Discrepancy Bounds for the Riemann Zeta Function

Anurag Sahay

PhD Oral Exam
University of Rochester

asahay@ur.rochester.edu

4th December 2020

References

Youness Lamzouri, Stephen Lester, and Maksym Radziwiłł.
Discrepancy bounds for the distribution of the Riemann zeta-function and applications.
J. Anal. Math., 139(2):453–494, 2019.

Harald Bohr and Börge Jessen,
Über die Werteverteilung der Riemannschen Zetafunktion, erste Mitteilung.
Acta Math. 54 (1930), no. 1, 1–35. MR1555301

E.C. Titchmarsh and D.R. Heath-Brown.
The theory of the Riemann zeta-function.
Oxford University Press, 1986.

Overview of the Talk

1. Introduction
   - The Limiting Distribution of $\log \zeta(s)$
   - Euler Product for the Riemann Zeta Function
   - The Random Model for the Riemann Zeta Function

2. Main Results
   - Discrepancy for $\log \zeta$
   - The Characteristic Function

3. Sketch of the Proof
   - Approximating Characteristic Functions
   - Beurling-Selberg Functions
   - Concluding the Discrepancy Bound

Notation and Preliminaries

In this talk, we use some notation divergent from [LLR19], in order to emphasize the probabilistic ideas behind their paper.

- We view $[T, 2T]$ as a probability space with the normalized Lebesgue measure, which we denote by $\mathbb{P}_T(\cdot)$.

Notation and Preliminaries

In this talk, we use some notation divergent from [LLR19], in order to emphasize the probabilistic ideas behind their paper.

- We view $[T, 2T]$ as a probability space with the normalized Lebesgue measure, which we denote by $\mathbb{P}_T(\cdot)$.
- We write $\mathbb{E}_T(\cdot)$ for the expectation against that measure. Clearly,

$$\mathbb{E}_T[f(t)] = \frac{1}{T}\int_T^{2T} f(t)\,dt$$

Notation and Preliminaries

In this talk, we use some notation divergent from [LLR19], in order to emphasize the probabilistic ideas behind their paper.

- We view $[T, 2T]$ as a probability space with the normalized Lebesgue measure, which we denote by $\mathbb{P}_T(\cdot)$.
- We write $\mathbb{E}_T(\cdot)$ for the expectation against that measure. Clearly,

$$\mathbb{E}_T[f(t)] = \frac{1}{T}\int_T^{2T} f(t)\,dt$$

- We will use $\mathbb{P}(\cdot)$ for probability and $\mathbb{E}(\cdot)$ for expectation associated with other sources of randomness.

Notation and Preliminaries

In this talk, we use some notation divergent from [LLR19], in order to emphasize the probabilistic ideas behind their paper.

- We view $[T, 2T]$ as a probability space with the normalized Lebesgue measure, which we denote by $\mathbb{P}_T(\cdot)$.
- We write $\mathbb{E}_T(\cdot)$ for the expectation against that measure. Clearly,

$$\mathbb{E}_T[f(t)] = \frac{1}{T}\int_T^{2T} f(t)\,dt$$

- We will use $\mathbb{P}(\cdot)$ for probability and $\mathbb{E}(\cdot)$ for expectation associated with other sources of randomness.
- Any limit of random variables in this talk will be in the sense of convergence in distribution.

How is $\log \zeta(\sigma + it)$ distributed for large t?

Fix $\sigma \in \mathbb{R}$. Then the map

$$t \mapsto \log \zeta(\sigma + it)$$

is a $\mathbb{C}$-valued random variable on $[T, 2T]$.

How is $\log \zeta(\sigma + it)$ distributed for large t?

Fix $\sigma \in \mathbb{R}$. Then the map

$$t \mapsto \log \zeta(\sigma + it)$$

is a $\mathbb{C}$-valued random variable on $[T, 2T]$.

A classical question in analytic number theory is the following: what is the distribution of this random variable for large T?

How is $\log \zeta(\sigma + it)$ distributed for large t?

Fix $\sigma \in \mathbb{R}$. Then the map

$$t \mapsto \log \zeta(\sigma + it)$$

is a $\mathbb{C}$-valued random variable on $[T, 2T]$.

A classical question in analytic number theory is the following: what is the distribution of this random variable for large T?

This question amounts to asking what is the distributional limit as $T \to \infty$ of the random variables

$$\{t \mapsto \log \zeta(\sigma + it) : t \in [T, 2T]\}_{T>0},$$

if it exists.

How is $\log \zeta(\sigma + it)$ distributed for large t?

In light of the functional equation, we restrict our discussion to $\sigma \geq 1/2$.

How is $\log\zeta(\sigma + it)$ distributed for large t?

In light of the functional equation, we restrict our discussion to $\sigma \geq 1/2$.

The case $\sigma > 1$ is relatively easy, since $\zeta(s)$ has a convergent Euler product and series representation.

How is $\log \zeta(\sigma + it)$ distributed for large t?

In light of the functional equation, we restrict our discussion to $\sigma \geq 1/2$.

The case $\sigma > 1$ is relatively easy, since $\zeta(s)$ has a convergent Euler product and series representation.

The case $\sigma = 1/2$ was considered by Selberg, who proved his Central Limit Theorem: loosely, it says that $\log|\zeta(1/2 + it)|$ is normally distributed with mean 0 and variance $\frac{1}{2}\log\log T$. We will not discuss this further.

How is $\log \zeta(\sigma + it)$ distributed for large t?

In light of the functional equation, we restrict our discussion to $\sigma \geq 1/2$.

The case $\sigma > 1$ is relatively easy, since $\zeta(s)$ has a convergent Euler product and series representation.

The case $\sigma = 1/2$ was considered by Selberg, who proved his Central Limit Theorem: loosely, it says that $\log |\zeta(1/2 + it)|$ is normally distributed with mean 0 and variance $\frac{1}{2} \log \log T$. We will not discuss this further.

Bohr and Jessen [BJ30] showed that if $\sigma > 1/2$, then the limiting distribution exists and is continuous. The main result of [LLR19] is an estimate on the rate of this convergence in the regime $1/2 < \sigma \leq 1$.

How is $\log \zeta(\sigma + it)$ distributed for large t?

In light of the functional equation, we restrict our discussion to $\sigma \geq 1/2$.

The case $\sigma > 1$ is relatively easy, since $\zeta(s)$ has a convergent Euler product and series representation.

The case $\sigma = 1/2$ was considered by Selberg, who proved his Central Limit Theorem: loosely, it says that $\log |\zeta(1/2 + it)|$ is normally distributed with mean 0 and variance $\frac{1}{2} \log \log T$. We will not discuss this further.

Bohr and Jessen [BJ30] showed that if $\sigma > 1/2$, then the limiting distribution exists and is continuous. The main result of [LLR19] is an estimate on the rate of this convergence in the regime $1/2 < \sigma \leq 1$.

For simplicity of exposition, we will not consider $\sigma = 1$ in this talk, although the same ideas apply, and are treated in [LLR19].

Euler Product for the Riemann Zeta Function

For $\sigma > 1$, we have the following convergent product formula for the Riemann zeta function due to to Euler:

$$\zeta(s) = \prod_p \left(\frac{1}{1 - \frac{1}{p^s}} \right)$$

Euler Product for the Riemann Zeta Function

For $\sigma > 1$, we have the following convergent product formula for the Riemann zeta function due to to Euler:

$$\zeta(s) = \prod_p \left(\frac{1}{1 - \frac{1}{p^s}} \right)$$

Putting $s = \sigma + it$, and rearranging a bit, we get that

$$\zeta(\sigma + it) = \prod_p \left(\frac{1}{1 - \frac{p^{-it}}{p^{\sigma}}} \right)$$

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
    - For fixed n, and large T, they are distributed approximately uniformly.

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
  - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
  - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?
  - 2^{-it} and 4^{-it}?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
    - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?
    - 2^{-it} and 4^{-it}?
    - $2^{-it}, 3^{-it}$ and 6^{-it}?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
  - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?
  - 2^{-it} and 4^{-it}?
  - $2^{-it}, 3^{-it}$ and 6^{-it}?
  - 2^{-it} and 3^{-it}?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
  - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?
  - 2^{-it} and 4^{-it}?
  - $2^{-it}, 3^{-it}$ and 6^{-it}?
  - 2^{-it} and 3^{-it}? 5^{-it} and 6^{-it}?

The Behaviour of p^{-it}

For $t \in \mathbb{R}$, we have that $n^{-it} \in \mathbb{T} \subseteq \mathbb{C}$, for $n \in \mathbb{N}$. n^{-it} is clearly distributed uniformly on $\mathbb{T}$ for every n.

What happens when $t \in [T, 2T]$?

- How do these behave individually?
  - For fixed n, and large T, they are distributed approximately uniformly.
- How do they interact for different values of n?
  - 2^{-it} and 4^{-it}?
  - $2^{-it}, 3^{-it}$ and 6^{-it}?
  - 2^{-it} and 3^{-it}? 5^{-it} and 6^{-it}?

  In general there is no reason to expect m^{-it} and n^{-it} to show any sort of relationship when m and n are coprime.

The Behaviour of p^{-it}

The heuristic for why m^{-it} and n^{-it} should behave independently when $(m, n) = 1$ comes from the following theorem from harmonic analysis:

Theorem (Kronecker-Weyl)

Let $\theta_1, \cdots, \theta_n \in \mathbb{R}$ be linearly independent over $\mathbb{Q}$. Then the set

$$\{(e(\theta_1 x), \cdots, e(\theta_n x) : x \in \mathbb{R}\}$$

is equidistributed on $\mathbb{T}^n$, where $e(\cdot) = e^{2\pi i(\cdot)}$ as usual.

Note that $\{1\} \cup \{\log p : p \text{ prime}\}$ is $\mathbb{Q}$-linearly independent – this is the fundamental theorem of arithmetic. We conclude that any finite subset of pair-wise coprime integers should behave independently.

The Definition of the Random Model $\zeta(\sigma, X)$

This behaviour of p^{-it} as approximately uniform and i.i.d. random variables on $[T, 2T]$ leads to the following definition:

Definition (Random Model)

Let X be random variable uniformly taking values in $\mathbb{T}^\infty$, indexed by the primes. In other words, $X = \{X(p)\}_p$ is a family of independent random variables uniformly distributed on the unit circle in $\mathbb{C}$, indexed by the primes. We define the $\mathbb{C}$-valued random variable $\zeta(\sigma, X)$ as follows

$$\zeta(\sigma, X) = \prod_p \left(\frac{1}{1 - \frac{X(p)}{p^\sigma}} \right)$$

The Random Model $\zeta(\sigma, X)$

Definition (Random Model)

Let X be random variable uniformly taking values in $\mathbb{T}^\infty$, indexed by the primes. Then, we define,

$$\zeta(\sigma, X) = \prod_p \left(\frac{1}{1 - \frac{X(p)}{p^\sigma}} \right)$$

The Random Model $\zeta(\sigma, X)$

Definition (Random Model)

Let X be random variable uniformly taking values in $\mathbb{T}^\infty$, indexed by the primes. Then, we define,

$$\zeta(\sigma, X) = \prod_p \left(\frac{1}{1 - \frac{X(p)}{p^\sigma}} \right)$$

It can be shown using probabilistic techniques (e.g., the Kolmogorov 3-series theorem or Chernoff-style concentration bounds) that the above product converges almost surely for $\sigma > 1/2$.

The Random Model $\zeta(\sigma, X)$

Definition (Random Model)

Let X be random variable uniformly taking values in $\mathbb{T}^\infty$, indexed by the primes. Then, we define,

$$\zeta(\sigma, X) = \prod_p \left(\frac{1}{1 - \frac{X(p)}{p^\sigma}} \right)$$

It can be shown using probabilistic techniques (e.g., the Kolmogorov 3-series theorem or Chernoff-style concentration bounds) that the above product converges almost surely for $\sigma > 1/2$.

Furthermore, for $\sigma > 1/2$, $\zeta(\sigma, X)$ is a $\mathbb{C}$-valued random variable with a continuous distribution, and Bohr-Jessen's result [BJ30] is essentially that $\{\log \zeta(\sigma + it)\}_{t \in [T, 2T]} \to \log \zeta(\sigma, X)$ as $T \to \infty$.

The Discrepancy Between $\log \zeta(\sigma + it)$ and $\log \zeta(\sigma, X)$

The limit $\{\log \zeta(\sigma + it)\}_{t \in [T, 2T]} \to \log \zeta(\sigma, X)$ as $T \to \infty$ naturally leads one to the question of how large the discrepancy between the distributions of true $\log \zeta$ and the the random model get for a fixed T.

The Discrepancy Between $\log\zeta(\sigma+it)$ and $\log\zeta(\sigma,X)$

The limit $\{\log\zeta(\sigma+it)\}_{t\in[T,2T]} \to \log\zeta(\sigma,X)$ as $T\to\infty$ naturally leads one to the question of how large the discrepancy between the distributions of true $\log\zeta$ and the the random model get for a fixed T.

Definition (Discrepancy)

Let $\sigma > 1/2$ be fixed, and T be large. Then,

$$\mathcal{D}_\sigma(T) = \sup_{\mathcal{R}} \left|\mathbb{P}_T\left(\log\zeta(\sigma+it)\in\mathcal{R}\right) - \mathbb{P}\left(\log\zeta(\sigma,X)\in\mathcal{R}\right)\right|$$

where the supremum runs over all axis-parallel rectangles $\mathcal{R}\subseteq\mathbb{C}$.

Clearly, the Bohr-Jessen result is $D_\sigma(T) = o(1)$.

Lamzouri, Lester and Radziwiłł prove the following bound:

Theorem

Let $1/2 < \sigma < 1$ *be fixed. Then*

$$D_\sigma(T) \ll_\sigma \frac{1}{(\log T)^\sigma}.$$

This improves on an earlier bound by Harman and Matsumoto.

The Characteristic Function of a Random Variable

For a real random variable ξ, the characteristic function $\Phi_\xi(x)$ is given by

$$\Phi_\xi(x) = \mathbb{E}\left(e^{ix\xi}\right).$$

The Characteristic Function of a Random Variable

For a real random variable ξ, the characteristic function $\Phi_\xi(x)$ is given by

$$\Phi_\xi(x) = \mathbb{E}\left(e^{ix\xi}\right).$$

The characteristic function is similar to the moment generating function $M_\xi(x) = \mathbb{E}(e^{x\xi})$ but with the advantage that Φ_ξ always exists for $x \in \mathbb{R}$, even though the moment generating function need not.

The Characteristic Function of a Random Variable

For a real random variable ξ, the characteristic function $\Phi_\xi(x)$ is given by

$$\Phi_\xi(x) = \mathbb{E}\left(e^{ix\xi}\right).$$

The characteristic function is similar to the moment generating function $M_\xi(x) = \mathbb{E}(e^{x\xi})$ but with the advantage that Φ_ξ always exists for $x \in \mathbb{R}$, even though the moment generating function need not.

If $F(u) = \mathbb{P}(\xi \leq u)$ is the distibution function of ξ on $\mathbb{R}$, then clearly,

$$\Phi_\xi(x) = \int_{-\infty}^{\infty} e^{ixu} dF(u)$$

and so Φ_ξ is just the Fourier transform of the measure dF.

The Characteristic Function of a Random Variable

When working with complex random variables ξ, the domain is extended to $z \in \mathbb{C}$, and the definition is changed to

$$\Phi_\xi(z) = \mathbb{E}\left(e^{i \operatorname{Re} \overline{z}\xi}\right)$$

The Characteristic Function of a Random Variable

When working with complex random variables ξ, the domain is extended to $z \in \mathbb{C}$, and the definition is changed to

$$\Phi_\xi(z) = \mathbb{E}\left(e^{i \operatorname{Re} \overline{z}\xi}\right)$$

Thinking of $z = u + iv$, and of Φ_ξ as a function of two real variables, this is the same as saying

$$\Phi_\xi(u, v) = \mathbb{E}\left(e^{i(u \operatorname{Re} \xi + v \operatorname{Im} \xi)}\right)$$

This leads us to the following two definitions:

The Characteristic Function of $\log \zeta$

This leads us to the following two definitions:

Definition (Characteristic Function of $\log \zeta(\sigma, X)$)

Let $\sigma > 1/2$ be fixed. Then, we define

$$\Phi_\sigma^r(u, v) = \mathbb{E}\left(\exp\left(iu \operatorname{Re} \log \zeta(\sigma, X) + iv \operatorname{Im} \log \zeta(\sigma, X)\right)\right).$$

The Characteristic Function of $\log \zeta$

This leads us to the following two definitions:

Definition (Characteristic Function of $\log \zeta(\sigma, X)$)

Let $\sigma > 1/2$ be fixed. Then, we define

$$\Phi_\sigma^r(u, v) = \mathbb{E}\left(\exp\left(iu \operatorname{Re}\log \zeta(\sigma, X) + iv \operatorname{Im}\log \zeta(\sigma, X)\right)\right).$$

Definition (Characteristic Function of $\log \zeta(\sigma + it)$)

Let $\sigma > 1/2$ and T large be fixed. Then, we define

$$\Phi_{\sigma,T}(u, v) = \mathbb{E}_T\left(\exp\left(iu \operatorname{Re}\log \zeta(\sigma + it) + iv \operatorname{Im}\log \zeta(\sigma + it)\right)\right)$$

The Characteristic Function of $\log \zeta$

This leads us to the following two definitions:

Definition (Characteristic Function of $\log \zeta(\sigma, X)$)

Let $\sigma > 1/2$ be fixed. Then, we define

$$\Phi^r_\sigma(u, v) = \mathbb{E}\left(\exp\left(iu \operatorname{Re}\log \zeta(\sigma, X) + iv \operatorname{Im}\log \zeta(\sigma, X)\right)\right).$$

Definition (Characteristic Function of $\log \zeta(\sigma + it)$)

Let $\sigma > 1/2$ and T large be fixed. Then, we define

$$\begin{aligned}\Phi_{\sigma,T}(u, v) &= \mathbb{E}_T\left(\exp\left(iu \operatorname{Re}\log \zeta(\sigma + it) + iv \operatorname{Im}\log \zeta(\sigma + it)\right)\right)\\ &= \frac{1}{T}\int_T^{2T} \exp\left(iu \operatorname{Re}\log \zeta(\sigma + it) + iv \operatorname{Im}\log \zeta(\sigma + it)\right)\, dt.\end{aligned}$$

Motivation: Lévy's Convergence Theorem

The motivation for considering the characteristic function of $\log \zeta$ comes from the following theorem from probability:

Theorem (Lévy's Convergence Theorem)

Let X_n be a sequence of $\mathbb{R}^n$-valued random variables, and X be an $\mathbb{R}^n$-valued random variable, with corresponding characteristic functions Φ_n and Φ. Then,

$$X_n \to X \text{ in distribution} \iff \Phi_n \to \Phi \text{ pointwise.}$$

Motivation: Lévy's Convergence Theorem

The motivation for considering the characteristic function of $\log \zeta$ comes from the following theorem from probability:

Theorem (Lévy's Convergence Theorem)

Let X_n be a sequence of $\mathbb{R}^n$-valued random variables, and X be an $\mathbb{R}^n$-valued random variable, with corresponding characteristic functions Φ_n and Φ. Then,

$$X_n \to X \text{ in distribution } \iff \Phi_n \to \Phi \text{ pointwise.}$$

Hence, to find the distributional discrepancy in $\log \zeta$, one looks for pointwise estimates for the characteristic functions.

Approximating $\Phi_{\sigma,T}$ by Φ_σ^r

We have the following theorem that tells us that these characteristic functions are not too far apart:

Theorem

Let $1/2 < \sigma < 1$ *and* $A \geq 1$ *be fixed. There exists a constant* $b = b(\sigma, A)$ *such that for all* $|u|, |v| \leq b(\log T)^\sigma$*, we have*

$$\Phi_{\sigma,T}(u,v) = \Phi_\sigma^r(u,v) + \mathcal{O}\left(\frac{1}{(\log T)^A}\right).$$

Approximating $\Phi_{\sigma,T}$ by Φ^r_σ: High Level Proof Idea

Proof Idea.

Let $Y \geq 0$ be a real number, and define the Dirichlet polynomial $R_Y(\sigma + it)$ by

$$R_Y(\sigma + it) = \sum_{n \leq Y} \frac{\Lambda(n)}{n^{\sigma+it} \log n} = \sum_{p^k \leq Y} \frac{1}{kp^{k(\sigma+it)}}$$

Approximating $\Phi_{\sigma,T}$ by Φ_σ^r: High Level Proof Idea

Proof Idea.

Let $Y \geq 0$ be a real number, and define the Dirichlet polynomial $R_Y(\sigma + it)$ by

$$R_Y(\sigma + it) = \sum_{n \leq Y} \frac{\Lambda(n)}{n^{\sigma+it} \log n} = \sum_{p^k \leq Y} \frac{1}{kp^{k(\sigma+it)}}$$

Correspondingly, we define the random Dirichlet polynomial by

$$R_Y(\sigma, X) = \sum_{n \leq Y} \frac{\Lambda(n)X(n)}{n^\sigma \log n} = \sum_{p^k \leq Y} \frac{X(p)^k}{kp^{k\sigma}}$$

□

Approximating $\Phi_{\sigma,T}$ by Φ_σ^r: High Level Proof Idea

Proof Idea.

Then, one can show that for $u, v \ll_{\sigma,A} (\log T)^\sigma$,

$$\Phi_{\sigma,T}(u,v) = \mathbb{E}_T\left(\exp\left(iu \operatorname{Re}\log\zeta(\sigma+it) + iv \operatorname{Im}\log\zeta(\sigma+it)\right)\right)$$

Approximating $\Phi_{\sigma,T}$ by Φ^r_σ: High Level Proof Idea

Proof Idea.

Then, one can show that for $u, v \ll_{\sigma,A} (\log T)^\sigma$,

$$\begin{aligned}\Phi_{\sigma,T}(u,v) &= \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\log\zeta(\sigma+it) + iv\,\mathrm{Im}\log\zeta(\sigma+it)\right)\right)\\ &\approx \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\,R_Y(\sigma+it) + iv\,\mathrm{Im}\,R_Y(\sigma+it)\right)\right)\end{aligned}$$

Approximating $\Phi_{\sigma,T}$ by Φ_σ^r: High Level Proof Idea

Proof Idea.

Then, one can show that for $u, v \ll_{\sigma,A} (\log T)^\sigma$,

$$\begin{aligned}\Phi_{\sigma,T}(u,v) &= \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\log\zeta(\sigma+it) + iv\,\mathrm{Im}\log\zeta(\sigma+it)\right)\right)\\ &\approx \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\,R_Y(\sigma+it) + iv\,\mathrm{Im}\,R_Y(\sigma+it)\right)\right)\\ &\approx \mathbb{E}\left(\exp\left(iu\,\mathrm{Re}\,R_Y(\sigma,X) + iv\,\mathrm{Im}\,R_Y(\sigma,X)\right)\right)\end{aligned}$$

Approximating $\Phi_{\sigma,T}$ by Φ^r_σ: High Level Proof Idea

Proof Idea.

Then, one can show that for $u, v \ll_{\sigma,A} (\log T)^\sigma$,

$$\begin{aligned}
\Phi_{\sigma,T}(u,v) &= \mathbb{E}_T\left(\exp\left(iu \operatorname{Re}\log\zeta(\sigma+it) + iv \operatorname{Im}\log\zeta(\sigma+it)\right)\right) \\
&\approx \mathbb{E}_T\left(\exp\left(iu \operatorname{Re} R_Y(\sigma+it) + iv \operatorname{Im} R_Y(\sigma+it)\right)\right) \\
&\approx \mathbb{E}\left(\exp\left(iu \operatorname{Re} R_Y(\sigma,X) + iv \operatorname{Im} R_Y(\sigma,X)\right)\right) \\
&\approx \mathbb{E}\left(\exp\left(iu \operatorname{Re}\log\zeta(\sigma,X) + iv \operatorname{Im}\log\zeta(\sigma,X)\right)\right)
\end{aligned}$$

Approximating $\Phi_{\sigma,T}$ by Φ_σ^r: High Level Proof Idea

Proof Idea.

Then, one can show that for $u, v \ll_{\sigma,A} (\log T)^\sigma$,

$$\begin{aligned}
\Phi_{\sigma,T}(u,v) &= \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\log\zeta(\sigma+it) + iv\,\mathrm{Im}\log\zeta(\sigma+it)\right)\right) \\
&\approx \mathbb{E}_T\left(\exp\left(iu\,\mathrm{Re}\,R_Y(\sigma+it) + iv\,\mathrm{Im}\,R_Y(\sigma+it)\right)\right) \\
&\approx \mathbb{E}\left(\exp\left(iu\,\mathrm{Re}\,R_Y(\sigma,X) + iv\,\mathrm{Im}\,R_Y(\sigma,X)\right)\right) \\
&\approx \mathbb{E}\left(\exp\left(iu\,\mathrm{Re}\log\zeta(\sigma,X) + iv\,\mathrm{Im}\log\zeta(\sigma,X)\right)\right) \\
&= \Phi_\sigma^r(u,v).
\end{aligned}$$

Here $\approx$ means up to an acceptable error. □

Beurling-Selberg Functions

Beurling-Selberg functions are a major tool in analytic number theory, which essentially arise as smoothed approximations to cutoff functions with prescribed Fourier support.

Beurling-Selberg Functions

Beurling-Selberg functions are a major tool in analytic number theory, which essentially arise as smoothed approximations to cutoff functions with prescribed Fourier support.

Lemma (Beurling-Selberg functions for a rectangle)

Let $\mathcal{R}$ be an axis-parallel rectangle in $\mathbb{C}$, and $L > 0$ be a real number. For any $z \in \mathbb{C}$ we have

$$1_{\mathcal{R}}(z) = W_{L,\mathcal{R}}(z) + E_{L,\mathcal{R}}(z)$$

where $W_{L,\mathcal{R}}(z)$ is smooth, and

$$E_{L,\mathcal{R}}(z) \ll (\operatorname{sinc} \pi L \theta_z)^2$$

with $\operatorname{sinc} x = \frac{\sin x}{x}$, and θ_z is the biggest among the distances of z from the sides of $\mathcal{R}$.

Extracting the Discrepancy Bound

Proof Idea.

Uniformly for axis-parallel rectangles $\mathcal{R}$, we want to bound

$$D_\sigma^{\mathcal{R}}(T) = \Big|\mathbb{P}_T(\log\zeta(\sigma+it)\in\mathcal{R}) - \mathbb{P}(\log\zeta(\sigma,X)\in\mathcal{R})\Big|.$$

Extracting the Discrepancy Bound

Proof Idea.

Uniformly for axis-parallel rectangles $\mathcal{R}$, we want to bound

$$D_\sigma^{\mathcal{R}}(T) = \left|\mathbb{P}_T(\log\zeta(\sigma+it)\in\mathcal{R}) - \mathbb{P}(\log\zeta(\sigma,X)\in\mathcal{R})\right|.$$

We can rewrite this as the magnitude of

$$\mathbb{E}_T\left[1_{\mathcal{R}}(\log\zeta(\sigma+it)\right] - \mathbb{E}\left[1_{\mathcal{R}}(\log\zeta(\sigma,X)\right]$$

Extracting the Discrepancy Bound

Proof Idea.

Uniformly for axis-parallel rectangles $\mathcal{R}$, we want to bound

$$D_\sigma^{\mathcal{R}}(T) = \Big|\mathbb{P}_T(\log\zeta(\sigma+it)\in\mathcal{R}) - \mathbb{P}(\log\zeta(\sigma,X)\in\mathcal{R})\Big|.$$

We can rewrite this as the magnitude of

$$\begin{aligned} &\mathbb{E}_T\Big[1_{\mathcal{R}}(\log\zeta(\sigma+it)\Big] - \mathbb{E}\Big[1_{\mathcal{R}}(\log\zeta(\sigma,X)\Big] \\ \approx\, &\mathbb{E}_T\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma+it)\Big] - \mathbb{E}\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma,X)\Big] \end{aligned}$$

where the error term $\mathbb{E}_T\Big[E_{L,\mathcal{R}}(\log\zeta(\sigma+it)\Big] - \mathbb{E}\Big[E_{L,\mathcal{R}}(\log\zeta(\sigma,X)\Big]$ can be shown to be $\ll 1/L$. □

Extracting the Discrepancy Bound

Proof Idea.

Explicitly, for $z = x + iy$, we have that $W_{L,\mathcal{R}}(z)$ is given by

$$\operatorname{Re} \int_0^L \int_0^L \frac{G\left(\frac{u}{L}\right) G\left(\frac{v}{L}\right) \left(e(ux - vy) f_{1,\mathcal{R}}(u,v) - e(ux + vy) f_{2,\mathcal{R}}(u,v)\right)}{2uv} \, du dv$$

where G is bounded on $[0, 1]$, and $f_{j,\mathcal{R}}(u, v) \ll \mu(\mathcal{R}) uv$.

Extracting the Discrepancy Bound

Proof Idea.

Explicitly, for $z = x + iy$, we have that $W_{L,\mathcal{R}}(z)$ is given by

$$\mathrm{Re} \int_0^L \int_0^L \frac{G\left(\frac{u}{L}\right) G\left(\frac{v}{L}\right) \left(e(ux - vy) f_{1,\mathcal{R}}(u,v) - e(ux + vy) f_{2,\mathcal{R}}(u,v)\right)}{2uv} dudv$$

where G is bounded on $[0,1]$, and $f_{j,\mathcal{R}}(u,v) \ll \mu(\mathcal{R})uv$. It follows that $\mathbb{E}\Big[W_{L,\mathcal{R}}(\log \zeta)\Big]$ is given by the real part of

$$\int_0^L \int_0^L \frac{G\left(\frac{u}{L}\right) G\left(\frac{v}{L}\right) \left(\Phi(2\pi u, -2\pi v) f_{1,\mathcal{R}}(u,v) - \Phi(2\pi u, 2\pi v) f_{2,\mathcal{R}}(u,v)\right)}{2uv} dudv$$

□

Extracting the Discrepancy Bound

Proof Idea.

From this, together with our estimate for characteristic functions we conclude that

$$\mathbb{E}_T\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma+it)\Big]-\mathbb{E}\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma,X)\Big]\ll_{\sigma,A}\frac{L^2\mu(\mathcal{R})}{(\log T)^A}$$

provided that $L\ll_{\sigma,A}(\log T)^{\sigma}$.

Extracting the Discrepancy Bound

Proof Idea.

From this, together with our estimate for characteristic functions we conclude that

$$\mathbb{E}_T\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma+it)\Big] - \mathbb{E}\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma,X)\Big] \ll_{\sigma,A} \frac{L^2\mu(\mathcal{R})}{(\log T)^A}$$

provided that $L \ll_{\sigma,A} (\log T)^\sigma$. In particular, choosing A large and $L \asymp_{\sigma,A} (\log T)^\sigma$, we see that

$$D_\sigma^{\mathcal{R}}(T) \ll_{\sigma,A} \frac{1}{(\log T)^\sigma} + \frac{\mu(\mathcal{R})}{(\log T)^{A-2\sigma}}.$$

Extracting the Discrepancy Bound

Proof Idea.

From this, together with our estimate for characteristic functions we conclude that

$$\mathbb{E}_T\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma+it)\Big]-\mathbb{E}\Big[W_{L,\mathcal{R}}(\log\zeta(\sigma,X)\Big]\ll_{\sigma,A}\frac{L^2\mu(\mathcal{R})}{(\log T)^A}$$

provided that $L \ll_{\sigma,A} (\log T)^\sigma$. In particular, choosing A large and $L \asymp_{\sigma,A} (\log T)^\sigma$, we see that

$$D_\sigma^{\mathcal{R}}(T) \ll_{\sigma,A} \frac{1}{(\log T)^\sigma} + \frac{\mu(\mathcal{R})}{(\log T)^{A-2\sigma}}.$$

To establish the theorem we need to remove the dependence on $\mathcal{R}$. This can be done by appealing to a large deviation estimate – morally this says that the extremal $\mathcal{R}$ maximising $D_\sigma^{\mathcal{R}}(T)$ satisfies $\mathcal{R} \subseteq [-\log_2 T, \log_2 T]^2$, completing the proof. □

The End