数理解析研究所講究録
第 1590 巻 2008 年 126-145

Three Notes on Connections between the Riemann Zeta Function and Probability Theory, in particular: Random Matrix Theory.

M.Yor [(1)(2)]

March 12, 2008

[(1)] Laboratoire de Probabilités et Modèles Aléatoires,
Universités Paris VI et VII, 4 Place Jussieu - Case 188,
F-75252 Paris Cedex 05
E-mail:deaproba@proba.jussieu.fr

[(2)] Institut Universitaire de France

A - Random Matrices and the Riemann Zeta function: the Keating-Snaith philosophy

B - A further note on Selberg's integrals, inspired by N. Snaith's results about the distribution of some characteristic polynomials

C - On the logarithm of the Riemann Zeta function: from Selberg's central limit theorem to total disorder

****

Some pertinent comments about each of the Notes A, B, C have been made by P. Bourgade; they are presented just after C.

A - Random Matrices and the Riemann Zeta function: the Keating-Snaith philosophy

M.Yor [(1)(2)]

March 3, 2008

[(1)] Laboratoire de Probabilités et Modèles Aléatoires,
Universités Paris VI et VII, 4 Place Jussieu - Case 188,
F-75252 Paris Cedex 05
E-mail:deaproba@proba.jussieu.fr

[(2)] Institut Universitaire de France

Abstract

The extremely precise conjecture of Keating-Snaith about the asymptotics of the moments of the Riemann Zeta function on the critical line, as the height T tends to $+\infty$ is presented, together with some striking similarities between the Riemann Zeta asymptotics, as $T \to \infty$, and asymptotics about the generic matrix A_N on the unitary group U_N, as $N \to \infty$. Explicit Mellin-Fourier computations done by Keating-Snaith about (A_N) are interpreted probabilistically. Further heuristics for the (KS) conjecture are also discussed.

1 The Keating-Snaith conjecture

(1.1) The importance of the Riemann Hypothesis:

$$\text{(RH)} \quad \boxed{\begin{array}{l}\text{All non-trivial zeros of the Zeta function} \\ (\zeta(s); s \in \mathbb{C} \setminus \{1\}) \text{ lie on the critical line: Re } (s) = \dfrac{1}{2}\end{array}}$$

justifies the intensive studies which keep being developed about the behavior of $\left\{\zeta\left(\frac{1}{2} + it\right); t \in \mathbb{R}\right\}$.

In particular, (RH) implies the (still unproven) Lindelöf hypothesis:

$$\left|\zeta\left(\frac{1}{2}+it\right)\right| = 0(t^{\varepsilon})\,,\ t \to \infty$$

for any $\varepsilon > 0$.
This conjecture can be shown to be equivalent to another one, relative to the moments of ζ on the critical line, namely: for every $k \in \mathbb{N}$,

$$I_k(T) \stackrel{\text{def}}{=} \frac{1}{T}\int_0^T ds \left|\zeta\left(\frac{1}{2}+is\right)\right|^{2k} = 0(T^{\varepsilon})\,,\ \text{as } T \to \infty$$

again for any $\varepsilon > 0$.
Until now, it has been shown rigorously that:

- $I_1(T) \underset{T\to\infty}{\sim} \log T$; [Hardy-Littlewood (1918)]
- $I_2(T) \underset{T\to\infty}{\sim} \frac{1}{2\pi^2}(\log T)^4$; [Ingham (1926)]

(1.2) These two results, together with a number of other arguments led Keating-Snaith [5] to formulate the extremely precise conjecture

$$\text{(KS)} \qquad \boxed{\forall k \in \mathbb{N}\,,\ I_k(T) \underset{T\to\infty}{\sim} H_{\mathcal{P}}(k) H_{\text{Mat}}(k)(\log T)^{k^2}}$$

where $H_{\mathcal{P}}(k)$ is a factor which takes care of the "arithmetic" of the set of primes $\mathcal{P}$, while $H_{\text{Mat}}(k)$ is a factor which takes more into account some hidden randomness and is associated with some asymptotics, as $N \to \infty$, of the characteristic polynomial

$$Z(A_N, \theta) = \det\left(I_N - A_N e^{-i\theta}\right)$$

where A_N is the generic unitary matrix, distributed with the Haar probability measure on U_N.
(1.3) The remainder of this Note is organized as follows:

- in Section 2, two strikingly similar results between asymptotics for:
  - ⋆ on one hand, the Riemann Zeta function, on the critical line, as the height T tends to $+\infty$;
  - ⋆ on the other hand, (A_N) asymptotics, as $N \to \infty$,

  are presented;
- in Section 3, explicit computations of Keating-Snaith related to (A_N) are discussed;

- in Section 4, I shall come back to the Keating-Snaith conjecture, and present some attempt by Gonek-Hughes-Keating [3] to "justify" the conjecture from a purely Riemann Zeta function perspective.

Convention: When discussing some points pertaining to the Riemann Zeta function, I shall use a box $\boxed{\text{NT}}$ meaning Number Theory, whereas when discussing some point about Random Matrix Theory, I shall use $\boxed{\text{RMT}}$.

2 Similarities between Riemann Zeta asymptotics, as $T \to \infty$, and (A_N) asymptotics, as $N \to \infty$

(2.1) The pair correlation laws of H. Montgomery (:NT) and F. Dyson (:RMT)

$\boxed{\text{NT}}$ For simplicity of exposition, let us assume here the validity of the Riemann hypothesis, and write all non-trivial roots of the Zeta function as:

$$\frac{1}{2} \pm it_n\,,\ 0 < t_1 < t_2 < \cdots < t_n < \cdots$$

Let $w_n = \frac{t_n}{2\pi} \log\left(\frac{t_n}{2\pi}\right)$; then, a key step towards an analytical proof of the prime number theorem is that, denoting: $N(W) = \sharp\{n; w_n \le W\}$, then:

$$\frac{N(W)}{W} \underset{W\to\infty}{\longrightarrow} 1$$

As a further step, the following quantities have been considered:

$$\frac{1}{W}\sharp\{(w_n, w_m) \in [0,W]^2; \alpha \le w_n - w_m \le \beta\}$$

and, more generally, for a generic function f:

$$R_2(f,W) = \frac{1}{W} \sum_{w_j, w_k \le W} f(w_j - w_k)$$

Then, there is the following

Theorem 1. *(H. Montgomery [6]; 1973)*
Define $\widehat{f}(\tau) = \int_{-\infty}^{+\infty} dx\, f(x)\exp(2i\pi\tau x)$*; then, if* $supp(\widehat{f}) \subset [-1,+1]$*, there is the limiting result:*

$$R_2(f,W) \underset{W\to\infty}{\longrightarrow} \int_{-\infty}^{+\infty} dx\, f(x) R_2(x)\,,\ \textit{where } R_2(x) \equiv 1 - \left(\frac{\sin(\pi x)}{\pi x}\right)^2$$

Furthermore, H. Montgomery has conjectured that this result holds true even if $\widehat{f}$ does not have compact support, but this is still open.

$\boxed{\text{RMT}}$ We now recall the strikingly similar result due to F. Dyson [2], about the pair correlation for the eigenvalues of $A_N \in U(N)$; denoting these eigenvalues by their arguments: $(\theta_1, \theta_2, \cdots, \theta_N)$, taking values in $[0, 2\pi[$, we let: $\phi_n = \frac{N}{2\pi}\theta_n$. Then, F. Dyson's result is:

$$\frac{1}{N}\int_{U(N)} d\mu_{U(N)}(A)\sharp\{(n,m); \alpha < \phi_n - \phi_m \leq \beta\}$$
$$\xrightarrow[N\to\infty]{} \int_\alpha^\beta dx \left(1 - \left(\frac{\sin(\pi x)}{\pi x}\right)^2\right)$$

(2.2) The central limit theorems of A. Selberg (:NT) and Keating-Snaith (:RMT)

$\boxed{\text{NT}}$ A. Selberg [7] proved the following result:

$$\frac{1}{T}\int_T^{2T} dt\, 1\left\{\frac{\log\zeta\left(\frac{1}{2}+it\right)}{\sqrt{\frac{1}{2}\log\log T}} \in \Gamma\right\} \xrightarrow[T\to\infty]{} \iint_\Gamma \frac{dxdy}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)$$

where Γ is any bounded regular Borel set in $\mathbb{C}$, ie: $\partial\Gamma$ is negligible for Lebesgue measure. This result translates as follows in probabilistic terms: if one considers, on the probability space $(u(\in [1,2]), du)$ the variables:

$$L_T(u) = \frac{\log\zeta\left(\frac{1}{2}+iuT\right)}{\sqrt{\frac{1}{2}\log(\log T)}},$$

then Selberg's result states that L_T converges in distribution to $\mathcal{N} \equiv N_1 + iN_2$, where N_1 and N_2 are two standard independent, centered, variance 1, Gaussian variables.
I discuss a multidimensional extension of Selberg's theorem in note **C**.

$\boxed{\text{RMT}}$ For $\theta \in [0, 2\pi[$, and A distributed with the Haar measure $\mu_{U(N)}$, consider

$$Z(A,\theta) = \det(I_N - Ae^{-i\theta}) \equiv \prod_{n=1}^N (1 - e^{i(\theta_n - \theta)}),$$

where $(e^{i\theta_1}, \cdots, e^{i\theta_N})$ are the eigenvalues of A.
Then, Keating-Snaith [5] have proven:

$$\int_{U(N)} d\mu_{U(N)}(A)\, 1_{\left\{\frac{\log Z(A,\theta)}{\sqrt{\frac{1}{2}\log N}} \in \Gamma\right\}} \xrightarrow[N\to\infty]{} \iint_\Gamma \frac{dxdy}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)$$

where again, Γ is a regular bounded Borel set in $\mathbb{C}$.
Formally, this result resembles Selberg's, when one takes: $T = \exp(N)$ (or, more generally, $\exp(N^\lambda)$).

3 Explicit results of Keating-Snaith about the law of Z_N, which lead to an interpretation of Z_N as a product of independent variables.

(3.1) Keating-Snaith [5] were able to calculate the generating function of the characteristic polynomial $Z(A,\theta)$, when A is distributed according to the Haar measure $\mu_{U(N)}(dA)$ (consequently, by rotational invariance, the law of $Z(A,\theta)$ does not depend on θ); precisely:

$$E[|Z_N(A,\theta)|^t \exp(is \arg Z_N(A,\theta))]$$

$$(1) \qquad = \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(t+j)}{\Gamma\left(j+\frac{t+s}{2}\right)\Gamma\left(j+\frac{t-s}{2}\right)}$$

which yields, in particular, taking $s = 0$:

$$(2) \qquad E[|Z_N(A,\theta)|^{2k}] \underset{N\to\infty}{\sim} H_{Mat}(k) N^{k^2} ,$$

where, thanks to the asymptotics of the gamma function:

$$(3) \qquad H_{Mat}(k) \equiv \prod_{j=1}^{k-1} \frac{j!}{(j+k)!}$$

Note that it is this constant $H_{Mat}(k)$ which is featured in the statement of the KS conjecture, presented in Section 1.
(3.2) We now give some indications about some representations of $Z_N(A,\theta)$ as a product of independent random variables; this may be done purely by interpreting formula (1) in terms of beta and gamma variables, or, independently from (1), by constructing the Haar measure $\mu_{U(N)}$ in a recursive manner.

• A beta-gamma interpretation of (1)
For simplicity, let us only consider $s = 0$ in (1), so that (1) now expresses the Mellin transform of $|Z_N(A,\theta)|$. It is immediate, from the expression of the Mellin transform of a γ_j variable, that is: a gamma variable with parameter j whose density is:

$$\frac{P(\gamma_j \in dt)}{dt} = \frac{t^{j-1}e^{-t}}{\Gamma(j)}$$

that (1) yields:

$$\prod_{j=1}^{N} \gamma_j \overset{\text{(law)}}{=} |Z_N| \prod_{j=1}^{N} (\gamma_j \gamma'_j)^{\frac{1}{2}} \tag{4}$$

where all the random variables in sight are assumed independent.

• A recursive construction of the Haar measures $\mu_{U(N)}$

This recursive construction, which is lifted here from [1], yields, as a consequence, the following stochastic representation of Z_N:

$$Z_N \overset{\text{(law)}}{=} \prod_{k=1}^{N} (1 + e^{i\theta_k} (\beta_{1,k-1})^{\frac{1}{2}}) , \tag{5}$$

where, on the RHS, the θ_k's are independent uniforms on $[0, 2\pi[$, independent of the beta variables with indicated parameters. We recall:

$$P(\beta_{a,b} \in du) = \frac{u^{a-1}(1-u)^{b-1}}{B(a,b)} \, du \qquad (0 < u < 1)$$

We leave to the interested reader the task of verifying that (5) implies (4).

I now explain the recursive construction, and how (5) follows from it.

a) Let $M \in U_N$ such that its first column M_1 is uniformly distributed on the unit complex sphere:

$$S_{\mathbb{C}}^{N-1} = \{(c_1, \ldots, c_N) \in \mathbb{C}^N \, ; \, |c_1|^2 + \ldots + |c_N|^2 = 1\}$$

Then, if $A_{N-1} \in U_{N-1}$ is chosen independently of M according to the Haar measure $\mu_{U_{N-1}}$, the matrix:

$$A_N \overset{\text{def}}{=} M \begin{pmatrix} 1 & 0 \\ 0 & A_{N-1} \end{pmatrix} \tag{6}$$

is distributed with the Haar measure μ_{U_N}.

b) One easily deduces from (6) that:

$$det(I_N - A_N) \overset{\text{(law)}}{=} (1 - M_{11}) det(I_{N-1} - A_{N-1}) \tag{7}$$

with M_{11} and A_{N-1} independent.

Since M_1 is uniform (see a) above), one obtains readily that:

$$M_{11} \overset{\text{(law)}}{=} e^{i\theta_N} (\beta_{1,N-1})^{\frac{1}{2}} , \tag{8}$$

where θ_N is uniform on $[0, 2\pi[$, and independent from $\beta_{1,N-1}$.

c) By iteration, (7) and (8) imply (5).

4 Further heuristics for the (KS) conjecture

A main difficulty inherent to the (KS) conjecture is: how to "see" the random matrix part in terms of the Riemann Zeta function?
This is the aim of the paper by Gonek-Hughes-Keating [3], which I only discuss in vague terms:
(i) In [3], the authors "factorize" approximately $\zeta\left(\frac{1}{2}+it\right)$ as:

$$\zeta\left(\frac{1}{2}+it\right) \sim P_X\left(\frac{1}{2}+it\right) Z_X\left(\frac{1}{2}+it\right) ,$$

where X is a real parameter, $X \geq 2$, and:

$$P_X\left(\frac{1}{2}+it\right) \sim \prod_{\substack{p\in\mathcal{P} \\ p\leq X}} (1-p^{-\frac{1}{2}-it})^{-1}$$

but I need to refer the reader to Theorem 1 of [3] for a precise definition of P_X and Z_X.
(ii) The authors make the Splitting Conjecture:

$$I_k(T,\zeta) \sim I_k(T,P_X)\bullet I_k(T,Z_X)$$

when X and T tend to $+\infty$, with $X = 0((\log T)^{2-\varepsilon})$ and we note:

$$I_k(T,f) \stackrel{\text{def}}{=} \frac{1}{T}\int_0^T dt \left|f\left(\frac{1}{2}+it\right)\right|^{2k}$$

(iii) They prove:

$$I_k(T,P_X) = H_{\mathcal{P}}(k)(e^{\gamma}\log X)^{k^2}\left(1+0_k\left(\frac{1}{\log X}\right)\right)$$

(iv) They conjecture:

$$I_k(T,Z_X) \sim H_{Mat}(k)\left(\frac{\log T}{e^{\gamma}\log X}\right)^{k^2} ,$$

when X and T tend to ∞, with $X = 0((\log T)^{2-\varepsilon})$.

* * * * * * * * * * * * *

Thus, clearly, the conjunction (ii), (iii), and (iv) yields the (KS) conjecture.

****

Comment: I apologize for this very rough "first aid" treatment of the (KS) conjecture. Despite quite some evidence, it is really tough to make NT and RMT meet there, but nonetheless, we are learning a number of "facts" in one or the other domain, on our way.

References

[1] P. Bourgade, C.P. Hughes, A. Nikeghbali, M. Yor. The characteristic polynomial of a random unitary matrix: a probabilistic approach. To appear in Duke Math. Journal (2008).

[2] F.J. Dyson. Statistical theory of the energy levels of complex systems, J. Math. Phys., 3, p. 140–175 (1962).

[3] S.M. Gonek, C.P. Hughes, J.P. Keating. A hybrid Euler-Hadamard product formula for the Riemann Zeta function. Duke Math. J. 136, p. 507–549, (2007).

[4] J.P. Keating. L-functions and the characteristic polynomials of Random Matrices. In: Recent Perspectives in Random Matrix Theory and Number Theory. London Math. Soc. Lecture Notes Series 322. eds: F. Mezzadri, N.C. Snaith. Cambridge Univ. press.

[5] J.P. Keating, N.C. Snaith. Random Matrix Theory and $\zeta\left(\frac{1}{2}+it\right)$, Comm. Math. Phys. 214, p. 57–89, (2000).

[6] H. Montgomery. The pair correlation of the zeta function. Proc. Symp. Pure Math. 24, p. 181–193, (1973).

[7] A. Selberg. Old and New conjectures and results about a class of Dirichlet series. Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), p. 367–385, Univ. Salerno, Salerno, (1992).

B - A further note on Selberg's integrals, inspired by N. Snaith's results about the distribution of some characteristic polynomials

M.Yor [(1)(2)]

March 3, 2008

[(1)] Laboratoire de Probabilités et Modèles Aléatoires,
Universités Paris VI et VII, 4 Place Jussieu - Case 188,
F-75252 Paris Cedex 05
E-mail:deaproba@proba.jussieu.fr

[(2)] Institut Universitaire de France

Abstract

The derivative at 1 of the characteristic polynomial of the generic random matrix valued in $SO(2N+1)$ is shown to be a product of N independent beta variables. A similar discussion is done with respect to the celebrated Selberg distributions.

1 A probabilistic discussion of some results from N. Snaith

(1.1) For a matrix $U \in SO(2N+1)$, distributed with the Haar measure:
a) the characteristic polynomial takes the form:

$$\Lambda_U(e^{i\theta}) = (1 - e^{-i\theta}) \prod_{n=1}^{N} (1 - e^{i(\theta_n - \theta)})(1 - e^{i(-\theta_n - \theta)})$$

b) hence, it admits the derivative at $e^{i0} = 1$:

$$\begin{aligned}\Lambda'_U(1) &= \frac{d}{d\alpha}\left[(1-e^{-\alpha})\prod_{n=1}^{N}(1-e^{i\theta_n-\alpha})(1-e^{-i\theta_n-\alpha})\right]\Big|_{\alpha=0} \\ &= \prod_{n=1}^{N}|1-e^{i\theta_n}|^2 = 2^N\prod_{n=1}^{N}(1-\cos\theta_n)\end{aligned}$$

c) N. Snaith ([3], bottom of p. 101) studies the distribution of $\Lambda'_U(1)$, starting from its Mellin transform:

$$(1) \qquad E[(\Lambda'_U(1))^s] = 2^{2Ns}\prod_{j=1}^{N}\left(\frac{\Gamma\left(\frac{1}{2}+s+j\right)}{\Gamma\left(\frac{1}{2}+j\right)}\right)\left(\frac{\Gamma(N+j)}{\Gamma(s+N+j)}\right)$$

The right-hand side of (1) is easily understood as the Mellin transform of a product of independent beta variables.
Indeed, recall the "beta-gamma algebra", in its most elementary form:

$$(2) \qquad \gamma_a \overset{\text{(law)}}{=} \beta_{a,b}\bullet\gamma_{a+b}\,,$$

where γ_a and γ_{a+b} denote two gamma variables, with respective parameters a and $(a+b)$, and $\beta_{a,b}$ a beta variable with parameters (a,b), i.e:

$$(3) \qquad \begin{cases} P(\gamma_a \in dt) = \dfrac{t^{a-1}e^{-t}dt}{\Gamma(a)}\,,\ t>0\,, \\[2ex] P(\beta_{a,b}\in du) = \dfrac{u^{a-1}(1-u)^{b-1}}{B(a,b)}\,du\,,\ u\in(0,1)\,, \end{cases}$$

On the RHS of (2), $\beta_{a,b}$ and $\gamma_{a,b}$ are assumed to be independent. Throughout this paper, products of r.v's will occur with, unless otherwise mentioned, independent r.v's.
It follows immediately from the Mellin transform of γ_a, which is:

$$(4) \qquad E[(\gamma_a)^s] = \frac{\Gamma(a+s)}{\Gamma(a)}\quad,\qquad s\geq 0$$

and from (2) that the Mellin transform of $\beta_{a,b}$ is:

$$(5) \qquad E[(\beta_{a,b})^s] = \left(\frac{\frac{\Gamma(a+s)}{\Gamma(a)}}{\frac{\Gamma(a+b+s)}{\Gamma(a+b)}}\right)\quad,\qquad s\geq 0$$

Consequently, we deduce from (1) that:

$$(6) \qquad \Lambda'_U(1) \overset{\text{(law)}}{=} 2^{2N}\prod_{j=1}^{N}\beta_{\left(\frac{1}{2}+j,\,N-\frac{1}{2}\right)}$$

where, as just indicated, the RHS only involves independent beta variables.

(1.2) With the help of the Mellin transform (1), N. Snaith [3] obtains a precise equivalent of the density of $\Lambda'_U(1)$, which we shall denote here by $\delta_N(y)$, near $y = 0$.

I shall now show how (6) provides this equivalent. For this purpose, I denote $\widetilde{\beta}_j = \beta_{(\frac{1}{2}+j, N-\frac{1}{2})}$, and its density by $(b_j(u), u \in [0,1])$.

To compute the density $\delta_N(y)$, we may write, for every $f : \mathbb{R}_+ \to \mathbb{R}_+$, Borel:

$$E[f(2^{2N}\prod_{j=1}^{N}\widetilde{\beta}_j)] = \int_0^1 du\, b_1(u) E[f(2^{2N}u\prod_{j=2}^{N}\widetilde{\beta}_j)]$$

and the change of variables:

$$u = \frac{y}{2^{2N}\prod_{j=2}^{N}\widetilde{\beta}_j}$$

together with Fubini's theorem, yields the formula:

$$\delta_N(y) = E\left[\frac{1}{2^{2N}\prod_{j=2}^{N}\widetilde{\beta}_j}\, b_1\left(\frac{y}{2^{2N}\prod_{j=2}^{N}\widetilde{\beta}_j}\right)\right] \tag{7}$$

From (3), we deduce:

$$b_1(u) = \frac{u^{1/2}(1-u)^{N-3/2}}{B\left(\frac{1}{2}, N-\frac{3}{2}\right)}$$

Thus, (7) yields the equivalent:

$$\delta_N(y) \underset{y\to 0}{\sim} \frac{y^{1/2}}{B\left(\frac{1}{2}, N-\frac{3}{2}\right)} E\left[\left(2^{2N}\prod_{j=2}^{N}\widetilde{\beta}_j\right)^{-3/2}\right] \tag{8}$$

The RHS of (8) equals:

$$\frac{y^{1/2}}{B\left(\frac{1}{2}, N-\frac{3}{2}\right)} 2^{-3N}\prod_{j=2}^{N} E[(\widetilde{\beta}_j)^{-3/2}] = y^{1/2} f(N)\,,$$

with:

$$f(N) = \frac{2^{-3N}}{B\left(\frac{1}{2}, N-\frac{3}{2}\right)}\prod_{j=2}^{N}\frac{B\left(j-1, N-\frac{1}{2}\right)}{B\left(\frac{1}{2}+j, N-\frac{1}{2}\right)}$$

This constant is also easily seen to be equal to:

$$f(N) = \frac{2^{-3N}}{B\left(\frac{1}{2}, N-\frac{3}{2}\right)}\frac{1}{\Gamma(N)}\prod_{j=2}^{N}\frac{\Gamma(j+N)\Gamma(j)}{\Gamma\left(j+\frac{1}{2}\right)\Gamma\left(j+N-\frac{3}{2}\right)}$$

which agrees with N. Snaith's constant given in formula (2.10) in [3].

2 Extending the discussion to Selberg's distributions

Here, we shall call Selberg's distributions, and denote these by ${}^{(N)}\sum\limits_{a,b}^{c}$, the following probabilities on $[-1,+1]^N$, indexed by $a>0, b>0, c\geq 0$:
(9)

$$ {}^{(N)}\sum_{a,b}^{c}(dx_1,\ldots,dx_N)=\frac{1}{C_{a,b}^{(c)}}|\Delta(x)|^{2c}\prod_{j=1}^{N}(1-x_j)^{a-1}(1+x_j)^{b-1}dx_1\ldots dx_N $$

where $C_{a,b}^{(c)}$ is the normalizing constant given by Selberg's formula:
(10)

$$ C_{a,b}^{(c)}=2^N[c(N-1)+a+b-1]\prod_{j=0}^{N-1}\frac{\Gamma(1+c(1+j))\Gamma(a+jc)\Gamma(b+jc)}{\Gamma(1+c)\Gamma(a+b+c(N+j-1))} $$

and

$$ \Delta(x)=\prod_{1\leq j<\ell\leq N}(x_j-x_\ell)\,, $$

for

$$ x=(x_j)_{1\leq j\leq N}\in[-1,+1]^N $$

It seems of interest (and this will allow us to relate the following with N. Snaith's results as presented above) to consider the joint distribution of
(11)

$$ {}^{(-)}X\stackrel{\text{def}}{=}\prod_{j=1}^{N}(1-x_j)\,;\ {}^{(+)}X\stackrel{\text{def}}{=}\prod_{j=1}^{N}(1+x_j)\,;\ |\Delta(x)|^2=\prod_{1\leq j<\ell\leq N}(x_j-x_\ell)^2 $$

under ${}^{(N)}\sum\limits_{a,b}^{c}$.

For this purpose, we may replace in (10), the triplet (a,b,c) by: $(a+s,b+t,c+u)$; to begin with, let us take $u=0$.
Then, we obtain:

$$ (12)\quad {}^{(N)}\sum_{a,b}^{c}\{({}^{(-)}X)^s({}^{(+)}X)^t\}=2^{N(s+t)}\prod_{j=0}^{N-1}\left(\frac{\Phi_j^{(N)}(a+s,b+t,c)}{\Phi_j^{(N)}(a,b,c)}\right) $$

where: $\Phi_j^{(N)}(a,b,c)=\dfrac{\Gamma(a+jc)\Gamma(b+jc)}{\Gamma(a+b+c(N+j-1))}$.
Recall again that the Mellin transform of a gamma variable is given by:

$$ E[(\gamma_a)^s]=\frac{\Gamma(a+s)}{\Gamma(a)} $$

Then, we can interpret (12) as follows:

$$
\begin{aligned}
& E\left[\left(\prod_{j=0}^{N-1}\gamma_{a+b+c(N+j-1)}\right)^{s+t}\left(\frac{1}{2^N}{}^{(-)}X\right)^s\left(\frac{1}{2^N}{}^{(+)}X\right)^t\right] \\
= \; & E\left[\prod_{j=0}^{N-1}(\gamma_{a+jc})^s(\gamma_{b+jc})^t\right]
\end{aligned}
$$

with all gamma variables independent between themselves, and independent of the pair $({}^{(-)}X, {}^{(+)}X)$; thus, with the same notation, and $({}^{(-)}X, {}^{(+)}X)$ being still considered under ${}^{(N)}\sum_{a,b}^{c}$, we see that:

$$
\begin{aligned}
& \left(\prod_{j=0}^{N-1}\gamma_{a+b+c(N+j-1)}\right)\frac{1}{2^N}({}^{(-)}X, {}^{(+)}X) \\
(13) \qquad & \overset{\text{(law)}}{=} \left(\prod_{j=0}^{N-1}\gamma_{a+jc}, \prod_{j=0}^{N-1}\gamma_{b+jc}\right)
\end{aligned}
$$

To simplify formula (13), we now use the beta-gamma algebra as follows:

$$(\maltese) \quad (\gamma_{a+jc}, \gamma_{b+jc}) \overset{\text{(law)}}{=} (\beta_{a+jc,b+jc}\gamma_{a+b+2jc}; (1-\beta_{a+jc,b+jc})\gamma_{a+b+2jc})$$

$$(\dagger) \qquad \gamma_{a+b+2jc} \overset{\text{(law)}}{=} \beta_{(a+b+2jc,c((N-1)-j))}\gamma_{a+b+c((N-1)+j)}$$

Importing $(\maltese)$ and $(\dagger)$ on the RHS of (13), we obtain, after simplification of both sides by:

$$\prod_{j=0}^{N-1}\gamma_{a+b+c(N+j-1)}$$

the identity in law:

$$
\begin{aligned}
(14)\ \frac{1}{2^N}({}^{(-)}X, {}^{(+)}X) \overset{\text{(law)}}{=} & \left(\prod_{j=0}^{N-1}\beta^{(j)}_{a+jc,b+jc}\cdot\beta'^{(j)}_{a+b+2jc,c((N-1)-j)}\,,\right. \\
& \left.\prod_{j=0}^{N-1}(1-\beta^{(j)}_{a+jc,b+jc})\cdot\beta'^{(j)}_{a+b+2jc,c((N-1)-j)}\right)
\end{aligned}
$$

From this identity (14), we may derive quite a number of consequences: a) with the help of the identity in law (which is easily derived from (2)):

$$\beta_{a,b}\,\beta'_{a+b,c} \overset{\text{(law)}}{=} \beta_{a+b,c}\,,$$

we obtain:

$$(\diamondsuit) \qquad \frac{1}{2^N}{}^{(-)}X \overset{\text{(law)}}{=} \prod_{j=0}^{N-1} \beta^{(j)}_{a+jc,b+c(N-1)}$$

$$(\S) \qquad \frac{1}{2^N}{}^{(+)}X \overset{\text{(law)}}{=} \prod_{j=0}^{N-1} \beta^{(j)}_{b+jc,a+c(N-1)}$$

Note the remarkable feature from $(\diamondsuit)$: although under ${}^{(N)}\sum_{a,b}^{c}$, the components $(1-x_j)$ are not independent, their product ${}^{(-)}X$ may be written as a product of independent beta variables; of course, we may make the same remark concerning ${}^{(+)}X$.

b) Going back to (14), we also note that:

$$(15) \qquad \frac{{}^{(-)}X}{{}^{(+)}X} \overset{\text{(law)}}{=} \prod_{j=0}^{N-1} \frac{\beta^{(j)}_{a+jc,b+jc}}{(1-\beta^{(j)}_{a+jc,b+jc})}$$

which, again, from the beta-gamma algebra, may be written as:

$$(\clubsuit) \qquad \frac{{}^{(-)}X}{{}^{(+)}X} \overset{\text{(law)}}{=} \frac{\prod_{j=0}^{N-1} \gamma^{(j)}_{a+jc}}{\prod_{j=0}^{N-1} \gamma^{(j)}_{b+jc}}$$

where, here, on the RHS, the numerator and denominator are independent.

Thus, similarly to the remark in a) above, although under ${}^{(N)}\sum_{a,b}^{c}$ these variables ${}^{(-)}X$ and ${}^{(+)}X$ are not independent, their ratio may be expressed as a ratio of independent variables.

c) The previous identities, e.g: (13) in particular, may also be used in order to obtain a recurrence relation between the laws of

$$({}^{(-)}X^{(N)}, {}^{(+)}X^{(N)}) \quad \text{and} \quad ({}^{(-)}X^{(N-1)}, {}^{(+)}X^{(N-1)})$$

(the parameters a and b may vary, but c remains fixed throughout).

3 From Selberg's generalized beta distributions to Selberg's generalized gamma distributions

(3.1) We first make an elementary change of variables in formula (9), i.e: $x_i = 1 - 2y_i$, so that now y_i takes values in $[0, 1]$. We denote by

${}^{(N)}\widetilde{\sum}^{c}_{a,b}$, the image of ${}^{(N)}\sum^{c}_{a,b}$ obtained from this change of variables

Thus:
(16)

$$ {}^{(N)}\widetilde{\sum}^{c}_{a,b}(dy_1, ..., dy_N) = \frac{1}{\widetilde{C}^{(c)}_{a,b}} |\Delta(y)|^{2c} \prod_{j=1}^{N} \left((y_j)^{a-1}(1-y_j)^{b-1} dy_j \right) $$

where $\widetilde{C}^{(c)}_{a,b} = \dfrac{C^{(c)}_{a,b}}{2^{N[c(N-1)+a+b-1]}}$.
To summarize the main result of Section 2, we simply write:

$$ \text{under } {}^{(N)}\widetilde{\sum}^{c}_{a,b}, $$

$$ (17) \qquad Y_N \overset{\text{def}}{=} \prod_{j=1}^{N} y_j \quad \text{is distributed as:} \quad \prod_{j=0}^{N-1} \beta^{(j)}_{a+jc,b+c(N-1)} $$

(3.2) We now wish to develop a similar discussion, when the beta variables are replaced by gamma ones. For this purpose, let us note that:

$$ \gamma_a \overset{\text{(law)}}{=} (b\,\beta_{a,b}) \left(\frac{\gamma_{a+b}}{b} \right) $$

so that, letting $b \to \infty$, we obtain:

$$ b\,\beta_{a,b} \xrightarrow[b\to\infty]{\text{(law)}} \gamma_a $$

This remark allows to introduce the probabilities:

$$ {}^{(N)}\Gamma^{c}_{a}(dy_1, ..., dy_N) = \frac{|\Delta(y)|^{2c}}{D^{(c)}_{a}} \prod_{j=1}^{N} (y_j^{a-1} e^{-y_j} dy_j) $$

where: $D^{(c)}_{a} = E[|\Delta(\gamma^{(j)}_{a}; j \le N)|^{2c}]$ and the result (17) now becomes:

$$ (18) \qquad \text{under } {}^{(N)}\Gamma^{c}_{a},\ Y_N \overset{\text{def}}{=} \prod_{j=1}^{N} y_j \text{ is distributed as: } \prod_{j=0}^{N-1} \gamma^{(j)}_{a+jc} $$

4 Final comments

(4.1) Prior to her paper [3], N. Snaith wrote [4], in which she calculated the Mellin transform of the n^{th} derivative of the characteristic polynomial averaged over the subset of matrices with n eigenvalues conditioned to lie at 1.
Again, as in the present discussion, the result can be interpreted in

terms of a product of independent variables.

(4.2) A crucial ingredient in N. Snaith's calculations is the use of the Selberg integrals; however, with the help of recursive constructions of the Haar measures, as the dimension increases, representations of the variables of interest as products of independent variables arise naturally. See [2] for a first development of this viewpoint, and P. Bourgade [1] for a more complete picture.

(4.3) The present discussion is much more modest, as it simply exploits the beta-gamma algebra in order to interpret a number of results due to N. Snaith, and obtained with analytic methods. For more in the same vein, see Yor [5].

References

[1] P. Bourgade. Circular ensembles and independence. In preparation, March 2008.

[2] P. Bourgade, C.P. Hughes, A. Nikeghbali, M. Yor. The characteristic polynomial of a random unitary matrix: a probabilistic approach. To appear in Duke Math. Journal (2008).

[3] N.C. Snaith. The derivative of $SO(2N+1)$ characteristic polynomials and rank 3 elliptic curves. In: London Math. Society, Lect. Notes Series 341, eds: J. Conrey, D. Farmer, F. Mezzadri, N.C. Snaith. Cambridge Univ. Press (2007), p.93–107.

[4] N.C. Snaith. Derivatives of a random matrix characteristic polynomials with applications to elliptic curves. J. Phys. A. Math. Gen 38 (2005), p. 10345–10360.

[5] M. Yor. A Note about Selberg's Integrals in Relation with the Beta-Gamma Algebra. In: Advances in Mathematical Finance, p. 49–58. Festschrift for D. Madan. Birkhäuser (2007).

C - On the logarithm of the Riemann Zeta function: from Selberg's central limit theorem to total disorder

M.Yor [(1)(2)]

March 3, 2008

[(1)] Laboratoire de Probabilités et Modèles Aléatoires,
Universités Paris VI et VII, 4 Place Jussieu - Case 188,
F-75252 Paris Cedex 05
E-mail:deaproba@proba.jussieu.fr

[(2)] Institut Universitaire de France

Abstract

Looking for a process version of the central limit theorem of Selberg for the logarithm of the Riemann Zeta function produces only "total disorder", and not a reasonable stochastic process. A number of comments about this result are made.

****

A well-known result of Selberg [2] states that the classical continuous determination of the logarithm of the Riemann Zeta function is asymptotically normally distributed, in the sense that, if Γ is a regular Borel measurable subset of $\mathbb{C}$, then:

$$(1)\quad \lim_{T\to\infty}\frac{1}{T}\int_T^{2T} 1_{\left\{\frac{\log\zeta\left(\frac{1}{2}+it\right)}{\sqrt{\frac{1}{2}\log\log T}}\in\Gamma\right\}}dt = \frac{1}{2\pi}\int_\Gamma \exp\left(-\frac{x^2+y^2}{2}\right)dxdy$$

where $1_{\{\Lambda\}}$ is the indicator of Λ, and regular means that the boundary of Γ has 0 Lebesgue measure.
If we let :

$$L_\lambda(N,u) = \frac{\log\left(\frac{1}{2}+iuN^\lambda\right)}{\sqrt{\log N}}$$

then Selberg's result may be stated as:

$$\lim_{N\to\infty}\int_1^2 1_{\{L_\lambda(N,u)\in\Gamma\}}du = P(G_\lambda\in\Gamma) \tag{2}$$

where $G_\lambda = G_\lambda^{(1)} + \mathrm{i}G_\lambda^{(2)}$ is a complex valued Gaussian random variable with mean 0 and variance $\left(\frac{\lambda}{2}\right)$, ie: $G_\lambda^{(1)}$ and $G_\lambda^{(2)}$ are independent, centered, and:

$$E[(G_\lambda^{(1)})^2] = E[(G_\lambda^{(2)})^2] = \frac{\lambda}{2}\,.$$

It is now a natural question, at least from a probabilistic standpoint, to look for an asymptotic distribution of the vectors (considered as r.v's on $([1,2],du)$) $(L_{\lambda_1}(N,\cdot),\ldots,L_{\lambda_k}(N,\cdot))$ for $0<\lambda_1<\lambda_2<\ldots<\lambda_k<\infty$.
This question has been resolved as follows:

Theorem 1. *([1]): For $0<\lambda_1<\lambda_2<\ldots<\lambda_k<\infty$, and for every $(\Gamma_j, j\le k)$ regular,*

$$\lim_{N\to\infty}\int_1^2 1_{\{\bigcap_{j=1}^k(L_{\lambda_j}(N,u)\in\Gamma_j)\}}du = \prod_{j=1}^k P(G_{\lambda_j}\in\Gamma_j) \tag{3}$$

The remainder of this Note shall consist in commenting about this result.

Comment 1. ***a)*** *If $(D_\lambda = D_\lambda^{(1)} + iD_\lambda^{(2)}, \lambda>0)$ is a totally disordered complex valued Gaussian process, meaning that $(D_\lambda^{(1)},\lambda>0)$ and $(D_\lambda^{(2)},\lambda>0)$ are two independent Gaussian processes all of whose coordinates are independent, with*

$$E[(D_\lambda^{(1)})^2] = E[(D_\lambda^{(2)})^2] = \frac{\lambda}{2}\,,$$

then the quantity on the RHS of (3) is:

$$P(D_{\lambda_1}\in\Gamma_1,\ldots,D_{\lambda_k}\in\Gamma_k)$$

b) *The totally disordered real-valued Gaussian process $(D_\lambda,\lambda>0)$ barely deserves the name of "process", as it does not admit any measurable version $(\widetilde{D}_\lambda)$; indeed, if so, by Fubini, this version would satisfy:*

$$\int_a^b \widetilde{D}_\lambda d\lambda = 0\,,\ a.s.\,,$$

hence: $\widetilde{D}_\lambda = 0$, $d\lambda dP$, which is absurd.

Comment 2. *a) In [1], Theorem 1 is proven using the* <u>*method of moments*</u>*, following carefully and adapting Selberg's original arguments to our multidimensional study.*
It might be interesting to be able to use another method, i.e: the characteristic function method of Paul Lévy.
b) The method of moments was used in the original proof by Kallianpur-Robbins of the following result:

$$\frac{1}{\log T}\int_0^T ds\, f(Z_s) \xrightarrow[T\to\infty]{(law)} \left(\frac{1}{2\pi}\overline{f}\right)\mathbf{e}\,, \tag{4}$$

where: $(Z_s, s \geq 0)$ *denotes planar Brownian motion,*

- $f : \mathbb{C} \to \mathbb{R}$ *is bounded, with compact support;* $\overline{f} = \int_{\mathbb{C}} dxdy\, f(x,y)$;
- $\mathbf{e}$ *is a standard exponential variable.*

However, "more Brownian" techniques allow to prove (4) via asymptotics of one-dimensional Brownian local times, and also - unlike the present study - to obtain an interesting <u>*process*</u> *result when replacing* T *in (4) by* N^λ, *for* $\lambda > 0$. *(For details, see, e.g., [3], Chap. XIII.) Thus, in this way,* $\log\left(\zeta\left(\frac{1}{2}+it\right)\right)$ *is more wildly random than Brownian motion.*

References

[1] C.P. Hughes, A. Nikeghbali, M. Yor. An arithmetic model for the total disorder process. Proba. Th. Rel. Fields, 141, p. 47–59, (2008).

[2] A. Selberg. Old and New conjectures and results about a class of Dirichlet series. Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), p. 367–385, Univ. Salerno, Salerno, (1992).

[3] D. Revuz, M. Yor. Continuous Martingales and Brownian Motion. Springer (1999), 3rd edition.