for the partial sums $\sum_{n \leq x} \frac{(-1)^n n}{p_n}$, $\sum_{10 \leq n \leq x} \frac{(-1)^{\pi(n)}}{n \log n}$ are $O\left(\frac{1}{(\log \log x)^{0.1}}\right)$, assuming Conjecture 1.3. Probabilistic heuristics suggest that the true rate of convergence is $O(1/\log x)$ and $O(1/x^{1/2-o(1)})$ respectively for the above two partial sums, though conjecturally the convergence of the first partial sum should also become $O(1/x^{1/2-o(1)})$ when one averages together consecutive partial sums to reduce oscillation.

Remark 3.4. Without the prime tuples conjecture, considerably less is known about the cancellation of $(-1)^{\pi(n)}$. In [1], [4] it is shown that $|\sum_{n \leq x} (-1)^{\pi(n)}| \leq \left(1 - \frac{1}{32} + o(1)\right)x$ unconditionally, which is a slight improvement over the trivial bound $|\sum_{n \leq x} (-1)^{\pi(n)}| \leq x$.

4. A GENERALIZATION

The questions and arguments in this section arose from several conversations with William Banks.

In this section we record a generalization of the above results, in which the base (-1) is replaced by other non-trivial unit phases. More precisely, we consider the following two questions:

Question 4.1. Let $z \in \mathbb{C}$ with $|z| = 1$ and $z \neq 1$. Does the series $\sum_{n=1}^{\infty} \frac{z^n n}{p_n}$ converge?

Question 4.2. Let $z \in \mathbb{C}$ with $|z| = 1$ and $z \neq 1$. Does the series $\sum_{n=2}^{\infty} \frac{z^{\pi(n)}}{n \log n}$ converge?

We first observe that these questions continue to be equivalent for any fixed z . Indeed one has a generalization

$$\sum_{n \leq x} \frac{z^n n}{p_n} = \frac{z}{z-1} \sum_{2 \leq m \leq x \log x} \frac{z^{\pi(m)}}{m \log m} + C_z + o(1)$$

of (2.1) as $x \rightarrow \infty$, where C_z is a constant depending on z and the rates of decay in asymptotic notation such as $o(1)$ is permitted to depend non-uniformly on z . We sketch the argument as follows. Writing $z^n = \frac{z}{z-1}(z^n - z^{n-1})$ and summing by parts, we have

$$\sum_{n \leq x} \frac{z^n n}{p_n} = -\frac{z}{2(z-1)} + \frac{z}{z-1} \sum_{n \leq x} z^n \left(\frac{n}{p_n} - \frac{n+1}{p_{n+1}}\right) + o(1).$$

As before we split

$$z^n \left(\frac{n+1}{p_{n+1}} - \frac{n}{p_n}\right) = z^n \frac{n(p_{n+1} - p_n)}{p_n p_{n+1}} - \frac{z^n}{p_n}.$$

A further summation by parts reveals that $\sum_{n=1}^{\infty} \frac{z^n}{p_n}$ converges, while from (2.3) we see that

$$\sum_{n=1}^{\infty} \left(\sum_{p_n \leq m < p_{n+1}} \frac{z^{\pi(m)}}{m \log m} - \frac{z^n n(p_{n+1} - p_n)}{p_n p_{n+1}} \right)$$

is absolutely convergent. The claim then follows along similar lines to the remaining arguments in Section 2.

Theorem 1.4 can also be extended to general z by adapting the arguments in Section 3. The Bonferroni-type inequality in Lemma 3.1 can be replaced with the weaker bound

$$z^N = \sum_{k=0}^r (z - 1)^k \binom{N}{k} + O\left(N 2^r \binom{N}{r}\right)$$

for any $r \geq 1$, which can be established by combining the unimodal nature of the function $k \mapsto |z - 1|^k \binom{N}{k}$ (or more precisely, that this function increases, then decreases), combined with the binomial formula $z^N = \sum_{k=0}^N (z - 1)^k \binom{N}{k}$. The rest of the argument proceeds with minor modifications; for instance, several factors of 2 appearing in expressions such as $2\mathbf{ES}_{p_n}$ will need to be replaced by $1 - z$ (or $\operatorname{Re}(1 - z)$ after an application of the triangle inequality). We leave the details of these changes to the interested reader.

To summarize the above discussion, we have shown that assuming Conjecture 1.3, the power series $\sum_{n=1}^{\infty} \frac{z^n n}{p_n}$ converges everywhere on the boundary of the disk $\{z : |z| \leq 1\}$ of convergence, except at $z = 1$ where of course it diverges, thanks to the prime number theorem. Based on analogy with random power series, it seems natural to conjecture⁹ that the unit circle is in fact the natural boundary of this series.

5. FURTHER QUESTIONS

In [7], the divergence and convergence of various other series relating to the primes is considered. In particular, the paper demonstrates the convergence of

$$(5.1) \quad \sum_{n=10}^{\infty} \frac{1}{n(\log \log n)^c (p_{n+1} - p_n)}$$

whenever $c > 2$, and conjectures divergence for $c = 2$; the paper also mentions Question 1.1, as well as the question of determining the convergence of the series

$$(5.2) \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{p_{n+1} - p_n}$$

and

$$(5.3) \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{n(p_{n+1} - p_n)}.$$

The divergence of (5.2) can now be established unconditionally, since Zhang's theorem [24] on bounded gaps between primes implies that the summand fails to converge to zero. The divergence of (5.1) when $c = 2$ can be deduced from Conjecture 1.3 by an

⁹We thank an anonymous commenter on the author's blog for this question.

argument that we sketch¹⁰ as follows. By dyadic decomposition and the prime number theorem, it suffices to establish the lower bound

$$\sum_{X \leq p_n < 2X} \frac{1}{p_{n+1} - p_n} \gg \frac{X \log \log X}{\log^2 X}$$

for all large X , since $\frac{1}{n(\log \log n)^2} \asymp \frac{\log X}{X(\log \log X)^2}$ for all n in this range, and $\frac{1}{\log X \log \log X}$ diverges when summed over powers of two. By further dyadic decomposition in the gap $p_{n+1} - p_n$, it suffices to obtain the bound

$$(5.4) \quad \#\{X \leq p_n < 2X : p_{n+1} - p_n \leq \lambda \log X\} \gg \frac{\lambda X}{\log X}$$

for any $\frac{2}{\log X} \leq \lambda \leq 1$. It suffices to treat the case where λ is small. The left-hand side can be written as

$$\sum_{X \leq p_n < 2X} 1_{\#\{p_n, p_n + \lambda \log X\} \cap \mathcal{P} \neq \emptyset}$$

Using the Bonferroni inequalities

$$m - \binom{m}{2} \leq 1_{m \neq 0} \leq m$$

for any $m \geq 0$, and rearranging, we can write this expression as

$$\sum_{2 \leq h \leq \lambda \log X} \sum_{X \leq n < 2X} 1_{n, n+h \in \mathcal{P}} - O\left(\sum_{2 \leq h < h' \leq \lambda \log X} \sum_{X \leq n < 2X} 1_{n, n+h, n+h' \in \mathcal{P}}\right),$$

so by Conjecture 1.3 it suffices to establish the bounds

$$\sum_{2 \leq h \leq \lambda \log X} \mathfrak{C}(\{0, h\}) \asymp \lambda \log X$$

and

$$\sum_{2 \leq h < h' \leq \lambda \log X} \mathfrak{C}(\{0, h, h'\}) \asymp \lambda^2 \log^2 X$$

since the contribution of the error terms are easily treated. But this follows from Gallagher's estimate [9, (3)] (or from the more refined estimates in [20], [15]). We remark that in the absence of the prime tuples conjecture, it should be possible to obtain the weaker version

$$\#\{X \leq p_n < 2X : p_{n+1} - p_n \leq \lambda \log X\} \gg \frac{\lambda^{49} X}{\log X}$$

of (5.4) for $\frac{246}{\log X} \leq \lambda \leq 1$ by combining the method of Maynard [17] with the calculations in [22]. This is however insufficient to settle the divergence of (5.2) in the $c = 2$ case. We also remark that a matching upper bound to (5.4) can be established from standard upper bound sieves.

The convergence of the sum (5.3) remains a challenging open problem, even with the assumption of Conjecture 1.3; a significant portion of this sum arises from the small prime gaps in which $p_{n+1} - p_n$ is much smaller than $\log n$, but these are a sparse set of n and so the van der Corput type methods used in this paper do not appear to be effective to induce cancellation from the $(-1)^n$ terms on this sparse set. However, standard

¹⁰We thank Kevin Ford for supplying a version of this argument.

probabilistic heuristics (for instance, replacing $(-1)^n$ by independent random signs and invoking Khintchine's inequality) suggest rather strongly that this series should be convergent. Indeed this heuristic suggests the stronger assertion that the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^{\theta}(p_{n+1} - p_n)}$$

should converge for $\theta > 1/2$ and diverge for $\theta \leq 1/2$.

ACKNOWLEDGMENTS

The author would like to thank Kevin Ford, Dan Goldston, Ofir Gorodetsky, and Kannan Soundararajan for supplying references, and William Banks for supplying references and suggesting some additional questions and generalizations to the argument. The author would also like to thank Binbin Hu, Keiju Sono, and the anonymous referee for corrections.

REFERENCES

- [1] Mihai Alboiu, *On the parity of the prime-counting function and related problems*, Ramanujan J. **38** (2015), no. 1, 179–187, DOI 10.1007/s11139-014-9596-1. MR3396493
- [2] William Banks, Kevin Ford, and Terence Tao, *Large prime gaps and probabilistic models*, Invent. Math. **233** (2023), no. 3, 1471–1518, DOI 10.1007/s00222-023-01199-0. MR4623547
- [3] Tszho Chan, *Pair correlation and distribution of prime numbers*, ProQuest LLC, Ann Arbor, MI, 2002. Thesis (Ph.D.)—University of Michigan. MR2703577
- [4] Ping Ngai Chung and Shiyu Li, *On the residue classes of $\pi(n)$ modulo t* , Integers **13** (2013), Paper No. A79, 10. MR3167926
- [5] M. J. Croft, *Square-free numbers in arithmetic progressions*, Proc. London Math. Soc. (3) **30** (1975), 143–159, DOI 10.1112/plms/s3-30.2.143. MR379402
- [6] Paul Erdős, *Some of my new and almost new problems and results in combinatorial number theory*, Number theory (Eger, 1996), de Gruyter, Berlin, 1998, pp. 169–180. MR1628841
- [7] Paul Erdős and Melvyn B. Nathanson, *On the sum of the reciprocals of the differences between consecutive primes*, Number theory (New York, 1991), Springer, New York, 1996, pp. 97–101, DOI 10.1007/978-1-4612-2418-1_7. MR1420202
- [8] Steven R. Finch, *Mathematical constants*, Encyclopedia of Mathematics and its Applications, vol. 94, Cambridge University Press, Cambridge, 2003. MR2003519
- [9] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), no. 1, 4–9, DOI 10.1112/S0025579300016442. MR409385
- [10] D. A. Goldston, *Linnik's theorem on Goldbach numbers in short intervals*, Glasgow Math. J. **32** (1990), no. 3, 285–297, DOI 10.1017/S001708950000937X. MR1073669
- [11] Daniel A. Goldston and Hugh L. Montgomery, *Pair correlation of zeros and primes in short intervals*, Analytic number theory and Diophantine problems (Stillwater, OK, 1984), Progr. Math., vol. 70, Birkhäuser Boston, Boston, MA, 1987, pp. 183–203. MR1018376
- [12] D. A. Goldston and Ade Irma Suriajaya, *A singular series average and the zeros of the Riemann zeta-function*, Acta Arith. **200** (2021), no. 1, 71–90, DOI 10.4064/aa200821-24-2. MR4319607
- [13] Richard K. Guy, *Unsolved problems in number theory*, 2nd ed., Problem Books in Mathematics, Springer-Verlag, New York, 1994. Unsolved Problems in Intuitive Mathematics, I, DOI 10.1007/978-1-4899-3585-4. MR1299330
- [14] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), no. 1, 1–70, DOI 10.1007/BF02403921. MR1555183
- [15] Vivian Kuperberg, *Sums of singular series with large sets and the tail of the distribution of primes*, Q. J. Math. **74** (2023), no. 4, 1457–1479, DOI 10.1093/qmath/haad030. MR4676553
- [16] Robert J. Lemke Oliver and Kannan Soundararajan, *Unexpected biases in the distribution of consecutive primes*, Proc. Natl. Acad. Sci. USA **113** (2016), no. 31, E4446–E4454, DOI 10.1073/pnas.1605366113. MR3624386
- [17] James Maynard, *Dense clusters of primes in subsets*, Compos. Math. **152** (2016), no. 7, 1517–1554, DOI 10.1112/S0010437X16007296. MR3530450

- [18] H. L. Montgomery, *Primes in arithmetic progressions*, Michigan Math. J. **17** (1970), 33–39. MR257005
- [19] H. L. Montgomery and K. Soundararajan, *Beyond pair correlation*, Paul Erdős and his mathematics, I (Budapest, 1999), Bolyai Soc. Math. Stud., vol. 11, János Bolyai Math. Soc., Budapest, 2002, pp. 507–514. MR1954710
- [20] Hugh L. Montgomery and K. Soundararajan, *Primes in short intervals*, Comm. Math. Phys. **252** (2004), no. 1-3, 589–617, DOI 10.1007/s00220-004-1222-4. MR2104891
- [21] G. Pólya, *Heuristic reasoning in the theory of numbers*, Amer. Math. Monthly **66** (1959), 375–384, DOI 10.2307/2308748. MR104639
- [22] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), Art. 12, 83, DOI 10.1186/s40687-014-0012-7. MR3373710
- [23] R. C. Vaughan, *On a variance associated with the distribution of primes in arithmetic progressions*, Proc. London Math. Soc. (3) **82** (2001), no. 3, 533–553, DOI 10.1112/plms/82.3.533. MR1816688
- [24] Yitang Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174, DOI 10.4007/annals.2014.179.3.7. MR3171761

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CALIFORNIA 90095-1555

Email address: tao@math.ucla.edu