Théorie des nombres/*Number Theory*

The n-level correlations of zeros of the zeta function

Zeév RUDNICK and Peter SARNAK

Abstract – We define the n-level correlation sums for the normalized zeros of Riemann's zeta function. A special case of these is the pair correlation function investigated by H. Montgomery. If one assumes the Riemann Hypothesis, then these sums measure the n-level correlations of the imaginary parts of the zeros, a statistic introduced by F. Dyson. It is shown that for a restricted class of test functions the n-level correlations follow the distribution predicted by the Gaussian Unitary Ensemble of random matrix theory. More generally, the same is true universally for the principal L-functions attached to cuspidal automorphic representations of $GL(m)$ over the rationals.

Les corrélations de niveau n des zéros de la fonction zêta

Résumé – *Nous introduisons les sommes des corrélations de niveau n pour les zéros normalisés de la fonction zêta de Riemann. Un cas particulier est la corrélation entre paires de zéros étudiée par H. Montgomery. Si l'on admet l'hypothèse de Riemann, ces sommes donnent la corrélation de niveau n entre les parties imaginaires des zéros, une statistique introduite par F. Dyson. On montre que, pour une classe restreinte de fonctions-test, les corrélations de niveau n ont la distribution prédite par l'Ensemble Unitaire Gaussien de la théorie des matrices aléatoires.*

To Paul Cohen on the occasion of his 60-th birthday

Version française abrégée – Dans cette Note, nous nous intéressons à la distribution des zéros de la fonction zêta de Riemann $\zeta(s)$. Soient $\rho_j = 1/2 + i\gamma_j$, $j = \pm 1, \pm 2 \dots$ les zéros non-triviaux de $\zeta(s)$. Pour motiver la présentation de nos résultats, nous commençons par admettre l'hypothèse de Riemann (RH) : $\gamma_j \in \mathbf{R}$. Ordonnons les parties imaginaires $\gamma_1 \leqq \gamma_2 \leqq \dots$ (en tenant compte des multiplicités). Leur nombre est asymptotiquement

$$(0.1) \qquad N(T) := \#\{ j \geqq 1 : \gamma_j \leqq T\} \sim \frac{1}{2\pi} T \log T, \qquad T \to \infty.$$

Il s'ensuit que les nombres $\tilde{\gamma}_j = (\gamma_j \log \gamma_j)/2\pi$ ont pour valeur moyenne d'espacement l'unité. Nous étudions les statistiques qui déterminent la distribution des $\tilde{\gamma}_j$.

D'après le travail de Montgomery [4] sur la corrélation entre paire de zéros, et les résultats numériques d'Odlyzko ([5], [6]), il est bien accepté maintenant que la distribution des espacements consécutifs suit l'Ensemble Unitaire Gaussien de la théorie des matrices aléatoires. Autrement dit, si nous définissons les espacements normalisés par $\delta_j = \tilde{\gamma}_{j+1} - \tilde{\gamma}_j$, alors pour toute fonction $f \in C_c^\infty(0, \infty)$, on doit avoir $(1/N) \sum_{j=1}^{N} f(\delta_j) \to \int_0^\infty f(x) P(x)\, dx$ où $P(x)\,dx$ est la distribution des espacements normalisés des valeurs propres des grandes matrices hermitiennes aléatoires (GUE distribution) [3].

La distribution des niveaux d'espacements consécutifs est déterminée par les corrélations de niveau n, notées R_n. Ces dernières mesurent la distribution des différences de n éléments de $B_N = \{\tilde{\gamma}_1 \leqq \dots \leqq \tilde{\gamma}_N\}$, $N \to \infty$.

Pour des fonctions-test lisses $f(x_1, \dots, x_n)$ satisfaisant :

TF 1. $f(x_1, \dots, x_n)$ *est symétrique*;

TF 2. $f(x + t(1, \dots, 1)) = f(x)$ *pour* $t \in \mathbf{R}$;

TF 3. $f(x) \to 0$ *rapidement quand* $|x| \to \infty$ *dans l'hyperplan* $\sum_j x_j = 0$,

---

Note présentée par Jean-Pierre SERRE.

0764-4442/94/03191027 $ 2.00

définissons

$$R_n(B_N, f) = \frac{n!}{N} \sum_{\substack{S \subset B_N \\ |S|=n}} f(S) \tag{0.2}$$

où $f(S) = f(a_1, \ldots, a_n)$ si $S = \{a_1, \ldots, a_n\}$, qui est bien définie d'après **TF 1**. La connaissance du comportement asymptotique de $R_n(B_N, f)$ quand $N \to \infty$ est équivalente à la connaissance de celui des corrélations lisses

$$R_n(T, f, h) = \sum_{j_1, \ldots, j_n}{}' h\left(\frac{\gamma_{j_1}}{T}\right) \ldots h\left(\frac{\gamma_{j_n}}{T}\right) f\left(\frac{L}{2\pi}\gamma_{j_1}, \ldots, \frac{L}{2\pi}\gamma_{j_n}\right) \tag{0.3}$$

pour une famille suffisamment riche de fonctions localisées (par exemple les fonctions de Schwartz). Ici $L = \log T$ et $\sum'$ indique la somme sur les indices distincts. Observons que, du fait que h est localisée, γ est d'ordre T, et par conséquent la normalisation $(L/2\pi)\gamma$ est la même que $\tilde{\gamma}$.

Dans [1], Dyson a déterminé la densité $W_n(x_1, \ldots, x_n)$ des corrélations de niveau n pour le modèle GUE. Il a montré que celle-ci est donnée par le déterminant $n \times n$

$$W_n(x_1, \ldots, x_n) = \det(K(x_i - x_j)), \qquad K(x) = \frac{\sin \pi x}{\pi x}. \tag{0.4}$$

THÉORÈME 0.1. – *Soit f satisfaisant TF 1, 2, 3 et supposons de plus que le support de $\hat{f}(\xi)$ est contenu dans $\sum_j |\xi_j| < 2$. Soient $g \in C_c^\infty(\mathbf{R})$ et $h(r) = \int_{-\infty}^{\infty} g(u)\, e^{iru}\, du$. Alors, quand $T \to \infty$,*

$$R_n(T, f, h) \sim \frac{1}{2\pi} T \log T \int_{-\infty}^{\infty} h(r)^n\, dr \times \int_{\mathbf{R}^n} f(x)\, W_n(x)\, \delta\left(\frac{x_1 + \ldots + x_n}{n}\right) dx_1 \ldots dx_n$$

où $\delta(x)$ est la mesure de Dirac en 0.

THÉORÈME 0.2. – *Admettons RH et supposons que le support de $\hat{f}(\xi)$ est contenu dans $\sum_j |\xi_j| < 2$. Alors, quand $N \to \infty$,*

$$R_n(B_N, f) \to \int_{\mathbf{R}^n} f(x)\, W_n(x)\, \delta\left(\frac{x_1 + \ldots + x_n}{n}\right) dx_1 \ldots dx_n.$$

---

1. THE RIEMANN ZETA FUNCTION. – This Note is concerned with the distribution of the zeros of the Riemann zeta function $\zeta(s)$. Let $\rho_j = 1/2 + i\gamma_j$, $j = \pm 1, \pm 2 \ldots$ denote the non-trivial zeros of $\zeta(s)$. To motivate the formulation of our results we begin by assuming the Riemann Hypothesis (RH) : $\gamma_j \in \mathbf{R}$. We order the imaginary parts by $\gamma_1 \leqq \gamma_2 \leqq \ldots$ (counted with multiplicities). Their number is asymptotically

$$N(T) := \#\{j \geqq 1 : \gamma_j \leqq T\} \sim \frac{1}{2\pi} T \log T, \qquad T \to \infty. \tag{1.1}$$

It follows that the numbers $\tilde{\gamma}_j := (\gamma_j \log \gamma_j)/2\pi$ have mean spacing unity. We investigate the statistics governing the distribution of the $\tilde{\gamma}_j$'s.

Following the original calculation by Montgomery [4] of the pair correlation (assuming *RH* and for a restricted class of test functions) and the extensive numerical investigations by Odlyzko ([5], [6]), it is by now well accepted that the distribution of consecutive spacings follow the Gaussian Unitary Ensemble of Random Matrix Theory. That is, if we define the normalized spacings $\delta_j := \tilde{\gamma}_{j+1} - \tilde{\gamma}_j$, then for any localized f on $(0, \infty)$, one should have: $(1/N) \sum_{n=1}^{N} f(\delta_n) \to \int_0^\infty f(x) P(x)\, dx$ where $P(x)\, dx$ is the distribution of normalized spacings of large random Hermitian matrices (GUE distribution), which was determined by Gaudin and Mehta [3].

The consecutive level spacing distribution is determined by the n-level correlation functions R_n which are defined as follows: Let $B_N = \{\tilde{\gamma}_1 \leqq \ldots \leqq \tilde{\gamma}_N\}$. The n-level correlations measure the distribution of the differences of n elements of B_N, $N \to \infty$. That is, for a box $Q \subset \mathbf{R}^{n-1}$ set

$$(1.2) \qquad R_n(B_N, Q) = \frac{1}{N} \#\{j_1, \ldots, j_n \leqq N, \text{ distinct} : (\tilde{\gamma}_{j_1} - \tilde{\gamma}_{j_2}, \ldots, \tilde{\gamma}_{j_{n-1}} - \tilde{\gamma}_{j_n}) \in Q\}.$$

An equivalent but technically more convenient way to measure this distribution is to use smooth test functions $f(x_1, \ldots, x_n)$ satisfying:

TF 1. $f(x_1, \ldots, x_n)$ is symmetric;

TF 2. $f(x + t(1, \ldots, 1)) = f(x)$ for $t \in \mathbf{R}$;

TF 3. $f(x) \to 0$ rapidly as $|x| \to \infty$ in the hyperplane $\sum_j x_j = 0$.

Define

$$(1.3) \qquad R_n(B_N, f) = \frac{n!}{N} \sum_{\substack{S \subset B_N \\ |S| = n}} f(S).$$

Here $f(S) = f(a_1, \ldots, a_n)$ if $S = \{a_1, \ldots, a_n\}$ and in view of **TF 1** this is well defined. **TF 2** asserts that f is a function of the successive differences and hence we recover what (1.2) seeks to measure. It turns out that knowing the asymptotic behaviour of $R_n(B_N, f)$ as $N \to \infty$ is equivalent to knowing that of the smoothed correlations

$$(1.4) \qquad R_n(T, f, h) = \sum_{j_1, \ldots, j_n}{}' h\left(\frac{\gamma_{j_1}}{T}\right) \ldots h\left(\frac{\gamma_{j_n}}{T}\right) f\left(\frac{L}{2\pi}\gamma_{j_1}, \ldots, \frac{L}{2\pi}\gamma_{j_n}\right)$$

for a sufficiently rich family of localized cutoff functions h (e.g. Schwartz functions). Here $L = \log T$, and $\sum'$ means sum over distinct indices. Note that since h localizes γ to be of order T, the normalization $(L/2\pi)\gamma$ is the same as $\tilde{\gamma}$.

In [1], Dyson determined the n-level correlation density $W_n(x_1, \ldots, x_n)$ for the GUE model. He showed it is given by $W_n(x_1, \ldots, x_n) = \det(K(x_i - x_j))$, $K(x) = \sin \pi x / \pi x$.

It is important to note that if h and f are defined for complex arguments and are localized then the sums (1.4) make sense even if we do not assume RH, and we still refer to these as the n-level correlations. As explained above, RH and the GUE model can be used to predict their asymptotic behaviour. Our first result proves that this prediction is correct for a restricted class of f's.

THEOREM 1.1. – *Let f satisfy **TF 1, 2, 3** and in addition assume that $\hat{f}(\xi)$ is supported in $\sum_j |\xi_j| < 2$. Let $g \in C_c^\infty(\mathbf{R})$ and $h(r) = \int_{-\infty}^{\infty} g(u)\, e^{iru}\, du$ (so that h and f are* *entire) and let* $\delta(x)$ *be the Dirac mass at* 0. *Then as* $T \to \infty$,

$$R_n(T, f, h) \sim \frac{1}{2\pi} T \log T \int_{-\infty}^{\infty} h(r)^n \, dr \times \int_{\mathbf{R}^n} f(x) W_n(x) \delta\left(\frac{x_1 + \dots + x_n}{n}\right) dx_1 \dots dx_n.$$

If we assume RH we can relax the smoothness condition on h and in fact choose a characteristic function of an interval. In this way we can prove that the n-level correlations of the zeros are GUE distributed at least for f's with restricted Fourier transforms.

THEOREM 1.2. – *Assume RH and that the support of* $\hat{f}(\xi)$ *is contained in* $\sum_j |\xi_j| < 2$. *Then as* $N \to \infty$

$$R_n(B_N, f) \to \int_{\mathbf{R}^n} f(x) W_n(x) \delta\left(\frac{x_1 + \dots + x_n}{n}\right) dx_1 \dots dx_n.$$

Remarks: 1) The condition $\sum_j |\xi_j| < 2$ is "natural" in the sense that it is exactly the region in which the "multi-diagonals" give the main term in the asymptotics. Beyond this region the off-diagonal terms contribute equally. One conjectures [4] that theorem 1.2 holds without the support restriction on $\hat{f}$–that is to say that the n-level correlations are those of GUE. E. Bogomolny and J. Keating have informed us that they have a heuristic proof of this (work in progress). The Euclidean volume of the region $\sum_{j=1}^{n} |\xi_j| < 2$ decreases rapidly as $n \to \infty$ and this unfortunately prevents us from using certain choices for f in theorem 1.2.

2) As mentioned before, theorem 1.2 for $n = 2$ is due to Montgomery. Hejhal [2] has recently established theorem 1.2 for the case $n = 3$ (in somewhat different form). Our approach is different to the one taken by these two authors. Firstly it directly establishes theorem 1.1, which allows us to avoid appealing to the Riemann Hypothesis, and secondly it considerably facilitates the computations for general n.

3) In [7], we also establish similar results for principal L-functions associated to cuspidal automorphic representations of $GL(m)_{/\mathbf{Q}}$, showing that the GUE model applies universally to the zeros of all these L-functions. The last is somewhat surprising since the coefficients of the series defining these L-functions do not have a universal statistical behaviour.

2. OUTLINE OF THE PROOFS. – We give a brief outline of the main steps in proving theorems 1.1 and 1.2. We use a smooth version of Riemann's Explicit Formula: If $g \in C_c^\infty(\mathbf{R})$ and $h(r) = \int_{-\infty}^{\infty} g(u) e^{iru} \, du$ then

$$\text{(2.1)} \quad \sum_\gamma h(\gamma) = h\left(\frac{i}{2}\right) + h\left(-\frac{i}{2}\right) + \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r) \left(\frac{\Gamma'_{\mathbf{R}}}{\Gamma_{\mathbf{R}}}\left(\frac{1}{2} + ir\right) + \frac{\Gamma'_{\mathbf{R}}}{\Gamma_{\mathbf{R}}}\left(\frac{1}{2} - ir\right)\right) dr - \sum_{n=1}^{\infty} \left(\frac{\Lambda(n)}{\sqrt{n}} g(\log n) + \frac{\Lambda(n)}{\sqrt{n}} g(-\log n)\right)$$

where $\Lambda(n) = \log p$, if $n = p^k$ is a prime power, and is 0 otherwise, and $\Gamma_{\mathbf{R}}(s) = \pi^{-s/2}\,\Gamma(s/2)$ is the Gamma factor appearing in the functional equation of $\zeta(s)$. Let

$$C_n(T, f, h) = \sum_{j_1, \dots, j_n} h\left(\frac{\gamma_{j_1}}{T}\right) \dots h\left(\frac{\gamma_{j_n}}{T}\right) f\left(\frac{L}{2\pi}\gamma_{j_1}, \dots, \frac{L}{2\pi}\gamma_{j_n}\right) \tag{2.2}$$

(note that here we are not assuming that the indices j_k are distinct). Then for f satisfying **TF 1, 2, 3** and $\hat{f}$ of compact support

$$C_n(T, f, h) = \int_{\mathbf{R}^n} \prod_{k=1}^{n} \left\{ \sum_{j_k} h\left(\frac{\gamma_{j_k}}{T}\right) e^{-iL\gamma_{j_k}\xi_k} \right\} d\mu(\xi) \tag{2.3}$$

where $d\mu(\xi) = \Phi(\xi)\,\delta(\xi_1 + \dots + \xi_n)\,d\xi$ is the Fourier transform of f. We apply the explicit formula (2.1) to the sum in (2.3). After multiplying out the resulting series over primes and a lengthy computation (which exploits the condition $\sum_j |\xi_j| < 2$) we find that as $T \to \infty$

$$\text{(2.4)} \quad \begin{cases} C_n(T, f, h) = \dfrac{TL}{2\pi} \displaystyle\int_{-\infty}^{\infty} h(r)^n\,dr \int_{\mathbf{R}^n} C_{\underline{O}}(v)\,\Phi(v)\,dv + O(T) \\ \displaystyle\int_{\mathbf{R}^n} C_{\underline{O}}(v)\,\Phi(v)\,dv = \Phi(0) + \sum_{r=1}^{[n/2]} \sum \int |v_1| \dots |v_r| \\ \qquad\qquad \times \Phi(v_1\,\mathbf{e}_{i(1), j(1)} + \dots + v_r\,\mathbf{e}_{i(r), j(r)})\,dv_1 \dots dv_r \end{cases}$$

where the sum is over all choices of r disjoint pairs of indices $i(t) < j(t)$ in $\{1, \dots, n\}$, and for $i < j$ we set $\mathbf{e}_{i,j} = \mathbf{e}_i - \mathbf{e}_j$, $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ are the standard basis vectors.

To prove theorem 1.1 one must now relate $C_n(T, f, h)$ which is a sum over all indices, to $R_n(T, f, h)$ which is a sum over distinct ones. To this end we need some combinatorial sieving. The crux is to establish certain identities between the distribution $C_{\underline{O}}(v)$ and certain related distributions and the GUE determinant W_n.

Let $\underline{F} = [F_1, \dots, F_\nu]$ be a set partition of $\{1, \dots, n\}$. For instance, we set $\underline{O} = [\{1\}, \dots, \{n\}]$. We define $C_{\underline{F}}(v)$ to be the distribution obtained from $C_{\underline{O}}(v)$ when the variables v_l, $l \in F_j$, are set equal to each other. Then

$$C_{\underline{F}}(v) = \delta_{\underline{F}}(v) + \sum_{r=1}^{[n/2]} \sum \prod_{t=1}^{r} \delta_{F_{i(t)} \cup F_{j(t)}}(v) \,\Big|\sum_{l \in F_{i(t)}} v_l\Big| \prod_{k \neq i(a), j(b)} \delta_{F_k}(v) \tag{2.5}$$

the sum being over all possible pairs of subsets $(F_{i(t)}, F_{j(t)})$, no repetitions allowed. The δ-functions are defined by $\delta_S(v) = \delta(\sum_{l \in S} v_l)$ for $S \subset \{1, \dots, n\}$ and for a set partition $\underline{F}$ as above, $\delta_{\underline{F}}(v) = \prod_{j=1}^{\nu} \delta_{F_j}(v)$.

The distribution $R_{\underline{O}}(v)$ that dictates the asymptotics of $R_n(T, f, h)$ may be expressed via Möbius inversion as

$$R_{\underline{O}}(v) = \sum_{\underline{F}} \mu(\underline{O}, \underline{F})\,C_{\underline{F}}(v) \tag{2.6}$$

where $\mu(\underline{O}, \underline{F}) = \prod_{j=1}^{\nu(\underline{F})} (-1)^{|F_j|-1} (|F_j| - 1)!$. The key identity is the following:

PROPOSITION 2.1. – *For* $\sum_j |v_j| < 2$, *the Fourier transform* $\widehat{W}_n(v)$ *equals* $R_{\underline{O}}(v)$.

Proposition 2.1 in turn is based on the following identity between piecewise linear functions. It couples the pairing structure (2.4), (2.5) and the cycle structure of the determinant defining W_n.

PROPOSITION 2.2. – *Let* $v \in \mathbf{R}^n$ *with* $\sum_j v_j = 0$. *Then*

$$\sum_{[F, F^c]} (|F| - 1)!(|F^c| - 1)! |\sum_{l \in F} v_l|$$
$$= \sum_{\{\theta\}} \max\{v_{\theta(1)}, v_{\theta(1)} + v_{\theta(2)}, \ldots\} - \min\{v_{\theta(1)}, v_{\theta(1)} + v_{\theta(2)}, \ldots\}$$

where the first sum is over all partitions of $\{1, \ldots, n\}$ *into two subsets, and the second sum is over all orderings* $\theta = (\theta(1), \ldots, \theta(n))$ *of* $\{1, \ldots, n\}$ *modulo rotations.*

The proof of proposition 2.2 is based on Spitzer's combinatorial method [8]. Finally, the passage from theorems 1.1 to 1.2 is a straight-forward approximation argument. Complete details of this proof, as well as the precise statements and proofs for the case of general L-functions will appear in [7].

Supported by NSF grants DMS-9102082 and DMS-9400163.

Note remise le 20 juillet 1994, acceptée le 15 septembre 1994.

REFERENCES

[1] F. J. DYSON, Statistical Theory of the Energy Levels of Complex Systems, III, *J. Math. Phys.*, 3, 1962, pp. 166-175.

[2] D. A. HEJHAL, On the Triple Correlations of Zeros of the Zeta Function, *preprint.*

[3] M. L. MEHTA, *Random Matrices*, Second Edition, Academic Press, 1991.

[4] H. L. MONTGOMERY, The Pair Correlation of Zeros of the Zeta Functions, *Proc. Sympos. Pure Math.*, 24, Amer. Math. Soc., Providence 1973, pp. 181-193.

[5] A. M. ODLYZKO, On the Distribution of Spacings Between Zeros of Zeta Functions, *Math. Comp.*, 48, 1987, pp. 273-308.

[6] A. M. ODLYZKO, The 10^{20} Zero of the Riemann Zeta Function and 70 Million of its Neighbors, *preprint*, A. T. & T., 1989.

[7] Z. RUDNICK and P. SARNAK, Zeros of Principal L-Functions and Random Matrix Theory, *preprint.*

[8] F. SPITZER, A Combinatorial Lemma and its Applications to Probability Theory, *Trans. Amer. Math. Soc.*, 82, 1956, pp. 323-339.

Department of Mathematics, Princeton University, Princeton NJ 08544, USA.