LARGE DEVIATIONS IN SELBERG'S CENTRAL LIMIT THEOREM

MAKSYM RADZIWIŁŁ

Abstract. Following Selberg [10] it is known that as $T \to \infty$,

$$\frac{1}{T} \operatorname*{meas}_{t \in [T;2T]} \left\{ \log |\zeta(\tfrac{1}{2} + it)| \geq \Delta \sqrt{\tfrac{1}{2} \log\log T} \right\} \sim \int_{\Delta}^{\infty} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}$$

uniformly in $\Delta \leq (\log\log\log T)^{1/2-\varepsilon}$. We extend the range of Δ to $\Delta \ll (\log\log T)^{1/10-\varepsilon}$. We also speculate on the size of the largest Δ for which the above normal approximation can hold and on the correct approximation beyond this point.

1. Introduction.

The value-distribution of $\log \zeta(\sigma + it)$ is a classical question in the theory of the Riemann zeta function. When $\sigma > 1/2$ this distribution is well-understood and is that of an almost surely convergent sequence of random variables (see [6], [3] , [7] or [13]).

The half-line is special, since for $\sigma = \frac{1}{2}$ we have a *central limit theorem*,

$$\frac{1}{T} \cdot \operatorname*{meas}_{t \in [T;2T]} \left\{ \frac{\log |\zeta(\frac{1}{2} + it)|}{\sqrt{\frac{1}{2} \log\log T}} \geqslant \Delta \right\} = \int_{\Delta}^{\infty} e^{-u^2/2} \cdot \frac{du}{\sqrt{2\pi}} + o(1) \ , \ T \to \infty \tag{1}$$

originally due to Selberg [10]. Whereas the distribution of large values of $\log |\zeta(\sigma + it)|$ with $\sigma > 1/2$ has been consistently studied since the pioneering work of Bohr and Jessen [1], the corresponding question on the half-line has only attracted more attention recently.

Conditionally on the Riemann Hypothesis, Soundararajan [12] obtained Gaussian upper bounds for the left-hand side of (1) (focusing mostly on the range $\Delta \gg \sqrt{\log\log T}$). As an application he derived near optimal upper bounds for moments of the Riemann zeta-function.

In this paper we will be interested in asymptotic formulas for the left-hand side of (1) when $\Delta \to \infty$ as $T \to \infty$. Previously asymptotic formulae of the form

$$\frac{1}{T} \cdot \operatorname*{meas}_{t \in [T;2T]} \left\{ \frac{\log |\zeta(\frac{1}{2} + it)|}{\sqrt{\frac{1}{2} \log\log T}} \geqslant \Delta \right\} \sim \int_{\Delta}^{\infty} e^{-u^2/2} \cdot \frac{du}{\sqrt{2\pi}} \tag{2}$$

where known only in the range $\Delta \ll (\log\log\log T)^{1/2-\varepsilon}$ as a consequence of Selberg's [15] near-optimal refinement of the error term in (1) [1] In this paper we introduce a new method that allows us to extend (2) to the large-deviations range $\Delta \ll (\log\log T)^{\alpha}$ for some small but fixed $\alpha > 0$.

Theorem 1. *The asymptotic formulae (2) holds for* $\Delta \ll (\log\log T)^{1/10-\varepsilon}$.

1991 *Mathematics Subject Classification.* Primary: 11M06, Secondary: 11M50.

The author is partially supported by a NSERC PGS-D award.

[1]Conjecturally (1) admits an asymptotic expansion in powers of $(\log\log T)^{-1/2}$. Selberg's obtained an error term of $O((\log\log\log T)^2 (\log\log T)^{-1/2})$ in (1), which is thus close to optimal in this sense.

Our method is very versatile and allows to extend most of the known distribution results about $\zeta(\sigma + \mathrm{i}t)$ (with $\sigma = 1/2 + o_{T\to\infty}(1)$) to a large deviations setting (for example [2] or the joint value distribution of $\Re \log \zeta(\frac12 + \mathrm{i}t)$ and $\Im \log \zeta(\frac12 + \mathrm{i}t)$). The method also adapts to the study of the distribution of large values of additive functions over sets of integers where only moments (but not moment generating functions) are available.

To prove Theorem 1 we use Selberg's work to reduce the problem to a question about Dirichlet polynomials. Theorem 1 then follows from the Proposition below, which might be of independent interest.

Proposition 1. *Let $x = T^{1/(\log\log T)^2}$. Then, uniformly in $\Delta = o(\sqrt{\log\log T})$,*

$$\frac{1}{T}\cdot \underset{t\in[T;2T]}{meas}\left\{\Re \sum_{p\leqslant x} \frac{1}{p^{\frac12+\mathrm{i}t}} \geqslant \Delta\cdot\sqrt{\tfrac12\sum_{p\leqslant x}\frac{1}{p}}\right\} \sim \int_{\Delta}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}} \tag{3}$$

Previously a result such as Proposition 1 was known only for fairly short Dirichlet polynomials (i.e $x \leqslant (\log T)^{\theta}$) (see [9] and also [8] for related work). The extension in the length of the Dirichlet polynomial is responsible for our improvement in Theorem 1.

The key idea in our proof of Proposition 1 is to work on a subset A of $[T;2T]$ on which the Dirichlet polynomial does not attain large values. We explain the idea briefly in more details in subsection (1.2) below.

Proposition 1 can be extended to the range $\Delta \asymp \sqrt{\log\log T}$. In that situation we obtain the following result.

Proposition 2. *Let $x = T^{1/(\log\log T)^2}$ and $k > 0$. Uniformly in $\Delta \sim k\sqrt{\frac12\log\log T}$,*

$$\frac{1}{T}\cdot \underset{t\in[T;2T]}{meas}\left\{\Re \sum_{p\leqslant x} \frac{1}{p^{\frac12+\mathrm{i}t}} \geqslant \Delta\cdot\sqrt{\tfrac12\sum_{p\leqslant x}\frac{1}{p}}\right\} \sim c_k\int_{\Delta}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}}$$

with $0 < c_k \neq 1$ a constant, depending only on k.

It is thus apparent that (3) does not persist for $\Delta \geqslant \varepsilon\sqrt{\log\log T}$. Similarly,

$$\frac{1}{T}\cdot \underset{t\in[T;2T]}{\mathrm{meas}}\left\{\frac{\log|\zeta(\frac12+\mathrm{i}t)|}{\sqrt{\frac12\log\log T}} \geqslant \Delta\right\} \sim \int_{\Delta}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}} \tag{4}$$

cannot be true for $\Delta \geqslant \varepsilon\sqrt{\log\log T}$ with $\varepsilon > 0$ small. Indeed, one can show (on the Riemann Hypothesis) that this would contradict the moment conjectures,

$$\int_T^{2T} |\zeta(\tfrac12+\mathrm{i}t)|^{2k}\mathrm{d}t \sim C_k\cdot T(\log T)^{k^2} \tag{5}$$

($T\to\infty$) since it is conjectured that $C_k \neq 1$ for $k \in (0,1)$ and if (4) holds for $\Delta \leq \varepsilon\sqrt{\log\log T}$ then $C_k = 1$ for $k \leq \varepsilon$ (the Riemann Hypothesis is used because we appeal to [12] to first restrict the range of integration in (5) to those t's at which $\log|\zeta(\frac12+\mathrm{i}t)| \sim k\log\log T$). Nonetheless, by analogy to Proposition 1 we expect (4) to hold for all smaller Δ.

Conjecture 1. *If $\Delta = o(\sqrt{\log\log T})$ then (4) holds.*

For larger values of Δ – in analogy to Proposition 2 – we conjecture that (4) deviates from the truth only by a "constant multiple".

Conjecture 2. *Let $k > 0$. If $\Delta \sim k\sqrt{\log\log T}$, then*

$$\frac{1}{T}\cdot \underset{t\in[T;2T]}{meas}\left\{\frac{\log|\zeta(\frac{1}{2}+\mathrm{i}t)|}{\sqrt{\frac{1}{2}\log\log T}}\geqslant \Delta\right\}\sim C_k\cdot\int_{\Delta}^{\infty}e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}}$$

with C_k the same constant as in (5)

An abelian argument (conditional on RH and using [12] to truncate the tails) shows that if the left-hand side of (4) is at all asymptotic to a constant κ times a standard Gaussian in the range $\Delta \sim k(\frac{1}{2}\log\log T)^{1/2}$ then $\int_T^{2T}|\zeta(\frac{1}{2}+\mathrm{i}t)|^{2k}\mathrm{d}t \sim \kappa T(\log T)^{k^2}$, hence $\kappa = C_k$. Thus there is only one reasonable choice for the constant in Conjecture 2 above.

For negative values of k Conjecture 2 is likely to be false as soon as $k \leqslant -1/2$. Indeed in that range the zeros of $\zeta(s)$ should force a transition to an exponential distribution, see [4] for a precise statement.

1.1. **Remarks.**

(1) There is no difficulty in adapting our proof to the study of negative values of $\log|\zeta(\frac{1}{2}+\mathrm{i}t)|$. Also, our argument carries over to $S(t) := \frac{1}{\pi}\Im\log\zeta(\frac{1}{2}+\mathrm{i}t)$ with little to no changes.
(2) Conditionally on the Riemann Hypothesis we obtain the better range $\Delta \leqslant (\log\log T)^{1/6-\varepsilon}$ for $S(t)$. (assuming the Riemann Hypothesis, the analogue of Lemma 1 for $S(t)$ has an k^{2k} instead of k^{4k}. This is responsible for the improvement).
(3) Assuming the relevant Riemann Hypothesis our result extends to elements of the Selberg class (see [11] for a definition).
(4) Finally, our method covers the case of the joint distribution of $\log|\zeta(\frac{1}{2}+\mathrm{i}t)|$ and $\Im\log\zeta(\frac{1}{2}+\mathrm{i}t)$. Similarly, one can recover Bourgade's [2] result on the joint distribution of shifts of the Riemann zeta function in a large deviation setting.
(5) The term k^{4k} in Lemma 1 is directly responsible the range $\Delta \ll (\log\log T)^{1/10-\varepsilon}$ in Theorem 1. For example, an improvement of the k^{4k} to k^{2k} in Lemma 1 would give Theorem 1 uniformly in $\Delta \ll (\log\log T)^{1/6-\varepsilon}$.

1.2. **Outline of the proof of Proposition 1.** The key idea in the proof of Proposition 1 is to initially work on a subset A of $[T;2T]$ on which the Dirichlet polynomial does not attain large values (say $\leq c\cdot\log\log T$). This allows us to express the moment generating function,

$$\int_A \exp\left(z\cdot\Re\sum_{p\leqslant x}\frac{1}{p^{1/2+\mathrm{i}t}}\right)\cdot\mathrm{d}t \quad,\quad |z|\ll 1 \tag{6}$$

in terms of only the first $\asymp \log\log T$ moments of $\Re\sum_{p\leqslant x}p^{-1/2-\mathrm{i}t}$ over $t\in A$. These moments can be taken over the full interval $[T;2T]$ (since A is very close in measure to T) and then they become easy to estimate. On adding up the contribution from the moments we obtain a very precise estimate for (6). From there, by standard probabilistic techniques, we obtain (3) with t restricted to A. Since A is very close in measure to T, we get (3) without the restriction to $t\in A$.

Acknowledgments. The author would like to thank his supervisor K. Soundararajan for advice and encouragements.

Notation. Throughout ε will denote an arbitrary small but fixed positive real. We allow ε to differ from line to line. Finally $\log_k T$ denotes the k-th iterated natural logarithm, so that $\log_k := \log\log_{k-1}T$ and $\log_1 T = \log T$.

2. Lemmata

Lemma 1 (Selberg, [15]). *Uniformly in $k \geqslant 0$,*

$$\int_T^{2T} \left| \log|\zeta(\tfrac{1}{2} + \mathrm{i}t)| - \Re \sum_{p \leqslant x} \frac{1}{p^{\frac{1}{2} + \mathrm{i}t}} \right|^{2k} \mathrm{d}t \ll A^k \cdot k^{4k} + A^k \cdot k^k \cdot (\log\log\log T)^k$$

where $x = T^{1/(\log\log T)^2}$ and $A > 0$ constant.

Proof. See Tsang's thesis [15], page 60. □

Lemma 2 (Hwang, [5]). *Let W_n be a sequence of distribution functions such that for $|s| \leqslant \varepsilon$,*

$$\int_{-\infty}^{\infty} e^{st} \mathrm{d}W_n(t) = F(s) \exp\left(\frac{s^2}{2} \cdot \phi(n) \right) \cdot \left(1 + O\left(\kappa_n^{-1}\right)\right)$$

with $F(s)$ a function analytic around $s = 0$ and equal to 1 at $s = 0$. Then, uniformly in $\Delta \leqslant o(\min(\kappa_n, \sqrt{\phi(n)})$,

$$1 - W_n(\Delta\sqrt{\phi(n)}) \sim \int_{\Delta}^{\infty} e^{-u^2/2} \cdot \frac{\mathrm{d}u}{\sqrt{2\pi}}$$

Proof. This is a special case of Theorem 1 in [5]. □

Lemma 3 (Soundararajan, [12]). *For $x \leqslant T^{1/4k}$,*

$$\int_T^{2T} \left| \Re \sum_{p \leqslant x} \frac{1}{p^{\frac{1}{2} + \mathrm{i}t}} \right|^{2k} \mathrm{d}t \ll k! \cdot T \left(\sum_{p \leqslant x} \frac{1}{p} \right)^k$$

Proof. This is a special case of Lemma 3 in [12] □

.

3. Proof of Proposition 1

Lemma 4. *Let $p_1, \ldots, p_k \leq x$ be primes. Then,*

$$\frac{1}{T} \int_T^{2T} \prod_{\ell=1}^{k} \cos(t \log p_\ell) \mathrm{d}t = f(p_1 \cdot \ldots \cdot p_k) + O(2^k x^k)$$

where f is a multiplicative function defined by $f(p^\alpha) = \frac{1}{2^\alpha}\binom{\alpha}{\alpha/2}$ By convention the binomial coefficient is zero when $\alpha/2$ is not an integer.

Proof. Write $n = p_1 \cdot \ldots \cdot p_k = q_1^{\alpha_1} \cdot \ldots q_r^{\alpha_r}$ with q_i mutually distinct primes. Notice that

$$\begin{aligned} \cos(t \log q_i)^{\alpha_i} &= \frac{1}{2^{\alpha_i}} \cdot \left(e^{\mathrm{i}t \log q_i} + e^{-\mathrm{i}t \log q_i}\right)^{\alpha_i} \\ &= \frac{1}{2^{\alpha_i}} \binom{\alpha_i}{\alpha_i/2} + \sum_{\alpha_i/2 \neq \ell \leq \alpha_i} \frac{1}{2^{\alpha_i}} \binom{\alpha_i}{\ell} e^{\mathrm{i}(\alpha_i - 2\ell)t \log q_i} \end{aligned}$$

Note that the leading coefficients in front of every $e^{\mathrm{i}(\alpha_i - 2\ell)}$ is ≤ 1 in absolute value. Furthermore $|\alpha_i - 2\ell| \leq \alpha_i$. Therefore

$$\frac{1}{T} \int_T^{2T} \prod_{\ell=1}^{k} \cos(t \log p_\ell) \mathrm{d}t = \frac{1}{T} \int_T^{2T} \prod_{i=1}^{r} \cos(t \log q_i)^{\alpha_i} \mathrm{d}t = f(n) + (\ldots)$$

where inside $(\ldots)$ there is a sum of $\leq (\alpha_1+1)\cdot\ldots\cdot(\alpha_r+1) \leq 2^{\alpha_1}\cdot\ldots\cdot 2^{\alpha_r} = 2^k$ terms of the form $\gamma e^{\mathrm{i}t(\beta_1 \log q_1 + \ldots \beta_r \log q_r)}$ with $|\gamma| \leq 1$ and integers $|\beta_i| \leq \alpha_i$. Since $\beta_1 \log q_1 + \ldots + \beta_r \log q_r \gg x^{-k}$ the integral over $T \leq t \leq 2T$ of each of these terms is at most $\ll x^k$. Since there is $\leq 2^k$ of these terms and each has a leading coefficient $|\gamma| \leq 1$ the integral over $T \leq t \leq 2T$ of all the terms inside $(\ldots)$ contributes at most $O(2^k x^k)$. □

Lemma 5. *Let $x = T^{1/(\log\log T)^2}$. Then,*

$$\frac{1}{T}\int_T^{2T}\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k \mathrm{d}t = \frac{1}{2\pi\mathrm{i}}\oint\prod_{p\leqslant x} I_0\left(\frac{z}{\sqrt{p}}\right)\cdot\frac{\mathrm{d}z}{z^{k+1}} + O\left((2x)^k\right)$$

where $I_0(z) = (1/\pi)\int_0^{\pi} e^{z\cos\theta}\mathrm{d}\theta = \sum_{n\geqslant 0}(z/2)^{2n}\cdot 1/(n!)^2$ is the modified 0-th order Bessel function.

Proof. Given an integer $n = p_1^{\alpha_1}\cdot\ldots\cdot p_\ell^{\alpha_\ell}$ with $\Omega(n) = k$ and $p_i \leqslant x$ there are $k!/(\alpha_1!\ldots\alpha_\ell!)$ ways in which this integer can be written as a product of k primes. We define a multiplicative function $g(n)$ by $g(p^\alpha) = 1/\alpha!$ and $g(p^\alpha) = 0$, $p > x$, so that $k!g(n)$ is equal to the number of ways in which an integer n with $\Omega(n) = k$ can be expressed as a product of k primes $\leqslant x$.

By Lemma 1 and the above observation,

$$\begin{aligned}\frac{1}{T}\int_T^{2T}\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k \mathrm{d}t &= \sum_{p_1,\ldots,p_k\leqslant x}\frac{f(p_1\cdot\ldots\cdot p_k)}{\sqrt{p_1\cdot\ldots\cdot p_k}} + O\left((2x)^{2k}\right)\\ &= k!\sum_{\Omega(n)=k}\frac{f(n)}{\sqrt{n}}\cdot g(n) + O\left((2x)^{2k}\right)\end{aligned}$$

We detect the condition $\Omega(n) = k$ by using Cauchy's integral formula. Thus the above is equal to

$$\frac{k!}{2\pi\mathrm{i}}\oint\sum_{n\geqslant 1}\frac{f(n)}{\sqrt{n}}\cdot g(n)z^{\Omega(n)}\cdot\frac{\mathrm{d}z}{z^{k+1}} + O\left((2x)^{2k}\right)$$

Since the functions f, g and $z^{\Omega(n)}$ are multiplicative, the above sum over $n \geqslant 1$ does factor into an Euler product,

$$\prod_{p\leqslant x}\left(1 + \sum_{\ell\geqslant 1}\left(\frac{z}{2\sqrt{p}}\right)^{2\ell}\frac{1}{(2\ell)!}\cdot\binom{2\ell}{\ell}\right) = \prod_{p\leqslant x} I_0\left(\frac{z}{\sqrt{p}}\right)$$

as desired. □

We are now ready to prove Proposition 1.

Proof of Proposition 1. By Lemma 3,

$$\int_T^{2T}\left|\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right|^{2k}\mathrm{d}t \ll Tk!\cdot\left(\sum_{p\leqslant x}\frac{1}{p}\right)^k$$

Therefore the set of those $t \in [T; 2T]$ for which $|\Re\sum_{p\leqslant x} p^{-\frac{1}{2}-\mathrm{i}t}| \geqslant \log_2 T$ has measure $\ll T(k/\log_2 T)^k$. Choosing $k = \lfloor \log_2 T/e \rfloor$, this measure is $\ll T(\log T)^{-\delta}$ where $\delta = 1/e$, which is negligible (also, the exact value of δ is unimportant). Thus, we can restrict our attention to the set

$$A := \left\{t \in [T; 2T] : \left|\Re\sum_{p\leq x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right| \leq \log_2 T\right\}$$

We denote its complement in $[T; 2T]$ by A^c. For complex $|z| \leqslant \varepsilon < \frac{1}{100}$,

$$\int_A \exp\left(z\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)\mathrm{d}t = \sum_{k\leqslant 3\mathcal{V}}\frac{z^k}{k!}\int_A\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k\mathrm{d}t + O(T(\log T)^{-3}) \tag{7}$$

where $\mathcal{V} = \log\log T$. The error term arises from bounding the terms with $k \geqslant 3\mathcal{V}$: each contributes at most $T(\varepsilon\log\log T)^k/k! \leq Te^{-k}$ since the integral over the set A is less than $T(\log\log T)^k$ by definition of A. To compute the moments with $k \leqslant 3\mathcal{V}$ we use Cauchy's inequality and notice that

$$\left|\int_A\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k\mathrm{d}t - \int_T^{2T}\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k\mathrm{d}t\right| \leqslant \tag{8}$$

$$\leqslant \sqrt{\mathrm{meas}\,(A^c)}\cdot\left(\int_T^{2T}\left|\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right|^{2k}\mathrm{d}t\right)^{\frac{1}{2}} \ll T\cdot(\log T)^{-\delta/2}\cdot\sqrt{k!}\cdot(\log\log T)^{k/2}$$

by Lemma 3. The integral over $T \leqslant t \leqslant 2T$ is readily available through Lemma 5. Thus,

$$\int_A\left(\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)^k\mathrm{d}t = \frac{Tk!}{2\pi\mathrm{i}}\oint\prod_{p\leqslant x}I_0\left(\frac{w}{\sqrt{p}}\right)\cdot\frac{\mathrm{d}w}{w^{k+1}} + O\left(\frac{T\sqrt{k!}\cdot(\log\log T)^{k/2}}{(\log T)^{\delta/2}}\right)$$

We choose the contour to be a circle of radius $r = 2$ (say) around the origin. Summing the above over $k \leqslant 3\mathcal{V}$ we conclude that (7) is equal to

$$\frac{T}{2\pi\mathrm{i}}\oint\prod_{p\leqslant x}I_0\left(\frac{w}{\sqrt{p}}\right)\sum_{k\leqslant 3\mathcal{V}}\frac{z^k}{w^{k+1}}\cdot\mathrm{d}w + O\left(T(\log T)^{-\delta/2+\varepsilon}\right)$$

Since $\prod_{p\leqslant x}I_0(w/\sqrt{p}) \ll (\log T)^2$ on $|w| = 2$, and $\sum_{k\geqslant 3\mathcal{V}}|z/w|^k \ll (\log T)^{-3\log 2+\varepsilon}$ we can complete the sum over $k \leqslant 3\mathcal{V}$ to all k, making an error of $T(\log T)^{-\delta/5}$. Thus the above (and hence (7)) is equal to

$$\frac{T}{2\pi\mathrm{i}}\oint\prod_{p\leqslant x}I_0\left(\frac{w}{\sqrt{p}}\right)\cdot\frac{\mathrm{d}w}{w-z} + O\left(T(\log T)^{-\delta/5}\right) = T\cdot\prod_{p\leqslant x}I_0\left(\frac{z}{\sqrt{p}}\right) + O\left(T(\log T)^{-\delta/5}\right) \tag{9}$$

by Cauchy's formula. But, $\prod_{p\leqslant x}I_0\left(\frac{z}{\sqrt{p}}\right) = F(z)\cdot\exp\left(\frac{z^2}{2}\cdot\frac{1}{2}\sum_{p\leq x}\frac{1}{p}\right)\cdot(1+O(1/\log T))$ with $F(z)$ analytic around $z = 0$ and equal to 1 at $z = 0$. Since $|z| \leqslant \varepsilon$ it is clear that the error term in (9) can be made relative at a small cost in the error term. We conclude that

$$\int_A\exp\left(z\cdot\Re\sum_{p\leq x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right)\mathrm{d}t = TF(z)\cdot\exp\left(\frac{z^2}{2}\cdot\frac{1}{2}\sum_{p\leq x}\frac{1}{p}\right)\cdot(1+O((\log T)^{-\delta/10})$$

uniformly in $|z| \leq \varepsilon$. Thus Lemma 2 is applicable, and it follows that

$$\frac{1}{\mathrm{meas}(A)}\cdot\operatorname*{meas}_{t\in A}\left\{\Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \geqslant \Delta\cdot\sqrt{\frac{1}{2}\sum_{p\leqslant x}\frac{1}{p}}\right\} \sim \frac{T}{meas(A)}\int_\Delta^\infty e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}}$$

Now $\mathrm{meas}(A^c) \ll T(\log T)^{-\delta}$ while $\int_{\Delta}^{\infty} e^{-u^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}} \gg (\log T)^{-o(1)}$ (because $\Delta = o(\sqrt{\log\log T})$. Hence the preceding equation becomes

$$\frac{1}{T}\cdot \underset{t\in[T;2T]}{\mathrm{meas}} \left\{ \Re \sum_{p\leqslant x} \frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \geqslant \Delta \cdot \sqrt{\tfrac{1}{2}\sum_{p\leqslant x}\frac{1}{p}} \right\} \sim \int_{\Delta}^{\infty} e^{-u^2/2}\cdot \frac{\mathrm{d}u}{\sqrt{2\pi}}$$

This establishes our claim □

4. Proof of Theorem 1.

Proof of Theorem 1. Let $x = T^{1/(\log\log T)^2}$. Let $1/4 > \delta > 0$ to be fixed later. By Lemma 1 the set of those $t \in [T;2T]$ for which

$$\left| \log|\zeta(\tfrac{1}{2}+\mathrm{i}t)| - \Re\sum_{p\leqslant x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \right| \geqslant \mathcal{L} := (\log\log T)^{\frac{1}{4}+\delta}$$

has measure $\ll T\cdot \mathcal{L}^{-2k}\cdot A^k\cdot k^{4k}\cdot(\log_3 T)^k$ for all $k \geq \log_3 T$. Choosing $k = \lfloor \mathcal{L}^{1/2}/Ae\rfloor$ we obtain a measure of $\ll T\exp(-c\mathcal{L}^{1/2})$ for some constant $c > 0$. Thus except for a set of measure $\ll T\exp(-c\mathcal{L}^{1/2})$,

$$\log|\zeta(\tfrac{1}{2}+\mathrm{i}t)| = \Re\sum_{p\leq x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} + \theta\mathcal{L}$$

with a $|\theta| \leqslant 1$ depending on t. The measure of those $t\in[T;2T]$ for which

$$\Re\sum_{p\leq x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \geqslant \Delta\cdot\sqrt{\tfrac{1}{2}\cdot\log\log T} + \theta\mathcal{L} \tag{10}$$

is (since $\theta \leq 1$) at least the measure of those $t\in[T;2T]$ for which

$$\Re\sum_{p\leq x}\frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \geq \Delta\cdot\sqrt{\tfrac{1}{2}\log\log T} + \mathcal{L}$$

and this measure is at least,

$$T(1+o(1))\int_{\Delta+2\mathcal{L}/\sqrt{\log\log T}}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}} \tag{11}$$

by Proposition 1. Similarly (using this time $\theta \geq -1$) the measure of those $t\in[T;2T]$ for which (10) holds is at most

$$T(1+o(1))\int_{\Delta-2\mathcal{L}/\sqrt{\log\log T}}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}}. \tag{12}$$

When $\Delta\mathcal{L} = o((\log\log T)^{1/2})$ both (11) and (12) are equal to $T\int_{\Delta}^{\infty} e^{-u^2/2}\cdot\frac{\mathrm{d}u}{\sqrt{2\pi}}\cdot(1+o(1))$. We conclude that,

$$\frac{1}{T}\cdot\underset{t\in[T;2T]}{\mathrm{meas}}\left\{ \frac{\log|\zeta(\frac{1}{2}+\mathrm{i}t)|}{\sqrt{\frac{1}{2}\log\log T}} \geqslant \Delta \right\} = \frac{(1+o(1))}{\sqrt{2\pi}}\int_{\Delta}^{\infty} e^{-u^2/2}\mathrm{d}u + O\left(e^{-c\mathcal{L}^{1/2}}\right)$$

provided that $\Delta\mathcal{L} = o(\log\log T)^{1/2}$. The main term dominates the error term as long as $\Delta \leqslant \mathcal{L}^{1/4-\varepsilon}$. Both conditions $\Delta\leqslant\mathcal{L}^{1/4-\varepsilon}$ and $\Delta\mathcal{L} = o(\log\log T)^{1/2}$ are met if we choose $\delta = 3/20$ and require that $\Delta \leqslant (\log\log T)^{1/16+\delta/4-\varepsilon} = (\log\log T)^{1/10-\varepsilon}$. □

5. Proof of Proposition 2

We will only give a brief sketch.

Lemma 6. *Let $C > 0$ be given. Then, uniformly in $0 \leq \Re z \leq C$, and $4C^2 \leq |\Im z|^2 \leq x^{1/8}$,*

$$\prod_{p \leq x} I_0\left(\frac{z}{\sqrt{p}}\right) \ll \exp(-c(\Im z)^2)$$

for some constant $c > 0$.

Proof. For $|z|^2 \leq x$ we write

$$\prod_{p \leq x} I_0\left(\frac{z}{\sqrt{p}}\right) = \prod_{p \leq |z|^2} I_0\left(\frac{z}{\sqrt{p}}\right) \cdot \prod_{|z|^2 \leq p \leq x} I_0\left(\frac{z}{\sqrt{p}}\right) e^{-(z/2\sqrt{p})^2} \cdot \exp\left(-\frac{z^2}{2} \cdot \frac{1}{2} \sum_{|z|^2 \leq p \leq x} \frac{1}{p}\right)$$

Using the bound $I_0(z) \ll \exp(-|\Im z|)$ the first term is $\ll \exp(-c|z|^{3/2}/\log|z|)$ Using the bound $I_0(z)e^{-(z/2)^2} \ll \exp(|z|^4)$ we get that the contribution of the second term is $\ll \exp(c|z|^2/\log|z|)$. Finally the third term is a Gaussian and thus contributes $\ll \exp((1/2)((\Re z)^2 - (\Im z)^2)(\log(\log x/2\log|z|))$. Under our assumptions on z and x the last bound dominates and the claim follows. □

Proof of Proposition 2. By a small modification of the proof of Proposition 1, for any fixed $\delta > 0$, uniformly in z,

$$\int_A \exp\left(z \cdot \Re \sum_{p \leq x} \frac{1}{p^{\frac{1}{2}+\mathrm{i}t}}\right) \cdot \mathrm{d}t = T \prod_{p \leq x} I_0\left(\frac{z}{\sqrt{p}}\right) + O((T\log T)^{-k^2-1})$$

with the set A and the function $F(\cdot)$ as in the proof of Proposition 1, except that now A is the set of those t's at which the Dirichlet polynomial $\Re \sum_{p \leq x} p^{-1/2-\mathrm{i}t} \leq C \log\log T$ with some C large enough but fixed. The constant C is choosen so large so as to guarantee an error of $T(\log T)^{-k^2-1}$ above and $\mathrm{meas}(A^c) \ll T(\log T)^{-k^2-1}$.

Let $H(z;x) = \prod_{p \leq x} I_0(z/\sqrt{p}) \exp(-z\Delta\sigma(x))$ where $\sigma(x) = (\frac{1}{2}\sum_{p \leq x} \frac{1}{p})^{1/2}$. By Tenenbaum's formula (see equation (10) and above in [14]),

$$\text{(13)}\quad \underset{t \in A}{\mathrm{meas}} \left\{ \Re \sum_{p \leq x} \frac{1}{p^{\frac{1}{2}+\mathrm{i}t}} \geq \Delta \sqrt{\frac{1}{2}\sum_{p \leq x}\frac{1}{p}} \right\} = \frac{\mathrm{meas}(A)}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} H(s;x) \frac{T\mathrm{d}s}{s(s+T)} + \\ + O\left(\left|\frac{T}{2\pi}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{H(s;x)T\mathrm{d}s}{(s+T)(s+2T)}\right|\right) + O(T(\log T)^{-k^2-1})$$

where $T = x^{1/16}$ and $c = \Delta \cdot (\sigma(x))^{-1/2}$. By the previous Lemma (used to handle the range $\Im s < x^{1/8}$ and the fast decay of $1/s(s+T)$ (used to handle $\Im s > x^{1/8}$) we can truncate the above integral at $|\Im s| \asymp \psi(x)/\sigma(x)$ with $\psi(x) \to \infty$ very slowly. This induces an negligible error term of $o(\int_\Delta^\infty e^{-u^2/2}\mathrm{d}u)$. Furthermore, since T is so large, the $O(\cdot)$ term in the above equation is also negligible. Write $\prod_{p \leq x} I_0(z/\sqrt{p}) = F(z) \cdot \exp(\frac{z^2}{2} \cdot \sigma(x)^2)$ with $F(z)$ analytic. Then, on completing the square and changing variables the first integral above (truncated at $|\Im s| \asymp \psi(x)/\sigma(x)$) can be rewritten as

$$\mathrm{meas}(A) \cdot \frac{e^{-\Delta^2/2}}{\sigma(x)} \cdot \frac{1}{2\pi} \int_{-\psi(x)}^{\psi(x)} F\left(c + \frac{\mathrm{i}v}{\sigma(x)}\right) \frac{e^{-v^2/2}\mathrm{d}v}{c + \mathrm{i}v/\sigma(x)}$$

Let $G(z) = F(z)/z$. Then $G(c + \mathrm{i}v/\sigma(x)) = G(c) + \mathrm{i}vG'(c)/\sigma(x)O(v^2/\sigma(x)^2)$. Plugging this into the integral the middle term vanishes while the last term contributes a negligible amount. Thus, we end up with

$$F(c) \cdot \frac{e^{-\Delta^2/2}}{\sqrt{2\pi}\Delta} \cdot (1 + o(1)) = F(c) \int_{\Delta}^{\infty} e^{-u^2/2} \frac{\mathrm{d}u}{\sqrt{2\pi}} (1 + o(1))$$

as a main term for (13). To conclude it suffices to notice that $F(c) \sim F(k)$ as $T \to \infty$, since $c \sim k$ and F is analytic. Furthermore as in the proof of Proposition 1 the restriction to $t \in A$ in (13) can be replaced by $t \in [T; 2T]$ because A is very close in measure to T (that is, $\mathrm{meas}(A^c) \ll T(\log T)^{-k^2-1}$).

□

References

[1] H. Bohr and B. Jessen. Uber die wertverteilung der Riemannschen zetafunktion, erste mitteilung. *Acta. Math.*, 50:1 – 35, 1930.

[2] P. Bourgade. Mesoscopic fluctuations of the zeta zeros. *Probab. Theory Related Fields*, 148 no. 3-4:479 – 500, 2010.

[3] T. Hattori and K. Matsumoto. A limit theorem for bohr-jessens probability measures of the riemann zeta-function. *J. Reine Angew. Math.*, 507:219232, 1999.

[4] C. P. Hughes. *On the characteristic polynomial of a random unitary matrix and the Riemann zeta function*. PhD thesis, University of Bristol, 2001.

[5] H. H. Hwang. Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.*, 6 (1):297 – 319, 1996.

[6] Y. Lamzouri. On the distribution of extreme values of zeta and L-functions in the strip $1/2 < \sigma < 1$. *Internat. Math. Res. Notices*, to appear.

[7] A. Laurinchikas. *Limit theorems for the Riemann Zeta-function*. Kluwer, Dordrecht.

[8] A. Laurinchikas. Distribution of trigonometric polynomials. *Litovsk. Mat. Sb.*, 21, no. 2:127 – 135, 1981.

[9] A. Laurinchikas. Large deviations of trigonometric polynomials. *Litovsk. Mat. Sb.*, 21, no. 3:52 – 61, 1981.

[10] A. Selberg. Contributions to the theory of the Riemann zeta-function. *Arch. Math. Naturvid.*, 48, no. 5:89 – 155, 1946.

[11] A. Selberg. Old and new conjectures and results about a class of Dirichlet series. *Proceedings of the Amalfi Conference on Analytic Number Theory*, pages 367 – 385, 1992.

[12] K. Soundararajan. Moments of the Riemann zeta function. *Annals of Math.*, 170 (2):981 – 993, 2009.

[13] J. Steuding. *Value distribution of L-functions*. Springer-Verlag, 2007.

[14] G. Tenenbaum. Sur la distribution conjointe de deux fonctions "nombre de facteurs premiers". *Aeq. Math.*, 35:55–68, 1988.

[15] K. M. Tsang. *The distribution of the values of the Riemann zeta function*. PhD thesis, Princeton University, Princeton, NJ, 1984.

Department of Mathematics, Stanford University, 450 Serra Mall, Bldg. 380, Stanford, CA 94305-2125

E-mail address: `maksym@stanford.edu`