Modern Stochastics: Theory and Applications 8 (2) (2021) 275–289
https://doi.org/10.15559/21-VMSTA172

Chaos expansion of uniformly distributed random variables and application to number theory

Ciprian Tudor

Université de Lille, CNRS UMR1824, Laboratoire Paul Painlevé, 59655 Villeneuve d'Ascq Cedex, France

ciprian.tudor@univ-lille.fr (F. Tudor)

Received: 29 January 2021, Revised: 13 February 2021, Accepted: 20 February 2021, Published online: 26 March 2021

Abstract The chaos expansion of a random variable with uniform distribution is given. This decomposition is applied to analyze the behavior of each chaos component of the random variable $\log \zeta$ on the so-called critical line, where ζ is the Riemann zeta function. This analysis gives a better understanding of a famous theorem by Selberg.

Keywords Riemann zeta function, Malliavin calculus, multiple Wiener–Itô integrals, Selberg theorem

2010 MSC 60F05, 60H05, 1M06

1 Introduction

Let H be a real and separable Hilbert space and let $(W(h), h \in H)$ be an isonormal Gaussian process on a standard probability space $(\Omega, \mathcal{F}, P)$. It is well-known that any square integrable random variable, measurable with respect to the sigma-algebra generated by W can be decomposed in chaos, i.e. it can be written as an orthogonal sum $F = \sum_{n\geq 0} F_n$ where for every $n \geq 0$, the random variable F_n is an element of the nth Wiener chaos. The knowledge of the concrete chaos expansion of F (i.e. the knowledge the exact form of F_n for every $n \geq 0$) gives an important information on the random variable F. If F is a random variable with a given common probability distribution, it is in general hard to get its exact chaos expansion, except in some particular case (Gaussian distribution, Gamma distribution, etc). In this work, our purpose is to find the chaos expansion of a random variable with uniform distribution and to apply this result to a well-known problem in number theory. We will consider the random variable F given by

$$F = e^{-\frac{1}{2}\left(W(f)^2+W(g)^2\right)} \tag{1}$$

with f, g orthonormal elements of the Hilbert space H. Then F follows the uniform distribution over the unit interval $[0, 1]$. By using the tools of Malliavin calculus, Stroock formula and the properties of the Hermite polynomials, we derive the chaos decomposition of the random variable defined by (1). Then we will apply our result to the Selberg theorem, which concerns the behavior of the Riemann zeta fumction on the critical line. Let us recall the context. The Riemann zeta function ζ is defined, for $\Re s > 1$, by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{2}$$

and when $\Re s \leq 1$, the function ζ is defined as an analytic continuation of (2). The distribution of the zeta zeros is one of the outstanding problems in mathematics. It is known that $\zeta(-2n) = 0$ for every $n \geq 1$. The points $-2n$ are called the *trivial zeros* of the Riemann zeta function. The most mysterious facts about the Riemann zeta function concerns the distribution of its *nontrivial zeros*. The Riemann hypothesis claims that all the nontrivial zeros of the Riemann zeta function lie on *the critical line* $\Re s = \frac{1}{2}$. Therefore, the behavior of the function ζ on the critical line and close to this critical line has been intensively studied.

A famous result by Atle Selberg says that, if $T > 0$ and t is a random variable uniformly distributed over the interval $[T, 2T]$, then the sequence

$$\frac{\log \zeta\left(\frac{1}{2} + \mathbf{i}t\right)}{\sqrt{\frac{1}{2}\log\log T}} \to X_1 + iX_2 \text{ in distribution as } T \to \infty \tag{3}$$

where $X_1 + iX_2$ is a complex-valued standard normal random variable, i.e. $X_1, X_2 \sim N(0, 1)$ are independent random variables. There are several versions of this theorem. In particular, the result (3) holds if $t \sim \mathcal{U}[0, T]$ or, more generally, if $t \sim \mathcal{U}[aT, bT]$ with $b > a \geq 0$. We will work with $t \sim \mathcal{U}[0, T]$ and we will assume throughout in the sequel $t = TU$ with $U \sim \mathcal{U}[0, 1]$. As in the literature (e.g., [7, 16]), we will still call Selberg's theorem the result concerning the convergence of (3) with $t = TU$.

Recall that for any $z \in \mathbb{C}$, $\log z = \log|z| + i \arg z$. Then the convergence (3) is actually equivalent to (4) and (5) below (see [11–13])

$$\frac{\log\left|\zeta\left(\frac{1}{2} + \mathbf{i}t\right)\right|}{\sqrt{\frac{1}{2}\log\log T}} \to^{(d)}_{T\to\infty} X_1 \sim N(0, 1) \tag{4}$$

and

$$\frac{\arg \zeta\left(\frac{1}{2} + \mathbf{i}t\right)}{\sqrt{\frac{1}{2}\log\log T}} \to^{(d)}_{T\to\infty} X_2 \sim N(0, 1) \tag{5}$$

where " $\to^{(d)}$ " stands for the convergence in distribution.

The idea behind the proof of the limit theorems (4) and (5) is (see, among others, [11–15]) to approximate $\log \zeta\left(\frac{1}{2}+\mathbf{i}t\right)$ by the (renormalized) Dirichlet series

$$\begin{aligned}
&\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\le T^{\varepsilon}}\frac{1}{p^{\frac{1}{2}+\mathbf{i}t}}\\
&\quad=\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\left(\sum_{p\le T^{\varepsilon}}\frac{\cos(t\log p)}{\sqrt{p}}+\mathbf{i}\sum_{p\le T^{\varepsilon}}\frac{\sin(t\log p)}{\sqrt{p}}\right).
\end{aligned}\tag{6}$$

with $t=TU$, $U\sim\mathcal{U}[0,1]$ and ε small enough. We will work throughout with $\varepsilon=1$. This approximation of $\log\zeta\left(\frac{1}{2}+\mathbf{i}t\right)$ by the Dirichlet series (6) is in L^2 sense, since (see [13], see also [7] for a detailed proof)

$$\mathbf{E}\left|\log\zeta\left(\frac{1}{2}+\mathbf{i}t\right)-\sum_{p\le T}\frac{1}{p^{\frac{1}{2}+\mathbf{i}t}}\right|^2\le C$$

where C is a strictly positive constant not depending on T. This implies that the sequence

$$\left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\left(\log\zeta\left(\frac{1}{2}+\mathbf{i}t\right)-\sum_{p\le T}\frac{1}{p^{\frac{1}{2}+\mathbf{i}t}}\right)\right)_{T>0}$$

converges to zero in $L^2(\Omega)$ as $T\to\infty$.

Our purpose is to bring a new contribution to understanding the limit theorems (4) and (5). More exactly, we will analyze the asymptotic behavior of each chaos component of $\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\log|\zeta\left(\frac{1}{2}+it\right)|$ with $t\sim\mathcal{U}[0,T]$. That is, if $J_n(T)$ denotes the projection of $\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\log\left|\zeta\left(\frac{1}{2}+it\right)\right|$ on the nth Wiener chaos, we want to study the limit behavior in distribution of $J_n(T)$ as $T\to\infty$ for each fixed $n\ge 0$. This will allow to understand which chaos projection $J_n(T)$ is dominant with respect to the others and determines the limit behavior of the $\log\zeta$.

Many other related works in the old and recent literature treated the distribution of the zeros of the Riemann zeta function. Some related results, among many others, are [1–4, 6].

Our analysis will be based on the study of the Dirichlet approximation (6). Using the chaos expansion of the uniformly distributed random variable U, we will find the chaos expansion of the random variable $\cos(TU\log p)$ with $U\sim\mathcal{U}[0,1]$ via Malliavin calculus and we will study the limit in distribution of each chaos component. We will see that every chaos converges to zero, but their sum tends in distribution to the Gaussian law. All the chaoses contribute to the limit and there is no term that is bigger than the others and gives the limit behavior.

Our work has the following structure. Section 2 contains some preliminaries on Wiener chaos and Malliavin calculus needed throughout the work. Section 3 is devoted to the chaos expansion of a uniformly distributed random variable U. In Section 4, we obtain the chaos decomposition of the Dirichlet series that approximates

$\log \zeta\left(\frac{1}{2}+\mathbf{i}TU\right)$ and we then analyze the asymptotic behavior, as $T \to \infty$, of each chaos component.

2 Preliminaries: the multiple stochastic integral and the Malliavin derivative

We also present the elements from the Malliavin calculus that will be used in the paper. We refer to [9] for a more complete exposition. Consider $\mathcal{H}$ as a real separable Hilbert space and $(B(\varphi), \varphi \in \mathcal{H})$ an isonormal Gaussian process on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, that is, a centered Gaussian family of random variables such that $\mathbf{E}\left(B(\varphi)B(\psi)\right) = \langle \varphi, \psi\rangle_{\mathcal{H}}$.

We denote by D the Malliavin derivative operator that acts on smooth functions of the form $F = g(B(\varphi_1), \ldots, B(\varphi_n))$ (g is a smooth function with compact support and $\varphi_i \in \mathcal{H}, i = 1, \ldots, n$)

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(B(\varphi_1), \ldots, B(\varphi_n))\varphi_i.$$

It can be checked that the operator D is closable from $\mathcal{S}$ (the space of smooth functionals as above) into $L^2(\Omega; \mathcal{H})$ and it can be extended to the space $\mathbb{D}^{1,p}$ which is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p}^{p} = \mathbf{E}F^p + \mathbf{E}\|DF\|_{\mathcal{H}}^{p}.$$

We denote by $\mathbb{D}^{k,\infty} := \cap_{p\geq}\mathbb{D}^{k,p}$ for every $k \geq 1$. In this paper, $\mathcal{H}$ will be the standard Hilbert space $L^2([0, T])$.

We will make use of the chain rule for the Malliavin derivative (see Proposition 1.2.4 in [9]). That is, if $\varphi : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function and $F \in \mathbb{D}^{1,2}$, then $\varphi(F) \in \mathbb{D}^{1,2}$ and

$$D\varphi(F) = \varphi'(F)DF. \tag{7}$$

Denote by I_n the multiple stochastic integral with respect to B (see [9]). This I_n is actually an isometry between the Hilbert space $H^{\odot n}$ (symmetric tensor product) equipped with the scaled norm $\sqrt{n!}\|\cdot\|_{H^{\otimes n}}$ and the Wiener chaos of order n which is defined as the closed linear span of the random variables $H_n(B(\varphi))$ where $\varphi \in \mathcal{H}$, $\|\varphi\|_H = 1$ and H_n is the Hermite polynomial of degree $n \geq 1$

$$H_n(x) = (-1)^n \exp\left(\frac{x^2}{2}\right)\frac{d^n}{dx^n}\left(\exp\left(-\frac{x^2}{2}\right)\right), \qquad x \in \mathbb{R}. \tag{8}$$

The isometry of multiple integrals can be written as: for m, n positive integers,

$$\begin{aligned} \mathbf{E}\left(I_n(f)I_m(g)\right) &= n!\langle \tilde{f}, \tilde{g}\rangle_{H^{\otimes n}} \quad \text{if } m = n,\\ \mathbf{E}\left(I_n(f)I_m(g)\right) &= 0 \quad \text{if } m \neq n. \end{aligned} \tag{9}$$

The Malliavin derivative D acts on the Wiener chaos as an annilihilation operator: if $F = I_n(f)$ with symmetric $f \in L^2([0, T]^n)$, then $D_tF = nI_{n-1}(f(\cdot, t))$ where "$\cdot$" stands for $n-1$ variables in $[0, T]$.

3 Chaos expansion of uniformly distributed random variables

The uniformly distributed random variable U will be chosen of a particular form that allows to use the techniques of the Malliavin calculus. Actually, in the sequel we will assume

$$U = e^{-\frac{1}{2}(W(f)^2+W(g)^2)} \tag{10}$$

with the following conditions fixed throughout our work: $f, g \in H$, $\|f\| = \|g\| = 1$ and $\langle f, g\rangle = 0$ (all the scalar products and norms in the paper will be considered in H if no further specification is made). In (10), $(W(h), h \in H = L^2([0,1]))$ stands for a Gaussian isonormal process as described in Section 2. In particular, this implies that $W(f)$ and $W(g)$ are independent standard normal random variables. The fact that the random variable (10) is uniformly distributed over $[0,1]$ follows from the simple computations: with $F : \mathbb{R} \to \mathbb{R}$ an arbitrary function such that $\mathbf{E}F(U) < \infty$,

$$\begin{aligned} \mathbf{E}F(U) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} F\left(e^{-\frac{x^2+y^2}{2}}\right) e^{-\frac{x^2+y^2}{2}}dxdy \\ &= \int_0^\infty d\rho F\left(e^{-\frac{\rho^2}{2}}\right)e^{-\frac{\rho^2}{2}}\rho = \int_0^1 F(u)du \end{aligned}$$

where we used the change of variables with polar coordinates.

We analyze the chaos expansion of

$$U^{2k} = e^{-k(W(f)^2+W(g)^2)} \tag{11}$$

for every $k > 0$. This will be done by using the techniques of the Malliavin calculus. We will also need some properties of the Hermite polynomials. Let H_n denote the nth Hermite polynomial, see (8). Recall that, if $Y \sim N(0, \sigma^2)$, then

$$\mathbf{E}H_{2m}(Y) = \frac{(\sigma^2-1)^m(2m)!}{2^m m!} \tag{12}$$

and $\mathbf{E}H_n(Y) = 0$ if n is odd.

We will use the following two auxiliary lemmas that concern the Hermite polynomials.

Lemma 1. *For every $s \geq 0$ and for every $k > 0$,*

$$\mathbf{E}\left[e^{-kW(f)^2}H_{2s}(\sqrt{2k}W(f))\right] = \frac{(-1)^s(2s)!}{2^s s!}(2k+1)^{-s-\frac{1}{2}}. \tag{13}$$

Proof. With the notation $\sigma^2 = \frac{2k}{2k+1}$, we have

$$\begin{aligned} \mathbf{E}\left[e^{-kW(f)^2}H_{2s}(\sqrt{2k}W(f))\right] &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{x^2(2k+1)}{2}}H_{2s}(\sqrt{2k}x)dx \\ &= \frac{1}{\sqrt{2k+1}}\frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}}H_{2s}(x)dx \\ &= \frac{1}{\sqrt{2k+1}}\mathbf{E}H_{2s}(Y) \end{aligned}$$

where $Y \sim N(0, \sigma^2)$. The conclusion comes from (12). $\square$

Recall that $D^{(s)}(s \geq 1)$ denotes the sth iterated Malliavin derivative. If $s = 0$ then by convention $f^{\otimes s} = 1$.

Lemma 2. *For any $k > 0$, consider the random variable $F = e^{-kW(f)^2}$. Then for every $s \geq 0$,*

$$D^{(s)}F = (2k)^{\frac{s}{2}}(-1)^s e^{-kW(f)^2} H_s\left(\sqrt{2k}W(f)\right) f^{\otimes s}. \tag{14}$$

Moreover, we have

$$\mathbf{E}D^{(s)}F = 0 \text{ if } s \text{ is odd} \tag{15}$$

and

$$\mathbf{E}D^{(2s)}F = \frac{(-1)^s(2s)!k^s}{s!}(2k+1)^{-s-\frac{1}{2}} f^{\otimes 2s}, \tag{16}$$

for every $s \geq 0$.

Proof. Denote by $g(x) = e^{-x^2}$ for $x \in \mathbb{R}$. By the chain rule of the Malliavin derivative (7), we have, for every $s \geq 0$,

$$D^{(s)}F = g^{(s)}\left(\sqrt{k}W(f)\right) k^{\frac{s}{2}} f^{\otimes s}. \tag{17}$$

We will use the following relation between the derivatives of the function g and the Hermite polynomials:

$$g^{(s)}(x) = (-1)^s e^{-x^2} 2^{\frac{s}{2}} H_s(\sqrt{2}x), \text{ for every } s \geq 0, x \in \mathbb{R}. \tag{18}$$

It is easy to see that if s is odd, then the of (14) expectation vanishes. If s is even, by plugging (18) into (17), we obtain (14). Consequently,

$$\mathbf{E}D^{(s)}F = (2k)^{\frac{s}{2}} s! f^{\otimes s}\mathbf{E}\left(e^{-kW(f)^2} H_s\left(\sqrt{2k}W(f)\right)\right)$$

and by using (13), we obtain (15) and (16). □

The next step is to get the chaos expansion of U^{2k} with U a random variable uniformly distributed over [0, 1]. If f, g are two functions, by $f\tilde{\otimes}g$ we denote the symmetrization of their tensor product.

Lemma 3. *Let U be given by (10). Then for every $k > 0$ the random variable U^{2k} admits the following Wiener chaos expansion:*

$$U^{2k} = \sum_{n \geq 0} \frac{k^n}{(2k+1)^{n+1}} I_{2n}(h_{2n}), \tag{19}$$

with

$$h_{2n} = (-1)^n \sum_{s=0}^{n} \frac{1}{s!(n-s)!} f^{\otimes 2s}\tilde{\otimes} g^{\otimes(2n-2s)}. \tag{20}$$

Proof. We will first show that for every $k > 0$

$$\mathbf{E}D^{(n)}U^{2k} = 0 \text{ if } n \text{ is odd} \tag{21}$$

and for every $n \geq 1$

$$\mathbf{E}D^{(2n)}U^{2k} = \frac{k^n}{(2k+1)^{n+1}}(2n)!h_{2n} \tag{22}$$

with h_n given by (20). By using the Leibniz rule for the Malliavin derivarive (see, e.g., [9], Exercise 1.2. 13), we can write, for $n, k \geq 1$,

$$\begin{aligned} D^{(n)}U^{2k} &= D^{(n)}(e^{-kW(f)^2}e^{-kW(g)^2}) \\ &= \sum_{s=0}^{n} C_n^s \left(D^{(s)}e^{-kW(f)^2}\right) \tilde{\otimes} \left(D^{(n-s)}e^{-kW(g)^2}\right), \end{aligned}$$

so

$$\mathbf{E}D^{(n)}U^{2k} = \sum_{s=0}^{n} C_n^s \left(\mathbf{E}D^{(s)}e^{-kW(f)^2}\right) \tilde{\otimes} \left(\mathbf{E}D^{(n-s)}e^{-kW(g)^2}\right) \tag{23}$$

where we used the independence of $W(f)$ and $W(g)$. By Lemma 2, we immediately get that $\mathbf{E}D^{(n)}U^{2k} = 0$ if n is odd and

$$\begin{aligned} \mathbf{E}D^{(2n)}U^{2k} &= \frac{(-1)^n k^n}{(2k+1)^{n+1}} \sum_{s=0}^{n} C_{2n}^{2s} \frac{(2s)!(2n-2s)!}{s!(n-s)!} f^{\otimes 2s} \tilde{\otimes} g^{\otimes(2n-2s)} \\ &= \frac{(-1)^n k^n}{(2k+1)^{n+1}} \sum_{s=0}^{n} \frac{(2n)!}{s!(n-s)!} f^{\otimes 2s} \tilde{\otimes} g^{\otimes(2n-2s)} \\ &= \frac{k^n}{(2k+1)^{n+1}}(2n)!h_{2n} \end{aligned}$$

with h_{2n} given by (20).

In order to obtain the Wiener chaos decomposition of U^{2k}, we will use the Stroock formula (see, e.g., [9]) to write

$$U^{2k} = \sum_{n\geq 0} \frac{1}{n!} I_n\left(\mathbf{E}D^{(n)}U^{2k}\right), \tag{24}$$

and using the formulas (21) and (22) for $\mathbf{E}D^{(n)}U^k$ with $n \geq 1$ and $k \geq 1$, we get (19). □

Remark 1. It is worth to point out that the (trivial) formulas $\mathbf{E}U^{2k} = \frac{1}{2k+1}$ and $\mathbf{E}U^{4k} = \frac{1}{4k+1}$ can also be checked through the chaos expansion (26). Indeed, for $n = 0$ in the right-hand side of (26) we get $\mathbf{E}U^{2k} = \frac{1}{2k+1}$ and, since $\|h_{2n}\|^2 = \sum_{s=0}^{n} \frac{(2s)!(2n-2s)!}{(2n)!} \frac{1}{(s!(n-s)!)^2}$ we get, via the isometry (9)

$$\mathbf{E}U^{4k} = \mathbf{E}(U^{2k})^2 = \frac{1}{(2k+1)^2} \sum_{n=0}^{\infty} \frac{k^{2n}}{(2k+1)^{2n}}(2n)!\|h_{2n}\|^2$$

$$
\begin{aligned}
&= \frac{1}{(2k+1)^2}\sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}\sum_{s=0}^{n}\frac{(2s)!(2n-2s)!}{(s!(n-s)!)^2}\\
&= \frac{1}{(2k+1)^2}\sum_{s=0}^{\infty}\frac{(2s)!}{s!^2}\sum_{n=s}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}\frac{(2n-2s)!}{(n-s)!)^2}\\
&= \frac{1}{(2k+1)^2}\left(\sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}\frac{(2n)!}{n!^2}\right)^2 = \frac{1}{4k+1},
\end{aligned}
$$

since, with $Z \sim N(0,1)$,

$$
\begin{aligned}
&\sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}\frac{(2n)!}{n!^2} = \sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}2^n\mathbf{E}Z^{2n}\\
&= \mathbf{E}\sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}\frac{(2n)!}{n!^2} = \sum_{n=0}^{\infty}\frac{k^{2n}}{(2k+1)^{2n}}2^nZ^{2n} = \mathbf{E}e^{\frac{2k^2}{(2k+1)^2}Z^2} = \frac{2k+1}{\sqrt{1+4k}}.
\end{aligned}
$$

4 Chaos analysis in Selberg' s theorem

We use the results in the previous section in order to get the chaos expansion of the Dirichlet series (6) and to study the behavior, as $T \to \infty$, of each chaos component of $\log\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)$ with U given by (10).

4.1 Chaos expansion of the Dirichlet series

We first obtain the Wiener–Itô chaos expansion of the Dirichlet series (6) that approximates $\log\zeta$ on the critical line (in the sense of (35)). We will actually focus on the analysis of the real part of (6) but we stress that a similar study can be done for the imaginary part. Consider the family $(X_T)_{T>0}$ given by

$$
X_T = \sum_{p\le T}\left[\frac{\cos(TU\log p)}{\sqrt{p}} - \mathbf{E}\frac{\cos(TU\log p)}{\sqrt{p}}\right], \tag{25}
$$

where the sum is taken over the primes p and U is $\mathcal{U}[0,1]$ distributed, of the form (10).

Proposition 1. *For $T>0$, let X_T be given by (25). Denote*

$$
\mathcal{X}_T := \frac{1}{\sqrt{\frac{1}{2}\log\log T}}X_T.
$$

Then

$$
\mathcal{X}_T = \sum_{n\ge 1}c_{2n}(T)I_{2n}(h_{2n}) \tag{26}
$$

where for every $n\ge 1$, h_n are defined by (20) and

$$
c_{2n}(T) = \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\le T}\frac{1}{\sqrt{p}}\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p)^{2k}\left(\frac{k}{2k+1}\right)^n. \tag{27}
$$

Proof. From (19) and the series expansion of the cosinus function $\cos x = \sum_{k\geq 0} \frac{(-1)^k}{(2k)!} x^{2k}$, we obtain the chaos decomposition of the random variable $\cos(T \log pU)$ where $T > 0$, p is a prime number and U is defined in (10). We actually have

$$\cos(TU \log p) = \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!} (T \log p)^{2k} \sum_{n\geq 0} I_{2n}(h_{2n}) \left(\frac{k}{2k+1}\right)^n .$$

Thus $\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \cos(T \log pU) = \sum_{n\geq 0} J_{2n}(T)$ with

$$J_{2n}(T) = I_{2n}\left(h_h\right) \sum_{p\leq T} \frac{1}{\sqrt{p}} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!} (T \log p)^{2k} \left(\frac{k}{2k+1}\right)^n \tag{28}$$

and h_n given by (20). Note that the chaos of order zero can be explicitly computed. In fact,

$$\begin{aligned} J_0(T) &= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!} (T \log p)^{2k} \\ &= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \frac{1}{T \log p} \sin(T \log p) = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \mathbf{E} X_T \end{aligned}$$

and therefore

$$\mathcal{X}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} X_T = \sum_{n=1}^{\infty} J_{2n}(T) = \sum_{n=1}^{\infty} c_{2n}(T) I_{2n}(h_{2n}).$$

□

Note that only Wiener chaoses of even order appear in the decomposition of X_T.

4.2 Asymptotic behavior for each chaos

We now analyze the limit behavior of the projection of $\Re \log \zeta\left(\frac{1}{2} + \mathbf{i}TU\right)$ on each Wiener chaos. The first step is to do this study for the Dirichlet series (6).

4.2.1 Chaoses of the Dirichlet series

From Proposition 1, we notice that the (renormalized) Dirichlet series (25) can be expanded into an infinite sum of random variables in chaoses of even orders. The projection on the $2n$th Wiener chaos is given by

$$J_{2n}(T) = c_{2n}(T) I_{2n}(h_{2n}) \tag{29}$$

with $c_{2n}(T)$, h_{2n} from (27), (20) respectively. We will show that for every $n \geq 1$,

$$J_{2n}(T) \rightarrow_{T\to\infty} 0 \text{ almost surely and in } L^2(\Omega),$$

i.e. the projection of each Wiener chaos of $\mathcal{X}_T$ converges to zero when $T \to \infty$. Note that in (29) only the coefficients $c_{2n}(T)$ depend on T and but the random parts $I_{2n}(h_{2n})$ do not. Therefore it suffices to study the behavior of $c_{2n}(T)$ as $T \to \infty$. This is done in the lemma below.

Lemma 4. *Let $T > 0$, $n \geq 1$ and let $c_{2n}(T)$ be defined by (27). Then for every $n \geq 1$,*

$$c_{2n}(T) \to_{T\to\infty} 0.$$

Proof. Let us first give the proof for $n = 1$. This will illustrate what happens in the general case. For $n = 1$, we have, writting $\frac{k}{2k+1} = \frac{1}{2}\left(1 - \frac{1}{2k+1}\right)$,

$$\begin{aligned} c_2(T) &= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k}\left(\frac{k}{2k+1}\right) \\ &= \frac{1}{2}\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k} \\ &\quad -\frac{1}{2}\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k}\left(\frac{1}{2k+1}\right). \end{aligned}$$

The two sums above can be calculated. Since

$$\sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k} = \frac{1}{T\log p}\sin(T\log p)$$

and

$$\begin{aligned} &\sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k}\left(\frac{1}{2k+1}\right) \\ &= \sum_{k\geq 0} \frac{(-1)^k}{(2k+1)!}(T\log p)^{2k}\frac{1}{(T\log p)^{2k+1}}\int_0^{T\log p} y^{2k}dy \\ &= \frac{1}{T\log p}\int_0^{T\log p}\frac{\sin y}{y}dy, \end{aligned}$$

we will have

$$c_2(T) = \frac{1}{2}\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T}\frac{1}{\sqrt{p}}\frac{1}{T\log p}\left(\sin(T\log p) - \int_0^{T\log p}\frac{\sin y}{y}dy\right).$$

By (34) and the trivial inequality

$$\left|\int_0^t \frac{\sin y}{y}dy - \frac{\pi}{2}\right| \leq \frac{C}{t}$$

we clearly get $c_2(T) \le CT^{-\frac{1}{2}} \to_{T\to\infty} 0$. Concerning the general case, we use again the identity $\frac{k}{2k+1} = \frac{1}{2}\left(1 - \frac{1}{2k+1}\right)$ and the Newton's formula to get

$$
\begin{aligned}
& c_{2n}(T) \qquad (30)\\
&= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k} \left(\frac{k}{2k+1}\right)^n\\
&= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k} \left(\frac{1}{2} - \frac{1}{2(2k+1)}\right)^n\\
&= \frac{(-1)^n}{2^n} \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k} \sum_{a=0}^{n} C_n^a (2k+1)^{-a} (-1)^{n-a}\\
&:= \frac{(-1)^n}{2^n} \sum_{a=0}^{n} C_n^a (-1)^{n-a} A_T(a) \qquad (31)
\end{aligned}
$$

where, for every $a = 0, 1, .., n$, we used the notation

$$
A_T(a) := \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k} (2k+1)^{-a}.
$$

The only element depending on T in the decomposition of $c_{2n}(T)$ is the one denoted by $A_T(a)$. We will show that for every $a = 0, 1, \ldots, n$,

$$
A_T(a) \to_{T\to\infty} 0. \qquad (32)
$$

We already proved the results for $a = 0, 1$, so we assume $a \ge 2$. We write

$$
\begin{aligned}
& A_T(a)\\
&= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k} (T\log p)^{-a(2k+1)}\\
& \left(\int_0^{T\log p} y^{2k}\right)^a\\
&= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} (T\log p)^{2k-2ka-a}\\
& \int_0^{T\log p} \ldots \int_0^{T\log p} dy_1 \ldots dy_a (y_1 y_2 \ldots y_a)^{2k}\\
&= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\le T} \frac{1}{\sqrt{p}} \int_0^{T\log p} \ldots \int_0^{T\log p} dy_1 \ldots y_a\\
& \frac{1}{T\log p} \frac{1}{y_1 .. y_a} \sum_{k\ge 0} \frac{(-1)^k}{(2k+1)!} \left(\frac{y_1 .. y_a}{(T\log p)^{a-1}}\right)^{2k+1}
\end{aligned}
$$

$$= \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \int_0^{T\log p} \dots \int_0^{T\log p} dy_1 \dots y_a$$
$$\frac{1}{T\log p} \frac{1}{y_1..y_a} \sin(y_1..y_a(T\log p)^{1-a}).$$

By the change of variables $\tilde{y}_i = \frac{y_i}{T\log p}$ for $i = 1, .., a$, we obtain

$$A_T(a) = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p\leq T} \frac{1}{\sqrt{p}} \int_0^1 \dots \int_0^1 dy_1 \dots y_a \frac{1}{T\log p} \frac{1}{y_1..y_a}$$
$$\times \sin(y_1..y_a(T\log p)).$$

By choosing $\delta \in (0, 1)$ small enough and by writting

$$\frac{\sin(y_1..y_a(T\log p))}{y_1..y_a(T\log p)} = \left(\frac{\sin(y_1..y_a(T\log p))}{y_1..y_a(T\log p)}\right)^\delta \left(\frac{\sin(y_1..y_a(T\log p))}{y_1..y_a(T\log p)}\right)^{1-\delta},$$

so that

$$\left|\frac{\sin(y_1..y_a(T\log p))}{y_1..y_a(T\log p)}\right| \leq \left(\frac{1}{y_1..y_a(T\log p)}\right)^{1-\delta},$$

we can bound $A_T(a)$ as follows:

$$\begin{aligned} A_T(a) &\leq C\frac{1}{\sqrt{\frac{1}{2}\log\log T}} T^{\delta-1} \sum_{p\leq T} \frac{1}{\sqrt{p}} (\log p)^{\delta-1} \\ &\leq C\frac{1}{\sqrt{\frac{1}{2}\log\log T}} T^{\delta-1} \sum_{p\leq T} \frac{1}{\sqrt{p}} \leq C\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \frac{T^{\delta-1/2}}{\log T} \end{aligned} \tag{33}$$

and this converges to zero as $T \to \infty$ if $\delta < \frac{1}{2}$. In the last inequality we used the following estimate (see, e.g., [16, 15]): for every $s \in \mathbb{C}$ with $\Re s < 1$ we have

$$\sum_{p\leq x} p^{-s} \sim \frac{x^{1-s}}{(1-s)\log x}. \tag{34}$$

From (31) and (33) we deduce that for every $n \geq 1$ and for T large enough, $|c_{2n}(T)| \leq C\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \frac{T^{\delta-1/2}}{\log T}$ with $\delta \in (0, \frac{1}{2})$ and the conclusion follows.

□

4.2.2 *On the chaos decomposition of* $\log\zeta$ *on the critical line*

Let us finish by some remarks concerning the asymptotic behavior of the chaos projection of $\log\left|\zeta\left(\frac{1}{2} + iTU\right)\right|$. This random variable is obviously square integrable. It is close to the Dirichlet series (6) in the sense that (see [13, 7])

$$\mathbf{E}\left|\log|\zeta\left(\frac{1}{2} + \mathbf{i}TU\right)| - \sum_{p\leq T} \frac{\cos(T\log pU)}{\sqrt{p}}\right|^2 \leq C,$$

so

$$\mathbf{E}\left|\frac{1}{\sqrt{\log\log T}}\log|\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)| - \frac{1}{\sqrt{\log\log T}}\sum_{p\leq T}\frac{\cos(T\log pU)}{\sqrt{p}}\right|^2$$
$$\leq C\frac{1}{\log\log T}\to_{T\to\infty} 0. \tag{35}$$

From the results in the previous paragraph, we can deduce the asymptotic behavior of the chaoses that compose $\log\zeta$ on the real line. We have the following result.

Proposition 2. *Let U be given by (10). Assume that for every $T>0$ the square integrable random variable $\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\log\left|\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)\right|$ admits the chaos decomposition*

$$\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\Re\log\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)=\sum_{n\geq 0}K_n(T).$$

Then for every $n\geq 1$, $K_n(T)\to_{T\to\infty} 0$ in $L^2(\Omega)$.

Proof. Since $\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_p\frac{1}{\sqrt{p}}\cos(T\log pU)=\sum_{n\geq 0}J_{2n}(T)$ with J_n from (28), we can write, by using (35),

$$\mathbf{E}\left|\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\Re\log\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)-\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{1}{\sqrt{p}}\cos(T\log pU)\right|^2$$
$$=\sum_{n\geq 0}\mathbf{E}\left|K_{2n}(T)-J_{2n}(T)\right|^2+\sum_{n\geq 0}\mathbf{E}|K_{2n+1}(T)|^2$$
$$\leq C\frac{1}{\frac{1}{2}\log\log T}.$$

This clearly implies that for every $n\geq 1$,

$$K_n(T)\to_{T\to\infty} 0 \text{ in } L^2(\Omega).$$

□

Let us make some comments on the result stated in Proposition 2. We denoted by $K_n(T)$ the nth chaos component of the random variable $F_T:=\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\times\log\left|\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)\right|$, i.e for every $T>0$ we have $F_T=\sum_{n\geq 0}K_n(T)$. We showed that for every $n\geq 0$, $K_n(T)\to_{T\to\infty} 0$ almost surely and in $L^2(\Omega)$. Let us discuss the meaning of this result. Suppose that we have a random sequence $(X_T)_{T>0}$ which converges in distribution, as $T\to\infty$, to the standard normal distribution. Assume that for each $T>0$ the random variable X_T admits a chaos decomposition

$$X_T=\sum_{n\geq 0}I_n(f_n(T)).$$

In many situations, for such a limit theorem, there exists a dominant chaos for X_T (see, e.g., Theorem 1.7 in [10] for an example on the average of solutions to some stochastic partial differential equations, or [8] for an example related to stochastic geometry). That is, there exists $N_0 \geq 1$ such that $I_{N_0}(f_{N_0}(T))$ converges in law, as $T \to \infty$, to $N(0, 1)$ while the other chaos components are negligible, i.e. $\mathbf{E} I_n(f_n(T))^2 \to_{T\to\infty} 0$ for $n \neq N_0$. This situation would be very convenient because, in order to understand the behavior of X_T, it suffices to look at the convergence of the chaos of order N_0.

In other situations, each chaos converges to a Gaussian limit, i.e. $I_n(f_n(T)) \to_{T\to\infty} N(0, \sigma_n^2)$ in law for every $n \geq 0$ and $\sum_{n\geq 0} \sigma_n^2 = 1$. We refer to, e.g., the paper [5] for such a situation.

We actually showed that we are not in the situations described above. We showed that for every $n \geq 0$, the behavior of $K_n(T)$ is very close to the behavior of the nth chaos component of the Dirichlet series (6) (denoted by $J_n(T)$ in (29)) and we proved that $J_n(T)^2 \to_{T\to\infty} 0$ almost surely and in L^2. This follows from the fact that $J_n(T) = 0$ if n is odd and $J_{2n}(T) = c_{2n}(T) I_{2n}(h_{2n})$ with $c_n(T)$ a deterministic sequences converging to zero as $T \to \infty$. This suggests the complexity of the Selberg theorem since the contribution to the limit comes from every chaos.

Acknowledgments

We will like to thank the referee for having carefully read our work.

Funding

C. Tudor was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01) and MATHAMSUD project SARC (19-MATH-06).

References

[1] Bourgade, P.: Mesoscopic fluctuations of the zeta zeros. Probab. Theory Relat. Fields **148**(3-4), 479–500 (2010). MR2678896. https://doi.org/10.1007/s00440-009-0237-3

[2] Coram, M., Diaconis, P.: New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function. J. Phys. A, Math. Gen. **36**, 2883–2906 (2003). Random matrix theory. MR1986397. https://doi.org/10.1088/0305-4470/36/12/302

[3] Erdös, P.: On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. Proc. Natl. Acad. Sci. USA **35**, 374–384 (1949). MR29411. https://doi.org/10.1073/pnas.35.7.374

[4] Hadamard, J.: Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. Bull. Soc. Math. Fr. **24**, 199–220 (1896). MR1504264

[5] Hu, Y., Nualart, D.: Renormalized self-intersection local time for fractional Brownian motion. Ann. Probab. **33**(3), 948–983 (2005). MR2135309. https://doi.org/10.1214/009117905000000017

[6] Hughes, C.P., Nikeghbali, A., Yor, M.: An arithmetic model for the total disorder process. Probab. Theory Relat. Fields **141**(1-2), 47–59 (2008). MR2372965. https://doi.org/10.1007/s00440-007-0079-9

[7] Kühn, P.: On Selberg's central limit theorem. Master Thesis, ETH Zurich.

[8] Nourdin, I., Peccati, G., Rossi, M.: Nodal statistics of planar random waves. Commun. Math. Phys. **369**(1), 99–151 (2019). MR3959555. https://doi.org/10.1007/s00220-019-03432-5

[9] Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Probability and its Applications (New York), p. 382. Springer (2006). MR2200233

[10] Nualart, D., Zheng, G.: Averaging Gaussian functionals. Electron. J. Probab. **25**, 48–54 (2020). MR4092767. https://doi.org/10.1214/20-ejp453

[11] Selberg, A.: On the normal density of primes in small intervals, and the difference between consecutive primes. Arch. Math. Naturvidensk. **47**(6), 87–105 (1943). MR12624

[12] Selberg, A.: On the remainder in the formula for $N(T)$, the number of zeros of $\zeta(s)$ in the strip $0 < t < T$. Avh. Nor. Vidensk.-Akad. Oslo, I **1944**(1), 27 (1944). MR15426

[13] Selberg, A.: Contributions to the theory of the Riemann zeta-function. Arch. Math. Naturvidensk. **48**(5), 89–155 (1946). MR20594

[14] Tao, T.: Selberg's limit theorem for the Riemann zeta function on the critical line. Blog article.

[15] Titchmarsh, E.C.: The Theory of the Riemann Zeta-function, 2nd edn., p. 412. The Clarendon Press, Oxford University Press, New York (1986). Edited and with a preface by D.R. Heath-Brown. MR882550

[16] Tsang, K.: Selberg's limit theorem for the Riemann zeta function on the critical line. Blog article.