Entropy of Riemann zeta zero sequence

O. Shanker *

Abstract

One of the key unsolved problems in mathematics is the proof or refutation of Riemann's remarkable 1858 hypothesis about the location of the roots of the Riemann zeta function. The statistical properties of the zero distribution have been studied intensively to get insight into the phenomenon. In this work we study the entropy of the sequence of zeros, which tells us how constrained is the pattern of zeros. A high value of the entropy would imply that the sequence has relatively low structure, and hence predicting the zeros is a difficult problem. A low value would give us encouragement that techniques in machine learning like neural networks would be helpful in studying the phenomenon.

*Mountain View, CA 94041, U. S. A. Email: oshanker@gmail.com

***AMO - Advanced Modeling and Optimization. 1841-4311**

zero count N	*Probability*
0	0.162
1	0.678
2	0.158
3	0.002

Table 1: Distribution of zero counts in a Gram interval

1 Introduction

The zero spacings of the Riemann Zeta function [1, 2, 3, 4] is a topic of deep abiding interest to mathematicians and physicists. In this work we study the entropy [5] of the sequence of zeros. Entropy is a measure of the role of probability in generating a sequence of values. It gives us an indication of how how constrained is the pattern of zeros. A high value of the entropy implies that predicting the zeros is a difficult problem, while a low value is an indication that the sequence has some structure, and hence techniques in machine learning like neural networks [6] would be helpful in studying the phenomenon.

In Section 2 we define the Gram interval and the sequence of zeros. Section 3 gives the calculation of the entropy for the sequence of zeros. The conclusions are presented in Section 4.

2 Gram Interval and zero distribution

In this section we define the Gram interval and the sequence of zeros. The concept of Gram interval is of particular importance, since it is the basis for defining the entropy. The Riemann Zeta function is defined for $\mathrm{Re}(s) > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \in primes} \left(1 - p^{-s}\right)^{-1}. \tag{1}$$

Eq. (1) converges for $\mathrm{Re}(s) > 1$. $\zeta(s)$ has a continuation to the complex plane and satisfies a functional equation

$$\xi(s) := \pi^{-s/2}\,\Gamma(s/2)\,\zeta(s) = \xi(1-s); \tag{2}$$

$\xi(s)$ is entire except for simple poles at $s = 0$ and 1. We write the zeroes of $\xi(s)$ as $1/2 + i\gamma$. The Riemann Hypothesis asserts that γ is real for the non-trivial zeroes. We order the γs in increasing order, with

$$\ldots\ldots\gamma_{-1} < 0 < \gamma_1 \leq \gamma_2 \ldots. \tag{3}$$

Then $\gamma_j = -\gamma_{-j}$ for $j = 1, 2, \ldots$, and γ_1, γ_2, … are roughly 14.1347, 21.0220, ….

Asymptotically, for the Riemann zeta function the mean number of zeros with height less than γ (the smoothed Riemann zeta staircase) is [4]

$$< \mathcal{N}_{\mathcal{R}}(\gamma) >= (\gamma/2\pi)(ln(\gamma/2\pi) - 1) - \frac{7}{8}. \tag{4}$$

Thus, the mean spacing of the zeros at height γ is $2\pi(\ln(\gamma/2\pi))^{-1}$. For the range of t values studied in this work this spacing is essentially constant at 0.109.

In our study an important role is played by the "Gram Points" and Gram intervals, which we now define. One defines

$$\theta(t) = arg(\pi^{-it/2}\Gamma(\frac{1}{4} + \frac{it}{2})), \tag{5}$$

where the argument is defined by continuous variation of t starting with the value 0 at $t = 0$. θ has the aymptotic expansion for large t:

$$\theta(t) = \frac{t}{2}\ln(\frac{t}{2\pi}) - \frac{t}{2} - \frac{\pi}{8} + \frac{1}{48t} - \frac{1}{5760t^3}. \tag{6}$$

Tthe function $Z(t) = exp(i\theta(t))\zeta(1/2 + it)$, known as the Riemann-Siegel Z-function, is real valued for real t and $|Z(t)| = |\zeta(1/2 + it)|$. Thus the zeros of $Z(t)$ are the imaginary part of the zeros of $\zeta(s)$ which lie on the critical line. Many of the zeros are separated by the "Gram points". When $t \geq 7$, the θ function Eq.(5) is monotonic increasing. For $n \geq 1$, the $n - th$ Gram point g_n is defined as the unique solution > 7 to $\theta(g_n) = n\pi$. The Gram points are as dense as the zeros of $\zeta(s)$ but are much more regularly distributed. Grams law is the empirical observation that $Z(t)$ usually changes its sign in each Gram interval $G_n = [g_n, g_{n+1})$. This law fails infinitely often, but it is true in a large proportion of cases. Eq. 4 implies that on average each Gram interval contains one zero. We can represent the zeros by a sequence $n_1n_2...n_k$ where n_i is the number of zeros in Gram interval i and k is the length of the sequence, i.e., the number of Gram intervals. Then from Eq. 4 the mean value of n_i over the sequence is 1, $\sum n_i = k$.

Odlyzko [7, 8] has made extensive numerical studies of the zeroes of the Riemann zeta function and their local spacings, and their relation to the random matrix models of physics. He confirmed numerically that the local spacings of the zeroes of the Riemann Zeta function obey the laws for the (scaled) spacings between the eigenvalues of a typical large unitary matrix. That is, they obey the laws of the Gaussian Unitary Ensemble (GUE) [9, 10, 11, 12]. Odlyzko's computations thus verified the discoveries and conjectures of Montgomery [13, 14, 15]. This has been extended to larger heights by Gourdon et al [16].

The author of this work studied the distributions of the zero spacings using Rescaled Range Analysis [17]. It has been shown that the long-range statistics of the zeroes of the Riemann zeta function are better described in terms of primes than by the GUE RMT. Berry [18, 19, 20, 21] has related this to a study of the semiclassical behaviour of classically chaotic physical systems. The primitive closed orbits of the physical system are analogous to the primes p.

The analogy comes from formulae that connect zeros of the zeta function and prime numbers [22, 23, 24] We collected statistics on a sample of 50000 zeros at $t = 10^{15}$. Table 1 shows the probability for a given Gram interval to contain a specified number of zeros.

The next section gives the calculation of the entropy for the sequence of zeros.

3 Entropy

Given a semi-infinite sequence of symbols like the sequence of zero counts in successive Gram intervals, one would like to measure the amount of structure present in the sequence. The concept of entropy provides such a measure. We will first review the general concept as presented by Shannon, and then apply it to the sequence of Riemann zeta zeros.

Following Shannon, consider a general source of information (like a source generating English sentences in a message). In general the sequences are not completely random. They have the statistical structure of the language. In English, the letter E occurs more frequently than Q, the sequence TH occurs more frequently than XP, etc. We can think of the information source as generating the message, emitting symbol after symbol. The source will choose successive symbols according to certain probabilities. The probabilities could depend on the preceding choices. Such a system is known as a stochastic process. The number of preceding symbols on which the probability of the next symbol depends gives the length to which the structure extends.

In the simplest case (albeit artificial) a choice depends only on the preceding symbol and not on the symbols before that. The statistical structure can then be described by a set of transition probabilities $p_i(j)$, the probability that symbol i is followed by the symbol j. The indices i and j range over all possible symbols. A second equivalent way of describing the structure is to give the "digram" probabilities $p(i, j)$, i.e., the relative frequency of the digram ij. More complex processes would involve trigram frequencies, tetragram frequencies, etc.

In our study, the symbols represent the number of zeros in a Gram interval, and the sequence is defined by specifying the number of zeros in successive Gram intervals. For example, if a given Gram interval contained 1 zero, and the next one contained no zeros, and the one after that contained 2 zeros, we represent the sequences as 102.

These processes can also be described as a state machine, with transition probabilities defined between the states. To make the state machine an information source we have to just specify the symbol that is emitted when a transition between states occurs. The states will correspond to the "residue of influence" from preceding letters. For example, we can represent the Riemann zeta zero sequence generating process as a state machine with three states, state A representing a sequence in which the number of zeros is exactly equal to the number of gram intervals (e.g., sequences like 1111 or 0121), state B representing a sequence in which the number of zeros is smaller than the number of gram

sequence length N	F_N	*Normalized* F_N	*excess zeros*						
			< -2	-2	-1	0	1	2	> 2
1	0.861	0.621	0	0	16189	67806	15822	183	0
2	0.773	0.558	0	417	17781	63698	17593	509	1
3	0.742	0.535	0	674	18568	61534	18529	693	0
4	0.725	0.523	1	882	18921	60372	18955	866	0
5	0.714	0.515	2	1086	18984	59780	19128	1015	1
6	0.706	0.509	2	1206	19178	59180	19266	1160	3
7	0.697	0.503	4	1377	19177	58742	19447	1246	1
8	0.686	0.495	5	1493	19158	58516	19489	1327	5
9	0.670	0.483	4	1577	19221	58213	19572	1402	3
10	0.646	0.466	5	1683	19101	58170	19587	1443	2
11	0.613	0.442	3	1792	19088	57964	19611	1530	2
12	0.568	0.410	9	1894	19078	57687	19759	1557	5
13	0.514	0.371	6	1989	18942	57700	19771	1577	3
14	0.453	0.327	7	2055	18907	57640	19724	1651	3
15	0.394	0.284	8	2108	18858	57615	19702	1691	4
16	0.338	0.244	13	2134	18894	57440	19809	1694	1
17	0.285	0.206	8	2242	18718	57587	19637	1787	5
18	0.236	0.170	10	2249	18805	57452	19605	1859	3
19	0.194	0.140	14	2256	18882	57287	19655	1884	4
20	0.158	0.114	8	2308	18801	57319	19655	1884	6
21	0.127	0.091	12	2269	18863	57399	19444	1987	6
22	0.100	0.072	6	2298	18896	57263	19529	1981	6
23	0.079	0.057	4	2331	18909	57144	19603	1983	4
24	0.061	0.044	12	2367	18823	57257	19453	2057	8
25	0.046	0.033	11	2436	18703	57319	19415	2085	7
26	0.034	0.025	13	2374	18857	57236	19355	2133	7
27	0.025	0.018	7	2349	18949	57135	19407	2121	6
28	0.018	0.013	7	2381	18885	57231	19277	2186	6
29	0.013	0.010	8	2378	18944	57065	19439	2130	8
30	0.010	0.007	11	2290	19027	57155	19342	2141	5
31	0.007	0.005	10	2295	19083	57029	19411	2135	7

Table 2: Approximations to H by considering sequences of N symbols.

intervals (e.g., sequences like 1011 or 0111), and state C representing a sequence in which the number of zeros is greater than the number of gram intervals (e.g., sequences like 1211 or 1121). It is known that the sequences of the Riemann zeta zeros are almost always in state A. Transitions from state A to state B, for example, would be accompanied by the emission of a 0 symbol, e.g. 1111 would transition to 11110 when the succeeding gram interval contains no zeros.

If P_i denotes the probability of being in a state i, and $p_i(j)$ denotes the probability of producing the next symbol j when in state i, then the entropy is defined as

$$H = \sum_i P_i p_i(j) \log(p_i(j)). \tag{7}$$

We make of the following theorem from Shannon to estimate the entropy.

Theorem 1. *Let $p(B_i, S_j)$ be the probability of sequence B_i followed by symbol S_j. Let $p_{B_i}(S_j) = p(B_i, S_j)/P(B_i)$ be the conditional probability of S_j after B_i. Let*

$$F_N = -\sum_{i,j} p(B_i, S_j) \log(p_{B_i}(S_j)). \tag{8}$$

where the sum is over all blocks B_i of $N-1$ symbols and over all symbols. Then F_N is a monotonic decreasing function of N, and $Lim_{N\to\infty} F_N = H$.

A series of approximations to H can be obtained by considering the statistical structure of the sequences extending over $1, 2, ..., N$ symbols. If there are no statistical influences extending over more than N symbols, then $F_N = H$. The ratio of the entropy of a source to the maximum it could have while still restricted to the same symbols is called its relative entropy.

For the sequence $n_1 n_2 ... n_k$ (where n_i is the number of zeros in Gram interval i and k is the length of the sequence, i.e., the number of Gram intervals), we have $\sum n_i = k$. We denote the excess zeros in a sequence of length k as $\sum n_i - k$. We collected statistics on a sample of 100000 zeros at $t = 10^{26}$ [25]. Table 2 shows the F_N for N from 1 to 31, as well as the distribution of the sequences classified by the excess zeros. We see that the entropy is very low, and the structure extends quite far out. Matiyasevich [26] has also found a remarkable ability to predict new zeros using the preceding zeros. This lends credence to the presence of high structure in the sequence of zeros.

4 Conclusions

We calculated the entropy of the sequence of Riemann zeta zeros. We find that the entropy is very low. This provides a quantitative measure of the large amount of structure present in the sequence of zeros. The presence of structure is encouraging for attempts to predict the position of the zeros using machine learning techniques.

References

[1] B. Riemann, "Über die Anzahl der Primzahlen uter Einer Gegebenen Gröbe," *Montasb. der Berliner Akad.*, (1858), 671-680.

[2] B. Riemann, "Gesammelte Werke", Teubner, Leipzig, (1892).

[3] E. Titchmarsh, "The Theory of the Riemann Zeta Function," Oxford University Press, Second Edition, (1986).

[4] H. M. Edwards, "Riemann's Zeta Function," Academic Press, (1974).

[5] C. E. Shannon, "A Mathematical Theory of Communication" *Bell Systems Technical Journal*, **27**, 279-423, (1948).

[6] O. Shanker, "Neural Network prediction of Riemann zeta zeros" *Advanced Modeling and Optimization*, **14**, 717-728, (2012).

[7] A. Odlyzko, "The 10^{20}-th Zero of the Riemann Zeta Function and 70 Million of its Neighbors," (preprint), A.T.T., (1989).

[8] A. Odlyzko, "Dynamical, Spectral, and Arithmetic Zeta Functions", Amer. Math. Soc., Contemporary Math. series, **290**, 139-144, (2001).

[9] E. Wigner, "Random Matrices in Physics," *Siam Review*, **9**, 1-23, (1967).

[10] M. Gaudin, M. Mehta, "On the Density of Eigenvalues of a Random Matrix," *Nucl. Phys.*, **18**, 420-427, (1960).

[11] M. Gaudin, "Sur la loi Limite de L'espacement de Valuers Propres D'une Matrics Aleatiore," *Nucl. Phys.*, **25**, 447-458, (1961).

[12] F. Dyson, "Statistical Theory of Energy Levels III," *J. Math. Phys.*, **3**, 166-175, (1962).

[13] H. Montgomery, "Topics in Multiplicative Number Theory," L.N.M., **227**, Springer, (1971).

[14] H. Montgomery, "The Pair Correlation of Zeroes of the Zeta Function," *Proc. Sym. Pure Math.*, **24**, *AMS*, 181-193, (1973).

[15] D. Goldston, H. Montgomery, "Pair Correlation of Zeros and Primes in Short Intervals," *Progress in Math.*, Vol. 70, Birkhauser, 183-203, (1987).

[16] Y. Saouter, X. Gourdon, P. Demichel, "AN IMPROVED LOWER BOUND FOR THE DE BRUIJN-NEWMAN CONSTANT" *MATHEMATICS OF COMPUTATION*, **80**, 22812287, (2011).

[17] O. Shanker, Generalised Zeta Functions and Self-Similarity of Zero Distributions, *J. Phys. A* **39**(2006), 13983-13997.

[18] M. V. Berry, "Semiclassical theory of spectral rigidity," *Proc. R. Soc.*, **A 400** , 229-251, (1985).

[19] M. V. Berry, "Riemann's zeta function: a model for quantum chaos?," *Quantum chaos and statistical nuclear physics (Springer Lecture Notes in Physics)*, **263** , 1-17, (1986).

[20] M. V. Berry, "Quantum Chaology," *Proc. R. Soc.* , **A 413** , 183-198, (1987).

[21] M. V. Berry, 'Number variance of the Riemann zeros'," *NonLinearity* , **1** ,399-407 , (1988).

[22] E. Landau, U:ber die Nullstellen der Zetafunktion, *Math. Ann. 71* (1911), 548-564.

[23] S. M. Gonek, A formula of Landau and mean values of $zeta(s)$, pp. 92-97 in *Topics in Analytic Number Theory*, S. W. Graham and J. D. Vaaler, eds., Univ. Texas Press, 1985.

[24] A. Odlyzko, "On the distribution of spacings between zeros of the zeta function", Math. Comp, **48**, 273-308, (1987).

[25] G. A. Hiary, "FAST METHODS TO COMPUTE THE RIEMANN ZETA FUNCTION", arxiv.org, math.NT, 0711.5005v4, (2011).

[26] Y. Matiyasevich, "Some non-standard methods to perform calculations with Riemanns zeta function', Globus seminar, Web report, http://www.mccme.ru/ium/globus.html, (2012).