Why moments? Moments of the ζ-function were introduced by Hardy and Littlewood [56] as an approach toward the Lindelöf Hypothesis 9.7.

RH implies LH, so it is natural to attack LH as an "easier" problem. LH is equivalent to

$$\int_0^T Z(t)^{2k}\,dt \ll_k T^{1+\varepsilon} \tag{13.3}$$

for all integer $k > 0$, which is why Hardy and Littlewood investigated moments.

Bounding such moments proved to be difficult: 100 years after Hardy and Littlewood we only know (13.3) for $k = 1$ and 2. But instead of becoming a dead-end in terms of proving LH, moments became a topic of independent interest. Other "families" of L-functions were introduced, with various types of moments. The focus shifted to finding precise formulas for the main terms in moments, not just on estimating the order of magnitude. Random matrix theory played a major role in understanding the leading order behavior and the structure of those moments, although ultimately number-theoretic heuristics provided the most precise conjectures. See [69][25][70][26][36].

Analysis of the argument. The claimed argument against RH arises from the error term in the $2k$th moment. Based on the conjectured best possible values $E_1 = \frac{1}{4}$ and $E_2 = \frac{1}{2}$ Ivić suggests [65][66] that the best possible error term in the $2k$th moment is $O(T^{k/4+\varepsilon})$. There is slightly more to Ivić's reasoning, such as the apparent inability of available techniques to provide bounds on high moments, but the main input is the conjectured values for E_1 and E_2.

If the error term (13.1) truly is of size $T^{k/4+\varepsilon}$, meaning that (13.1) is false with $k/4$ replaced by a smaller number, then (13.3) is also false. Therefore LH is false, so RH is false. That is the essence of that argument against RH.

Basing an argument on such limited data for small k is problematic, and the conjecture of $E_k = k/4$ is not based on any underlying principles. Suppose someone instead conjectured $E_k = \frac{3}{4} - 2^{-k}$ for $k \geq 1$. That also is a simple formula which fits the limited data, and it has the added benefit of implying LH. But without any good evidence, that conjecture should also not be taken seriously. Thus, this proposed reason to doubt RH is not based on a mathematical foundation, but on naive pattern matching.

Before the reader criticizes the author for being unkind, let us note that this author fully recognizes the folly of speculating on the size of error terms, and indeed is guilty of similarly unwise speculation. In a foundational paper on moments of L-functions [26] this author and coauthors conjecture that in great generality the error term in any conjectured higher moment is $E_k = \frac{1}{2}$. The conjecture of Ivić is as pessimistic as possible, and the conjecture in [26] is as naively optimistic as possible. Unfortunately, that optimistic error term is not always correct. Specifically, in some cases [35][36] there is a secondary main term of size $X^{\frac{3}{4}}$. It may even be that similar secondary main terms are common. The lesson is that speculating on the error term in moment conjectures is difficult, and it is folly to draw conclusions from such error term conjectures.

The structure of moments. Immediately after making a poorly founded conjecture about error terms, Ivić goes on [65][66] to make the perceptive speculation:

> the shape of the asymptotic formula...changes when $k = 4$.

That speculation has been borne out by subsequent work on the structure of moments.

The $2k$th moment (13.1) is only known for $k = 0$, 1, and 2. Conjectures for larger moments are a fairly recent development. One reason those moments are difficult is that the natural way to proceed is via an "approximate functional equation" for $Z(t)^k$, which is a sum of two Dirichlet polynomials of length $t^{k/2}$. (The Riemann-Siegel formula (4.1) can be viewed as an example: there are $\approx t^{\frac{1}{2}}$ terms in the sum.) When used to evaluate $\int_0^T Z(t)^{2k}\,dt$ the approximate functional equation will have length $T^{k/2}$. The difficulty is the lack of tools to handle Dirichlet polynomials where the length of the polynomial is greater than the length of the integral.

Conrey and Ghosh [29] developed a heuristic approach which gave a conjecture when $k = 3$, and Conrey and Gonek [31] pushed the method to its limit to make a conjecture when $k = 4$. That truly is the limit of the method, because when applied to the case $k = 5$ those methods produce an answer which is negative! Thus, Ivić was correct: some new phenomenon appears, beginning at $k = 4$. But that new phenomenon has more implications for that main term than for the error term.

13.2 Reason 5: The Deuring-Heilbronn phenomenon

The Riemann ζ-function is the simplest of an infinite family of functions known as "L-functions". The next simplest examples are the Dirichlet L-functions, $L(s, \chi)$, where χ is a primitive Dirichlet character. RH for all $L(s, \chi)$ is known as the **Generalized Riemann Hypothesis**.

Informally, by **L-function** we mean a Dirichlet series with a functional equation and an Euler product of a particular form. To give a more precise definition, there are two primary approaches. The *axiomatic approach* considers L-functions as analytic functions with certain properties. Within the axiomatic approach at one extreme we have the Selberg class [95][30], which has a minimal set of general axioms. At the other extreme are the tempered balanced analytic L-functions [45] which have a large number of precise axioms.

The *structural approach* describes L-functions as arising from arithmetic or automorphic objects. For the perspective in this paper, the important fact is that conjecturally all approaches describe the same set of functions, and *all such L-functions have an analogue of the Riemann Hypothesis*. The collection of all such Riemann Hypotheses is called the **grand Riemann Hypothesis**.

Under any of the above definitions, each L-function has a functional equation analogous to (3.3), a ξ_L-function analogous to (3.4), and a Z_L-function analogous to (3.6). Each Z_L-function is a real valued function which wiggles with the same type of randomness as $Z(t)$. It also has carrier waves as described in the earlier parts of this paper.

Linear combinations. The Deuring-Heilbronn phenomenon concerns functions which are not L-functions, but rather are linear combinations of L-functions. As we will see in Principle 13.2, a nontrivial linear combination of L-functions never satisfies RH, in the worst possible way.

Consider two L-functions which are being added, and focus on a small interval of the real line (say, where one expects to see a few dozen zeros of each Z-function). If one of the Z-functions happens to be much larger than the other over that entire interval, then the sum looks just like the larger Z-function. In particular, if the larger Z-function satisfies RH, then their sum satisfies RH on that interval. That is the typical situation and this was the motivation for developing the idea of carrier waves: the logarithm of each Z-function has a Gaussian distribution, and because of the carrier waves its size does not change too rapidly. Thus, 100% of the time one of them is much larger than the

other and stays larger over a span of many zeros. Under some mild additional hypotheses, that is how Bombieri and Hejhal [20] proved that if two L-functions have the same functional equation and each satisfies RH, then their sum has 100% of its zeros on the critical line.

But, 100% is not everything. As the two Z-functions take turns being the largest, occasionally they are about the same size over the span of a few zeros. When that happens, and we are adding two *different* functions that look like the top graph in Figure 3.1 and are similar in size, then not all of the zeros of the sum will be real. In other words:

Principle 13.1 *A nontrivial linear combination of L-functions will not satisfy the Riemann Hypothesis.*

Naturally occurring examples. The above discussion makes it seem obvious that a nontrivial linear combination of L-functions will never satisfy RH. But the situation was not always so clear. Let $\chi_{5,2}$ be the Dirichlet character mod 5 with $\chi(2) = i$, and define θ by $\tan\theta = (\sqrt{10 - 2\sqrt{5}} - 2)/(\sqrt{5} - 1)$. The **Deuring-Heilbronn function** is given by

$$F_{DH}(s) = \tfrac{1}{2}\sec\theta\left(e^{-i\theta}L(s,\chi_{5,2}) + e^{i\theta}L(s,\overline{\chi}_{5,2})\right) \tag{13.4}$$

which satisfies the functional equation

$$\xi_{DH}(s) := 5^{\frac{1}{2}s}\pi^{-\frac{1}{2}s}\Gamma(\tfrac{1}{2}s + \tfrac{1}{2})F_{DH}(s) = \xi_{DH}(1-s). \tag{13.5}$$

Except for the sign, that is the same functional equation satisfied by $L(s,\chi_{5,2})$. However $F_{DH}(s)$ does not satisfy RH: its first zero off the critical line is near $0.8085 + 85.699i$. Furthermore [101] it has infinitely many zeros in $\sigma > 1$, although it does not seem that an explicit example has ever been computed.

We have two Dirichlet series, $F_{DH}(s)$ and $L(s,\chi_{5,2})$, which satisfy similar functional equations. We know that one of them does not satisfy RH, so why should the other? That is the essence of the argument in Section 3 of [66]: why should we believe RH for one function when we know it is false for another, and those functions (quoting from [66]) "share many common properties"?

The Euler product. The standard answer to the above question (as noted in [66]) is that the functions are fundamentally different because $L(s,\chi_{5,2})$ is an L-function, but $F_{DH}(s)$ is not — because it does not have an Euler product. It is an exercise to check that every axiom in the Selberg class [95] is essential: omit any one of them and there will be functions which do not satisfy RH. The Euler product condition is not more important than the other axioms, but it is important, and without it one has no expectation of RH.

The Deuring-Heilbronn function is one of many examples constructed by taking linear combinations of L-functions. It is an exercise to show that a nontrivial linear combination of Euler products is not an Euler product. Thus, any such linear combination is not expected to satisfy RH, even if the constituent functions do satisfy RH. This explains why examples like $F_{DH}(s)$ do not cast doubt on RH. However, there is a bit more to be said if we consider the locations of the zeros of those functions.

Zeros outside the critical strip. The Deuring-Heilbronn function failed RH in the worst possible way: it has zeros in the region where the Dirichlet series converges absolutely. In fact it has $\gg T$ zeros in $\sigma > 1$ up to height T, but that is not any extra information because Dirichlet series are almost periodic functions (of t) in $\sigma > 1$: so once there is one zero in that region, there must

be $\gg T$ zeros in that region. Those zeros in $\sigma > 1$ are a general phenomenon, so we obtain a refinement of Principle 13.1:

Principle 13.2 *A nontrivial linear combination of L-functions will have infinitely many zeros in $\sigma > 1$.*

Principle 13.2 is a theorem for Dirichlet L-functions [90], for automorphic L-functions [22], and in an axiomatic setting [86]. Those results are interesting because one has a vector space of Dirichlet series, all of which satisfy the same functional equation. Conjecturally only the functions in the span of individual basis elements (i.e. the actual L-functions) satisfy RH, yet every other element of the space fails RH in the worst possible way. This highlights the key role of the Euler product.

Why exactly is the Euler product so important? Maybe it isn't. In terms of the analytic properties of the L-function, the most obvious consequence is that the Euler product prevents the L-function from having any zeros in $\sigma > 1$. So, is it the Euler product that matters, or the lack of zeros in $\sigma > 1$?

14 Discussion of the Mistaken Notions

We revisit the topics in Section 4 in the context of the Principles. Specifically, we re-examine the mistaken notions that the largest values of the ζ-function arise from particularly large zero gaps, that those large values occur when many of the initial terms in the Riemann-Siegel formula have the same sign, that large values are places to look for possible counterexamples to RH, and that the Gram points are special.

14.1 Large values and large gaps

Mistaken Notion 4.3, equating the largest values of the ζ-function with the largest gaps, is the core idea for three of the Notions. As has already been adequately addressed by the discussion of carrier waves, see Principle 5.2, that Notion does not accurately describe the typical large values, nor the truly largest values.

However, Mistaken Notion 4.3 *does* describe the behavior in the range of current computations, and sometimes a wrong idea can lead to the right answer. Kotnik [67] sought computational evidence that the known Ω-result (3.14) was not the true order of growth. The method was to find the sequence of largest values of $|\zeta(\frac{1}{2} + it)|$ for small t. Those computations did make it appear plausible that the true order of growth is faster than the Ω-result. Since those large values arise from large gaps, the computations were not revealing the true order of growth. But by Principle 9.2, the maxima in the large gaps should be $\gg e^c\sqrt{\log t}$, which is larger than the current best Ω-result.

14.2 Finding large values

Mistaken Notion 4.4 asserts that the largest values of the ζ-function occur when many initial terms in the Riemann-Siegel formula have the same sign. The Notion makes sense: the first few terms are the largest, so if you want the entire sum to be large, it is efficient to focus on the initial terms. That argument ignores the fact that there are $\gg t^{\frac{1}{2}}$ terms of size $\gg t^{-\frac{1}{4}}$. Even a small bias in such a large number of terms could contribute much more than the initial terms. Indeed, one of the arguments for the conjectured maximum size of the

ζ-function is based on a random model for the tail of that sum, which is larger than can be obtained from a strong bias in the initial terms.

Nevertheless, Mistaken Notion 4.4 is how the largest known values of the ζ-function have been found [17][100]. It is possible to take that idea too far. Consider

$$F(t, g(t)) = 2 \sum_{n<g(t)} n^{-\frac{1}{2}} \cos(\theta(t) - t \log n). \tag{14.1}$$

The Riemann-Siegel formula has $g(t) = \sqrt{t/2\pi}$. One might expect that $F(t, t^\varepsilon)$ is a good approximation for the large values of $Z(t)$ because (assuming the Lindelöf Hypothesis) a similarly short sum is a good approximation to $\zeta(s)$ in $\sigma > \frac{1}{2}$. Kotnik [67] considered $F(t, t^\delta)$ for $\delta = \frac{1}{3}$ and $\delta = \frac{1}{4}$. Those functions are faster to evaluate than the Riemann-Siegel formula, and this sped up the computations by more quickly locating regions where $Z(t)$ was more likely to be large.

Tihanyi [99] took it a step too far by considering $F(t, \log t/2\pi)$. Since $F(t, g(t)) \ll \sqrt{g(t)}$, Tihanyi's approximation has no hope of capturing the largest values. Yet, using $F(t, \log t/2\pi)$ as a filter (the same strategy as Kotnik [67]) turned out to be useful for finding many points where $Z(t)$ was larger than 1000, and in subsequent work [100], found a value of $Z(t)$ even larger than that found by Bober and Hiary [17]. So, a Notion may be wrong or misleading, yet still serve as the seed for a new discovery.

Although it is out of reach to compute in regions where the carrier wave, and not the local zero gaps, provide the main contribution to the size of $Z(t)$, it would still be of theoretical interest to find heuristics for locating truly large values.

14.3 Trying to refute RH

If RH were not true, how would one go about searching for a pair of zeros off the line? Mistaken Notion 4.5 suggests that a large gap might "push some zeros off the line". That is contrary to the idea behind Principle 12.1: random unitary polynomials can have large zero gaps — conjecturally with the same size and frequency as the ζ-function — which influence the nearby gaps but do not destroy the fact that the polynomial is unitary.

We offer two more arguments against Mistaken Notion 4.5. The first is similar to an idea used in the discussion of carrier waves in Section 5: run the argument in reverse. Suppose there is a pair of zeros off the line. Would one expect there to be a particularly large gap nearby? Surely not. That pair off the line, if they were not too far from the line, would only cause a slightly larger than average gap in zeros on the critical line, and furthermore the function would be small (not large) within that zero gap.

The second argument concerns the fact that no plausible suggestion has been made for how RH would fail if it were false. Maybe the zeros in $\sigma > \frac{1}{2}$, if they exist, all lie on $\sigma = \frac{3}{4}$? Is that a particularly absurd suggestion? Zeros that far from the line have a very small influence on the behavior of $Z(t)$ — in particular there is no associated large value or large zero gap. It is perfectly reasonable to seek fame and fortune by trying to computationally disprove RH. But searching for zeros off the line should be accompanied by, and tailored to, a model for how the failure will occur. A scenario in which RH fails, and those failures are typically associated to a large zero gap, is arguably even less plausible than having all non-critical zeros on $\sigma = \frac{3}{4}$.

The carrier wave frequently causes the Z-function to stay very small over a wide span of zeros. In that region a slight perturbation would cause many zeros to fall off the critical line. Maybe attempts to computationally disprove RH

should begin by looking for regions where the initial terms in the Riemann-Siegel formula mostly cancel to nothing?

14.4 Gram's law, as it was originally formulated

Mistaken Notion 4.6 contains two statements: that Gram points are special, and that Gram's law is helpful for understanding properties of the zeros. First we address Gram's law. For reasons why Gram points are not special, see Subsection 14.5.

Gram's law, as usually stated, misses a key point of Gram's original observation. What Gram noted is that the zeros and the Gram points "separate each other". In other words, in the range computed by Gram not only does the nth Gram interval (g_n, g_{n+1}) contain exactly one zero, that zero is the $n+2$nd zero. (The +2 arises because the definition of Gram point has $g_0 \approx 17.84$.) We propose:

Reformulation 14.1 The Precise Gram's Law. *The* ***precise Gram's law holds*** *for the* ***Gram interval*** (g_m, g_{m+1}) *if that interval contains the* $m+2$*nd zero of the* Z*-function and no other zeros.*

That version of Gram's law also follows from Principle 5.8.

Gram's observation was based on the first 15 zeros, but note that he suggested the interlacing would not continue to hold indefinitely. In fact, Gram's law (either version) fails after the 126th zero, $\gamma_{126} \approx 279.2$, because there is a large zero gap and the 127th zero falls in the 126th Gram interval. The Gram interval (g_{3357}, g_{3358}) contains exactly one zero, but the precise Gram's law fails because the previous interval contains two zeros. That is the first time there is a difference between the two versions of the law. There are 16 differences among the first $10,000$ zeros.

That the precise Gram's law holds initially can be seen either as a consequence of $S(t)$ being small, or that the first term in the Riemann-Siegel formula is providing a good approximation for the zeros. A similarity between those two reasons is that the sign and the Γ factor in the functional equation give rise both to the first term in the Riemann-Siegel formula, and also to the main part of the counting function of the zeros. The Dirichlet series gives rise to the other terms, and also to $S(t)$.

However, we suggest that $S(t)$ provides a more fundamental insight into the behavior of the ζ-function and its zeros. A large height, the Riemann-Siegel formula needs more terms, but it is not immediately obvious that this causes the precise Gram's law to fail most of the time. On the other hand, at large height $|S(t)|$ typically is big, which implies that the precise Gram's law holds rarely:

Observation 14.2 A necessary (but not sufficient) condition for the precise Gram's law to hold on a Gram interval (g_n, g_{n+1}) is that $|S^+(g_n)| \leq 1$.

The frequency of that occurrence at height t can be estimated using Selberg's central limit theorem. It is clear that in the limit it occurs 0% of the time. The slow growth of $S(t)$ makes this effect difficult to observe: even for characteristic polynomials of random matrices in $U(1000)$, as in the examples in Section 8, the precise Gram's law still holds more than 40% of the time.

The fact that Gram's law (the non-precise version) holds around 66% of the time should be viewed as an aspect of the random matrix model of zeros of the ζ-function: the normalized nearest neighbor spacing of zeros/eigenvalues, see the leftmost plot in Figure 8.9, is concentrated around 1. So it is reasonably likely that *any* interval of normalized length 1 contains exactly one zero/eigenvalue

(although at low height a Gram interval is more likely than a random interval, see [55] or apply Principle 5.8).

The precise formulation of Gram's law describes a rigidity which only holds initially.

Principle 14.3 The sloshing water model. *At larger heights the zeros* ***slosh back-and-forth****. That term is meant to evoke what happens when one tries to walk while carrying a wide but shallow pan of water. Large groups of zeros are shifted to the left or the right of their expected location, even though locally all sets of zeros behave similarly.*

Because of the sloshing, $S(t)$ is large (in absolute value) most of the time. Long runs of zeros miss their intended Gram intervals: many in a row falling to the right, and later, many in a row falling to the left. (But locally the zeros are still obeying random matrix statistics, so a Gram interval (or any other interval of that length) will contain exactly one zero more than 66% of the time.) The Z-function is wiggling with great amplitude for a while, and then wiggling with small amplitude. The size of the carrier wave is independent of whether the nearby zeros arrive before or after their Gram interval, that is, independent of the size or the sign of $S(t)$. All of these phenomena are manifestations of the density wave, a wave which cannot exist if the precise Gram's law were true too often.

Problem 14.4 Develop a random model for the density wave, equivalently the carrier wave, preferably as a theorem in the case of the CUE. □

A good solution to that problem will allow generating examples from systems far larger than possible if one must first produce a large random matrix and then find its eigenvalues.

A solution to that problem must grapple with the issue that the definition of "density wave" depends on the range over which the density is being measured. A solution to that problem, along the lines of directly creating a density wave as in Figure 8.5, could be a step toward the next level of understanding the behavior of the ζ-function a large height. By that we mean: in the first level we recognize that the local distribution of zeros is random; that randomness is now understood. This is occurs in both the L-function and the characteristic polynomial worlds. The next level is carrier/density waves. This is beyond the reach of the L-function world, and observable but not understood in the random matrix world. Once carrier waves are understood, as in Problem 14.4, then it is natural to ask whether those waves are sitting in a yet larger structure. The analogy looks like

$$\text{local zero spacing} : \text{carrier wave} \ :: \ \text{carrier wave} : ?????. \tag{14.2}$$

If one could directly produce density/carrier waves, then perhaps those can be generated without needing to directly produce a large random matrix. This may exhibit interesting (self similar?) structures, as mentioned in Subsection 11.4.

14.5 Gram points are of transient interest

A recent paper of Shanker [96] used computational data for the ζ-function to make two conjectures about Gram intervals, and to suggest that "Gram points have interesting properties which distinguish them from random points on the critical line". We examine that work in the context of the Principles, reaching a different conclusion.

The conjectures in [96] are:

1. If $\{g_n\}$ are the Gram points, then the distribution of $Z(g_{2n})$ is the negative of the distribution of $Z(g_{2n+1})$.

2. The properties of the Z-function on a set of consecutive Gram points is the same if that set is reversed.

   For example, the probability that the number of zeros between consecutive Gram points is $3, 1, 4, 1, 5, 9$ is the same as the probability that the number of zeros is $9, 5, 1, 4, 1, 3$.

We will see that both conjectures are true for random unitary matrices, therefore both conjectures are immediate consequences of the Keating-Snaith Law.

The second part of [96] uses two sets of data to suggest that Gram points are special. Say that a Gram point g_n is **good** if $(-1)^n Z(g_n) > 0$ and **bad** otherwise. Table 14.5 contains data from Table 2 in [96].

Table 14.5 The proportion of good and bad Gram intervals containing $0, 1, 2$, or 3 zeros, and the combined proportion ignoring whether the interval is good or bad. Data from 10^7 zeros near $t = 10^{28}$. The bottom row is the limiting prediction from the GUE Hypothesis (Principle 7.3).

	$m = 0$	1	2	3	**size**
good	0.1097	0.7829	0.1071	0.00016	7373998
bad	0.3083	0.3838	0.3006	0.00709	2626002
all	0.1619	0.6781	0.1579	0.00198	10000000
GUE	0.1702	0.6614	0.1664	0.00186	

The second set of data concerns consecutive Gram intervals which contain either $2, 1, \dots, 1, 0$ or $0, 1, \dots, 1, 2$ zeros. Those **Gram blocks** are called **Type I** and **Type II**, respectively. As mentioned above, Type I and Type II, of the same length, occur with equal probability. However, what is considered in [96] are intervals which are shifted from Gram intervals by $k\delta$ where δ is the length of a Gram interval and $k \in \{-0.2, -0.1, 0, 0.1, 0.2\}$. Table 14.6 contains part of Table 5 from [96].

Table 14.6 The ratio Type II/Type I intervals for displaced Gram points, for 10^7 zeros near $t = 10^{28}$.

Length	-0.2δ	-0.1δ	0	0.1δ	0.2δ
2	2.268	1.504	0.999	0.663	0.441
3	3.624	1.896	0.998	0.526	0.274
4	5.588	2.367	1.001	0.426	0.178
5	8.849	2.923	1.011	0.343	0.115
6	14.373	3.728	1.008	0.266	0.070

The third conjecture in [96] is that in each row of Table 14.6, the product of the entries for $k\delta$ and $-k\delta$ equals 1. For characteristic polynomials, that conjecture is immediate from the formula for Haar measure; see below for a more complete explanation.

The numbers Table 14.5 were used to suggest that, while the behavior of all Gram points seems close to the limiting GUE prediction, the difference between the good and bad intervals casts doubt on the random matrix prediction. The numbers in Table 14.6 were used to suggest that the Gram points are special, distinguishing them from typical points having the same spacing. We will reach different conclusions by invoking the Keating-Snaith Law and performing similar experiments with random unitary polynomials.

Gram point for unitary polynomials. Recall that Gram points for the ζ-function are the extrema of the first term in the Riemann-Siegel formula (4.1).

As described in [55] we define the **Gram points** for a unitary polynomial in an analogous way. Recall that $\mathcal{Z}_A(\theta)$, defined in (7.15), is real for $\theta \in \mathbb{R}$. Therefore there exists θ_A such that

$$\mathcal{Z}_A(\theta) = 2\cos\Big(\frac{N}{2}(\theta - \theta_A)\Big) + \sum_{0 \le n < N/2} a_n \cos(n(\theta - \theta_A)), \tag{14.3}$$

with $a_n \in \mathbb{R}$ and where for convenience we have assumed N is even. The leading term is analogous to the leading term $2\cos(\theta(t))$ in the Riemann-Siegel formula, so we define the **Gram points** for the matrix A to be the points where the leading term in (14.3) has a maximum or minimum: $(\theta_A + 2\pi n/N)_{n=0,1,2,\dots}$. Note that if $\det A = 1$ then $\theta_A = 0$. See [55] for a more detailed discussion.

Now it is straightforward to explain the first two conjectures in [96]. The first conjecture follows from Principle 3.2. The first two conjectures are immediate from the Keating-Snaith Law and Lemma 7.9, using the above definition of Gram point.

The third conjecture, concerning the statistics for a sequence of Gram intervals shifted by $k\delta$ and $-k\delta$ follows from Lemma 7.9 because any statistic for a sequence of Gram intervals shifted by $k\delta$ is formally the same as the reverse sequence shifted by $-k\delta$.

We now address the larger issue of interpreting the data in Table 14.5 and Table 14.6.

By the Keating-Snaith Law, we match the size of the matrices using the rule $N = \log(t/2\pi)$. In the context of the data from [96], this means that the appropriate matrix size is $N = 62$. Table 14.7 and Table 14.8 show the results from 10^7 random eigenvalues from matrices in $U(N)$ for $N = 62$, 250, and 1000, analogous to Table 14.5 and Table 14.6, respectively

Table 14.7 The proportion of good and bad Gram intervals containing $0, 1, 2$, or 3 zeros, and the combined proportion ignoring whether the interval is good or bad. Data from 10^7 eigenvalues of Haar-random matrices in $U(N)$ for $N = 62$, 250, and 1000.

N		$m = 0$	1	2	3	**size**
62	good	0.1111	0.7802	0.1083	0.00021	7295392
	bad	0.2996	0.4009	0.2924	0.00687	2704608
	all	0.1621	0.6776	0.1581	0.00201	10000000
250	good	0.1269	0.7484	0.1241	0.00050	6621923
	bad	0.2488	0.5031	0.2433	0.00461	3378077
	all	0.1681	0.6655	0.1643	0.00189	10000000
1000	good	0.1377	0.7262	0.1351	0.00081	6147041
	bad	0.2205	0.5610	0.2148	0.00356	3852959
	all	0.1696	0.6625	0.1658	0.00187	10000000

Table 14.8 The ratio Type II/Type I intervals for displaced Gram points, among 10^7 eigenvalues of Haar-random matrices in $U(N)$ for $N = 62$, 250, and 1000.

N	**Length**	-0.2δ	-0.1δ	0	0.1δ	0.2δ
62	2	2.245	1.498	1.001	0.670	0.448
	3	3.499	1.876	0.998	0.534	0.288
	4	5.048	2.257	1.003	0.444	0.198
	5	7.372	2.732	1.011	0.372	0.138
	6	9.756	3.155	0.994	0.318	0.100
250	2	1.730	1.322	1.002	0.762	0.579
	3	2.216	1.500	1.005	0.667	0.454
	4	2.673	1.638	0.998	0.611	0.373
	5	3.221	1.812	1.011	0.560	0.318
	6	3.683	1.987	1.018	0.524	0.273
1000	2	1.484	1.218	0.995	0.816	0.672
	3	1.750	1.337	1.000	0.749	0.570
	4	2.013	1.432	1.993	0.705	0.502
	5	2.148	1.447	0.982	0.664	0.451
	6	2.424	1.566	1.024	0.661	0.436

The data are consistent, therefore Gram points are not important. Table 14.5 and Table 14.7 both show that the good and bad gram points behave differently, and the proportions in each table are similar. However, this is not evidence against the prediction that the distinction between good and bad gram points is meaningless in the long run. Indeed, a matrix size of $N = 62$ is small, particularly when some of the limiting behavior reveals itself on the scale of $\sqrt{\log N}$. The $N = 250$ and 1000 sections of Table 14.7 provide a nice illustration of the fact that, as $N \to \infty$, the good and bad intervals become equally likely. As described in Subsection 14.4, at large height most zeros have slid far away from their expected location, so it is meaningless to try to separate "good" from "bad". Yet the local spacing between zeros still looks the same everywhere, and that is the real reason that Grams law (but not the precise Grams law) continues to hold more than 66% of the time.

Table 14.8 shows that at $N = 62$, shifting the Gram points has a significant effect. But 62 is a small number, and the $N = 250$ and 1000 cases make it less surprising that as $N \to \infty$ every entry in that table will approach 1.

The data differ significantly. We have argued that the numbers in Table 14.5 and the $N = 62$ portion of Table 14.7 are similar, but in fact they differ significantly. The sample sizes are 10^7, so those numbers differ by more than 10 standard deviations. This is a consequence of Principle 7.7. The lower order terms mentioned in that principle can be (statistically) significant, particularly when we are sampling at a small matrix size such as $N = 62$. Note that the differences are in the direction indicated by that principle: for the random matrices larger gaps and clusters of zeros are more likely.

The differences between Table 14.6 and the $N = 62$ portion of Table 14.8 are even more pronounced. The differences are greatest for the unlikely events: there are 10^7 data points, but the lowest row of each table corresponding to length = 6 only occurs a few thousand times in the sample. So, those apparent discrepancies are also explained by the principle that for random matrices there are secondary terms which make extreme events more likely.

15 The chicken or the egg?

Which came first: the ζ-function or its zeros? That sounds like a meaningless question, but we suggest that the answer factors into whether or not one is skeptical of RH.

Suppose you believe the function is what matters, either the ζ-function given by its Dirichlet series or some other expression, or the Z-function given by the Riemann-Siegel formula. From that perspective, the zeros are determined by the function, and it would take an unlikely conspiracy for RH to be true. It might still bother you that RH is "barely true". The suggestion at the very end of Subsection 14.3 might be viewed as an intriguing idea worth exploring. A failure of RH could occur for accidental reasons.

Now suppose you believe it is the zeros which matter. The zeros determine the function, with the density wave setting the scale and the local arrangement providing lower-order adjustments. Every zero is important and meaningful: a purported list of zeros containing one tiny error would be detected (by Weil's explicit formula or some other means). If there were zeros off the line, they would be there to serve a specific purpose and their location would be meaningful. The suggestion at the end of Subsection 14.3 is just silly: the Z-function stays small over a wide region only because the density of zeros is higher.

It is hoped that this paper helps cultivate an appreciation for the explanatory power which comes from viewing the zeros as the basic object.

Bibliography

[**1**] Louis-Pierre Arguin, Paul Bourgade, and Maksym Radziwiłł, *The Fyodorov-Hiary-Keating Conjecture. I*, arXiv: 2007.00988; (2020)

[**2**] Louis-Pierre Arguin, Paul Bourgade, and Maksym Radziwiłł, *The Fyodorov-Hiary-Keating Conjecture. II*, arXiv: 2307.00982; (2023)

[**3**] Louis-Pierre Arguin, Guillaume Dubach, and Lisa Hartung, *Maxima of a Random Model of the Riemann Zeta Function over Intervals of Varying Length*, Ann. Inst. Henri Poincaré Probab. Stat. no. 1, **60** (2024) pp. 174-200. arXiv: 2103.04817;

[**4**] F. V. Atkinson, *The mean value of the zeta-function on the critical line*, Proc. London Math. Soc. no. 2, **47** (1941) pp. 174-200.

[**5**] R. Backlund, *Über die Beziehung zwischen Anwachsen und Nullstellen der Zetafunktion*, Öfversigt Finska Vetensk. Soc. no. 9, **61** (1918-19)

[**6**] E.C. Bailey and J.P. Keating, *Maxima of log-correlated fields: some recent developments*, J. Phys. A: Math. Theor. **55** no. 5, arXiv: 2106:15141; (2022)

[**7**] Yacine Barhoumi-Andréani, Christopher Hughes, Joseph Najnudel, and Ashkan Nikeghbali, *On the number of zeros of linear combinations of independent characteristic polynomials of random unitary matrices*, Int. Math. Res. Not. IMRN no. 23, (2015) pp. 12366-12404.

[**8**] Gérard Ben Arous and Paul Bourgade, *Extreme gaps between eigenvalues of random matrices*, Ann. Probab. **41** no. 4, (2013) pp. 2648-2681.

[**9**] M. V. Berry, *Universal oscillations of high derivatives*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **461** no. 2058, (2005) pp. 1735-1751.

[**10**] M. V. Berry, *Superoscillations and leaky spectra*, J. Phys. A: Math. Theor.

52 (2019) pp. 1-11.

[**11**] M. V. Berry, *Riemann zeros in radiation patterns: II. Fourier transforms of zeta*, J. Phys. A: Math. Theor. **48** (2015)

[**12**] M. V. Berry, *Semiclassical formula for the number variance of the Riemann zeros*, Nonlinearity **1** (1988) pp. 399-407.

[**13**] M. V. Berry and J. P. Keating, *The Riemann zeros and eigenvalue asymptotics*, SIAM Review **41** (1999) pp. 236-266.

[**14**] Philippe Blanc, *A new reason for doubting the Riemann Hypothesis*, Exp. Math. **31** no. 1, (2022) pp. 88-92. 10.1080/10586458.2019.1586600

[**15**] Philippe Blanc, *An unexpected property of odd order derivatives of Hardy's function*, Publ. Inst. Math. (Beograd) (N.S.) **95** no. 109, (2014) pp. 173-188.

[**16**] Philippe Blanc, *Optimal upper bound for the maximum of the k-th derivative of Hardy's function*, JNT **154** (2015) pp. 105-117.

[**17**] Jonathan Bober and Ghaith Hiary, *New computations of the Riemann zeta function on the critical line*, Exp. Math. **27** no. 2, (2018) pp. 125-137.

[**18**] Andriy Bondarenko and Kristian Seip, *Extreme values of the Riemann zeta function and its argument*, Math. Ann. **372** no. 3-4, (2018) pp. 999-1015.

[**19**] E. Bombieri, *Problems of the Millennium: the Riemann Hypothesis*, https://www.claymath.org/wp-content/uploads/2022/05/riemann.pdf;

[**20**] E. Bombieri and D. Hejhal, *On the distribution of zeros of linear combinations of Euler products*, Duke Math. J. **80** no. 3, (1995) pp. 821-862.

[**21**] Andrew Booker, *Artin's Conjecture, Turing's Method, and the Riemann Hypothesis*, Experiment. Math. **15** no. 4, (2006) pp. 385-408.

[**22**] Andrew Booker and Frank Thorne, *Zeros of L-functions outside the critical strip*, Algebra Number Theory **8** no. 9, (2014) pp. 2027-2042.

[**23**] Richard P. Brent, *On the zeros of the Riemann zeta function in the critical strip*, Math. Comp. **33** no. 148, (1979) pp. 1361-1372.

[**24**] J. B. Conrey, *More than two fifths of the zeros of the Riemann zeta function are on the critical line*, J. Reine Angew. Math. **399** (1986) pp. 1-26.

[**25**] J. B. Conrey and D. W. Farmer, *Mean values of L-functions and symmetry*, Internat. Math. Res. Notices (2000) **17** pp. 883-908.

[**26**] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith, *Autocorrelation of random matrix polynomials*, Commun. Math. Phys **237** (2003) no. 3, pp. 365-395.

[**27**] J. B. Conrey, D. W. Farmer, and M. R. Zirnbauer, *Autocorrelation of ratios of L-functions*, Commun. Number Theory Phys **2** (2008) no. 3, pp. 593-636.

[**28**] J. B. Conrey and A. Ghosh, *Mean values of the Riemann zeta-function*, Mathematika **31** (1984) pp. 159-161.

[**29**] J. B. Conrey and A. Ghosh, *A conjecture for the sixth power moment of the Riemann zeta-function*, Int. Math. Res. Not. **15** (1998) pp. 775-780.

[**30**] J. B. Conrey and A. Ghosh, *On the Selberg class of Dirichlet series: small degrees*, Duke Math. J. **72** (1993) pp. 673-693.

[**31**] J. B. Conrey and S. M. Gonek, *High moments of the Riemann zeta-*

function, Duke Math. Jour. (2001) **107** pp. 577-604.

[32] J. B. Conrey and H. Iwaniec, *Spacing of zeros of Hecke L-functions and the class number problem*, Acta Arith. (2002) **103** no. 3, pp. 259-312.

[33] J. B. Conrey and N. C. Snaith, *Applications of the L-functions ratios conjectures*, Proc. Lond. Math. Soc. (2007) **94** no. 3, pp. 594-646.

[34] P. Diaconis and M. Shahshahani, *On the eigenvalues of random matrices*, J. Appl. Probab. **31A** (1994) pp. 49-62.

[35] A. Diaconu, *On the third moment of $L(\frac{1}{2}, \chi_d)$ I: The rational function field case*, J. Number Theory. **198** (2019) pp. 1-42.

[36] A. Diaconu, D. Goldfeld, and J. Hoffstein, *Multiple Dirichlet series and moments of zeta- and L-functions*, Compositio Math. **139** (2003) no. 3, pp. 297-360.

[37] Alexander Dobner, *A proof of Newman's conjecture for the extended Selberg class*, Acta Arith. **201** no. 1, pp. 29-62. (2021)

[38] Freeman Dyson, *A Brownian-motion model for the eigenvalues of a random matrix*, J. Mathematical Phys. **3** pp. 1191-1198. (1962)

[39] Ioana Dumitriu and Alan Edelman, *Matrix models for beta ensembles*, J. Math. Phys. **43** no. 11, pp. 5830-5847. (2002)

[40] David W. Farmer, *Long mollifiers of the Riemann zeta-function*, Mathematika **40** no. 1, pp. 71-87. (1993)

[41] David W. Farmer, *Jensen polynomials are not a plausible route to proving the Riemann Hypothesis*, Adv. Math. **411** (2022) 14 pages. Paper No. 108781 https://arxiv.org/abs/2008.07206;

[42] David W. Farmer, *When are the zeros of a polynomial distinct and real: a graphical view*, Preprint (2021) https://arxiv.org/abs/2010.15608;

[43] D.W. Farmer, S.M. Gonek, and C.P. Hughes, *The maximum size of L-functions*, J. Reine Angew. Math. **609** (2007) pp. 215-236.

[44] David W. Farmer, Francesco Mezzadri, and Nina Snaith, *Random polynomials, random matrices and L-functions*, Nonlinearity **19** (2006) no. 4, pp. 919-936.

[45] David W. Farmer, Ameya Pitale, Nathan Ryan, and Ralf Schmidt, *Analytic L-functions: definitions, theorems, and connections*, Bull. Amer. Math. Soc. (N.S.) **56** (2019) no. 2, pp. 261-280.

[46] David W. Farmer and Robert Rhoades, *Differentiation evens out zero spacings*, Trans. Amer. Math. Soc. **357** (2005) no. 9, pp. 3789-3811.

[47] David W. Farmer, Sally Koutsoliotas, Stefan Lemurell, and David P. Roberts, *Patterns in some landscapes of L-functions*, preprint;

[48] Shaoji Feng, *Zeros of the Riemann zeta function on the critical line*, J. Number Theory **132** (2012) no. 4, pp. 511-542.

[49] Yan Fyodorov and Jonathan Keating, *Freezing transition and extreme values: random matrix theory, $\zeta(\frac{1}{2} + it)$ and disordered landscapes*, Phil. Trans. R. Soc. A **372** (2014) no. 20120503, 32.

[50] Yan Fyodorov, Sven Gnutzmann, and Jonathan Keating, *Extreme values of CUE characteristic polynomials: a numerical study*, J. Phys. A **51** (2018) no. 46, 22.

[51] Yan Fyodorov, Ghaith Hiary, and Jonathan Keating, *Freezing Transition, Characteristic Polynomials of Random Matrices, and the Riemann Zeta-*

Function, Phys. Rev. Lett. **108** (2012) no. 170601,

[52] Gonek, S. M.; Hughes, C. P.; Keating, J. P., *A hybrid Euler-Hadamard product for the Riemann zeta function*, Duke Math. J. **136** no. 3, (2007) pp. 507-549.

[53] Xavier Gourdon, *Computation of zeros of the Zeta function*, preprint; http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeroscompute.html (2004)

[54] J. Gram, *Sur les zéros de la fonction $\zeta(s)$ de Riemann*, Acta Math **27** (1903) pp. 289-304.

[55] Cătălin Hanga and Christopher Hughes, *Probabilistic models for Gram's Law*, preprint; arXiv: 1911.03190; (2020)

[56] G. H. Hardy and J. E. Littlewood, *Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes*, (1918) Acta Mathematica pp. 119-196. **41**

[57] A. Harper, *On the partition function of the Riemann zeta function, and the Fyodorov--Hiary--Keating conjecture*, preprint; arXiv: 1906.05783; (2019)

[58] Dennis Hejhal, *On the distribution of zeros of a certain class of Dirichlet series*, Internat. Math. Res. Notices no. 4, (1992) pp. 83-91.

[59] Ghaith Hiary, *Personal website*, https://people.math.osu.edu/hiary.1/fastmethods.html (Accessed March 3, 2022)

[60] Ghaith Hiary and Andrew Odlyzko, *The zeta function on the critical line: numerical evidence for moments and random matrix theory models*, Math. Comp. **81** no. 279, (2012) pp. 1723-1752.

[61] A. Huckleberry, A. Püttmann, and M. R. Zirnbauer, *Haar expectations of ratios of random characteristic polynomials*, Complex Anal. Synerg. no. 1, (2016) **2** Paper No. 1, 73 pp..

[62] C.P. Hughes, J.P. Keating, and Neil O'Connell, *Random matrix theory and the derivative of the Riemann zeta function*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. no. 2003, **456** (2000) pp. 2611-2627.

[63] J. I. Hutchinson, *On the roots of the Riemann zeta-function*, Trans. Amer. Math. Soc. no. 27, (1925) pp. 49-60.

[64] A. E. Ingham, *Mean-value theorems in the theory of the Riemann zeta-function*, Proceedings of the London Mathematical Society no. 92, (1926) **27** pp. 273-300.

[65] A. Ivić, *On some results concerning the Riemann Hypothesis*, *Analytic Number Theory* LMS LNS no. 247, Cambridge University Press; Y. Motohashi, (1997) pp. 139-167.

[66] A. Ivić, *On some reasons for doubting the Riemann Hypothesis*, preprint; arXiv: 0311162; (2003)

[67] Tadej Kotnik, *Computational estimation of the order of $\zeta(1/2+it)$*, Math. Comp. **73** no. 246, (2004) pp. 949-956.

[68] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, AMS Colloquium Publications **45** AMS, Providence, RI; (1999)

[69] J. P. Keating and N. C. Snaith, *Random matrix theory and $\zeta(\frac{1}{2}+it)$*, Comm. Math. Phys. **214** (2000) pp. 57-89.

[**70**] J. P. Keating and N. C. Snaith, *Random matrix theory and L-functions at* $s = \frac{1}{2}$, Comm. Math. Phys. **214** (2000) pp. 91-110.

[**71**] Haseo Ki, *The Riemann* Xi*-function under repeated differentiation*, J. Number Theory **120** no. 1, (2006) pp. 120-131.

[**72**] Rowan Killip and Irina Nenciu, *Matrix models for circular ensembles.*, Int. Math. Res. Not. no. 50, (2004) pp. 2665-2701.

[**73**] Norman Levinson, *More than one third of zeros of Riemann's zeta-function are on* $\sigma = 1/2$, Advances in Math. **13** (1974) pp. 383-436.

[**74**] Meghann Moriah Lugar, Micah B. Milinovich, and Emily Quesada-Herrera, *On the number variance of zeta zeros and a conjecture of Berry*, Mathematika **69** 2 arXiv 2211.14918; (2023)

[**75**] M. Mehta, *Random Matrix Theory*, Academic Press. Boston; (1991)

[**76**] Francesco Mezzadri, *How to generate random matrices from the classical compact groups*, Notices Amer. Math. Soc. **54** no. 5, (2007) pp. 592-604.

[**77**] H.L. Montgomery, *The pair correlation of zeros of the Riemann zeta-function*, Proc. Symp. Pure Math. **24** (1973) pp. 81-93.

[**78**] Charles M. Newman, *Fourier transforms with only real zeros*, Proc. Amer. Math. Soc. **61** no. 2, (1977) pp. 245-251.

[**79**] A. Odlyzko, *The* 10^{20}*th zero of the Riemann zeta-function and 70 million of its neighbors.*, Preprint (1989)

[**80**] A. Odlyzko, *Zeros number* $10^{12}+1$ *through* $10^{12}+10^4$ *of the Riemann zeta function*, available at http://www.dtc.umn.edu/~odlyzko/zeta_tables;

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

[**82**] Dave Platt and Tim Trudgian, *The Riemann hypothesis is true up to* $3 \cdot 10^{12}$, Bull. Lond. Math. Soc. **53** no. 3, (2021) pp. 792-797.

[**83**] Kyle Pratt, Nicolas Robles, Alexandru Zaharescu, and Dirk Zeindler, *More than five-twelfths of the zeros of* ζ *are on the critical line*, Res. Math. Sci. **7** no. 2, (2020) paper no. 2, 74.

[**84**] The PreTeXt project, *The PreTeXt document authoring system*, https://pretextbook.org

[**85**] Brad Rodgers and Terence Tao, *The de Bruijn-Newman constant is non-negative.*, Forum Math. Pi no. e6, **8** (2020) 62.

[**86**] Mattia Righetti, *Zeros of combinations of Euler products for* $\sigma > 1$, Monatsh. Math. **180** no. 2, (2016) pp. 337-356.

[**87**] Michael Rubinstein, *Low-lying zeros of L-functions and random matrix theory*, Duke Math. J. **109** no. 1, (2001) pp. 147-181.

[**88**] Michael Rubinstein, *Computational methods and experiments in analytic number theory*, *Recent perspectives in random matrix theory and number theory, London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press; no. 322, (2005) pp. 425-506.

[**89**] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996) pp. 269-322.

[**90**] Eric Saias and Andreas J. Weingartner, *Zeros of Dirichlet series with periodic coefficients*, Acta Arithmetica no. 4, **140** (2008) pp. 335-344.

[**91**] E. Saksman and C. Webb, *The Riemann zeta function and Gaussian*

multiplicative chaos: statistics on the critical line, Ann. Probab. no. 2, **48** (2020) pp. 2680-2754.

[92] E. Saksman and C. Webb, *On the Riemann zeta function and Gaussian multiplicative chaos*, Advancements in complex analysis -- from theory to practice Springer; (2020) pp. 473-496.

[93] Will Sawin, *A refined random matrix model for function field L-functions*, Preprint (2024) https://arxiv.org/abs/2409.02876;

[94] A. Selberg, *Contributions to the theory of the Riemann zeta-function*, Arch. Math. Naturvid. **48** no. 5, (1946) pp. 89-155.

[95] 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)* Univ. Salerno, Salerno; (1992)

[96] O. Shanker, *Good-to-bad Gram point ratio for Riemann zeta-function*, Experimental Math. **30** no. 1, (2021) pp. 76-85.

[97] K. Soundararajan, *The distribution of values of zeta and L-functions*, ICM--International Congress of Mathematicians. Vol. 2. Plenary lectures pp. 1260-1310. (2023) arXiv 2112.03389;

[98] A. Speiser, *Geometrisches zur Riemannschen Zetafunktion*, Math Ann. **110** (1935) pp. 514-521.

[99] Norbert Tihanyi, *Fast method for locating peak values of the Riemann zeta function on the critical line*, 2014 16th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing IEEE; (2014) pp. 91-96.

[100] Norbert Tihanyi, Atilla Kovács, and József Kovács, *Computing extremely large values of the Riemann zeta function*, J. Grid Computing **15** (2017) pp. 527-534.

[101] E. C. Titchmarsh, *The theory of the Riemann zeta-function. Second edition. Edited and with a preface by D. R. Heath-Brown*, The Clarendon Press, Oxford University Press, New York; (1986)

[102] T. Trudgian, *On the success and failure of Gram's law and the Rosser rule*, Acta Arith. **148** (2011) pp. 225-256.