Riemann Hypothesis

Peter Sarnak

CMI Millennium Prize Lecture
Harvard, April 15, 2026.

STATEMENT (RIEMANN 1859)

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p} (1 - p^{-s})^{-1}$$

(UNIQUE FACTORIZATION)

Extends to a holomorphic (except for a simple pole at $s = 1$) function of the complex variable s and satisfies

Functional equation :

$$\Xi(s) := \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \Xi(1 - s)$$

Riemann Hypothesis (RH) : All the zeros of $\zeta(s)$ have $\Re(\rho) = \frac{1}{2}$.

- As uncovered by Siegel (1932), Riemann computed the first few zeros and checked that they satisfied his hypothesis. How ? He needed a good approximation to $\zeta(s)$ in $0 < \Re(s) < 1$. Now known as the Riemann–Siegel formula, which remains the gold standard in such calculations (conductor up to 10^{12}).

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f'(z)}{f(z)} dz = 1$$

- The largest zero that appears to have been computed is around the 10^{36}th zero (Böser–Hiary). All zeros that have been computed satisfy RH and are simple (if not, we would not be able to verify it numerically).

For Riemann it was a working hypothesis

Since with it he could complete his aim from his explicit formula :

$$\psi(x) := \sum_{p \leq x} \log p = x - \sum_{\rho} \frac{x^{\rho}}{\rho}$$

(all zeros) that $\psi(x) \sim x$ as $x \to \infty$. That is the Prime Number Theorem (PNT).

While Riemann did not revisit his hypothesis as far as we know, his approach to the PNT via $\zeta(s)$ led to a proof of the latter by Hadamard and de la Vallée Poussin (1896). They showed that

$$\Re(\rho) < 1 \quad \text{for all zeros } \rho.$$

RH has characteristics of a good problem :

(a) Elegant ; easy to state ; appealing.

(b) Falsifiable.

But why all the fuss ?

- Proving RH would be a monumental achievement. Earn the Clay prize money and much more.
- We expect that it would entail new understandings about the integers and that it would apply to more general zeta functions below. If not, it would be disappointing as $\zeta(s)$ itself concerns the Archimedean distribution of prime numbers, which is limited.

A typical sharp consequence is :

$$\text{Almost all intervals } [x, x + (\log x)^2]$$

contain a prime number (Selberg, 1943).

- **Disproving RH** by finding a ρ with $\Re(\rho) > \frac{1}{2}$ and at height say 10^5 should earn the prize money (however read the Clay legal fine-print carefully !).
- **What would the impact be ?** (Nick Treffen)

— Curiously, because of a feature of "zero repulsion" and Landau–Siegel zeros, such a ρ could probably be used to resolve some effectivity questions.

- But the real impact would be for those like Museuf who believe that RH and much more is true : one such zero would burst our bubble.
- While the numerical evidence for RH appears to be very convincing, one should temper this with the "complexity" that is the local density of zeros at height T is only $\log T$, which is not so large.

Generalized Riemann Hypothesis (GRH)

The Riemann zeta function has natural generalizations to Dedekind zeta functions and L-functions of automorphic forms on GL_n. The analogue of RH for these is a powerful working hypothesis and, if true, a profound feature of wave numbers.

For $K/\mathbb{Q}$ an extension of degree n and discriminant Δ_K, its zeta function is

$$\zeta_K(s) = \sum_{\mathfrak{a}} N(\mathfrak{a})^{-s} = \prod_{\mathfrak{p}} (1 - N(\mathfrak{p})^{-s})^{-1}.$$

Statistics (say K has r_1 real embeddings) :

$$Z_K(s) := |\Delta_K|^{s/2} \left(\pi^{-s/2}\right)^{r_1} \zeta_K(s) = Z_K(1-s).$$

Here $D = \Delta_K$ is the "conductor" measuring ramification complexity, and $\log D$ is the complexity in terms of the local density of zeros.

GRH asserts that the zeros of $Z_K(s)$ (and of all automorphic L-functions) have their real part equal to $\frac{1}{2}$.

The L-Function and Modular Forms Data Base : www.lmfdb.org contains a wealth of information about such zeta functions of small conductor and small n, including checks of GRH.

The modern analytic theory of L-functions consists of approximations of GRH :

- Non-vanishing theorems.
- Density theorems.
- Subconvexity.

Which often serve as substitutes for GRH in applications.

A sample :

PNT assuming RH (Riemann)	Proved unconditionally (1896).
There is a polynomial in $\log n$ time algorithm to determine whether a large n is prime or not assuming GRH (Miller, 1975)	Unconditionally Agrawal–Kayal–Saxena (2002).
The André–Oort conjecture concerns the Zariski closure of a sequence of special (CM-points) points in Shimura varieties. Resolved by Edixhoven, Klingler–Yafaev assuming GRH (2014).	Unconditionally Pila–Shahkar–Tsimerman (Esnault–Groechenig) (2022). Though to make the unconditional proof effective, one needs GRH.

- Irreducible polynomials over $\mathbb{Q}$:

$$P(x) = x^d + a_1 x^{d-1} + \cdots + a_d x + 1$$

with $a_j \in \{0, 1\}$ chosen with probability $1/2$ (i.i.d.). Then assuming GRH,

$$\text{Prob}(P \text{ is irreducible}) \to 1 \quad \text{as } d \to \infty$$

(proved by Breuillard–Varjú ; conjectured by Odlyzko–Poonen).

- Cubic Diophantine equations : A central critical problem is to give sharp bounds for the number of integers $(x_1, \ldots, x_6) \in \mathbb{Z}^6$, $|x_j| \leq T$, satisfying

$$x_1^3 + x_2^3 + x_3^3 = x_4^3 + x_5^3 + x_6^3.$$

Hua showed $N(T) \ll T^{7/2}$. Hooley, assuming GRH for Hasse–Weil zeta functions of associated 3-folds :

$$N(T) \ll T^{3+\varepsilon}, \quad \varepsilon > 0$$

which is sharp in the exponent.

Analogues

The setting in which great progress has been made is that of **function fields**.
Rationals $\mathbb{Q}$ replaced by $k = \mathbb{F}_q(t)$, finite field. Integers by $\mathbb{F}_q[t]$ polynomials. Primes : irreducible polynomials.

$$\mathrm{Norm}(f) = q^{\deg f}, \qquad T = q^{-s}.$$

$$\sum_{(f)} \mathrm{Norm}(f)^{-s} = \sum_{(f)} T^{\deg f} = \frac{1}{(1-T)(1-qT)}.$$

If K is a finite extension of k (or a curve over $\mathbb{F}_q$), then its zeta function takes the shape :

$$Z_K(T) = \frac{P_K(T)}{(1-T)(1-qT)}.$$

P_K has degree $2g$, where g is the genus of K ; so # of zeros is $2g$.

$$Z_K(T) \quad \text{has a functional equation invariance under } T \mapsto \frac{1}{qT}.$$

Artin constructed the analogue of GRH : all the zeros have $|\rho| = \sqrt{q}$.

Weil formulated and conjectured generalizations to varieties of $\sqrt{q}$. The proofs of these is one of the highlights of 20th century mathematics.

- Weil for curves.
- Grothendieck, Deligne for varieties.

Crucial to these developments is the **linearization** which realizes the zeros as eigenvalues of the Frobenius morphism $x \mapsto x^q$ acting on cohomology groups.

It is worth noting that the passage from linearization to showing that the zeros lie on a line (or circle) is a major step (no remarkable "Hilbert–Pólya" type self-adjointness miracle). In general (Deligne) it involves deforming the variety in a family and using the glue of monodromy symmetry to go from one to all.

Checking the criteria for a good problem : GRH

(a) Elegant and falsifiable.

(b) If true it is a powerful truth about numbers.

(c) It has profound analogues and generalizations in related fields.

RH and GRH : Where do we stand today ?

I review some and add to my 2018 Bristol lecture "Commentary and comparisons of some approaches to GRH".

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p (1-p^{-s})^{-1}.$$

(i) The LHS allows one to study $\zeta(s)$ using the additive structure of $\mathbb{R}$ and of $\mathbb{Z}$ as a discrete subgroup with compact quotient. In particular, the Poisson summation formula leads to the analytic properties.

(ii) The RHS captures the multiplicative structure of $\mathbb{Z}$.

That one has to use both (i) and (ii) is made clear by

Davenport–Heilbronn (1936) (no Euler product) : Let $Q(x, y)$ be an integral binary quadratic form of discriminant $d < 0$ for which $R(d)$ is odd and not 1. Then

$$Q_0(s) = \sum_{(x,y)\neq(0,0)} Q(x, y)^{-s}$$

has infinitely many zeros in $\Re(s) > 1$.

Deligne (1980) showed that $Q_0(s)$ has a positive proportion of its zeros on $\Re(s) = \frac{1}{2}$ (matching his result on $\zeta(s)$), showing that such results are statistical in nature.

Mixing the additive and multiplicative structures effectively remains a challenge.

- In the first 30 or 50 years after Riemann's paper it was not clear if it was a difficult problem or not (as Riemann himself expressed). However, by now it is clear that if there is some elementary solution it is well hidden.
- There are plenty equivalences of RH in the literature. It is difficult to predict whether any of these offer insights or not. As long as they yield no new results that have applications along the lines that RH is a working hypothesis, their relevance remains unclear.

For example, if $\lambda(n) = (-1)^{\Omega(n)}$ is the parity of the total number of prime factors of n, then

(i') $\sum_{n\leq x} \lambda(n) = o(x)$ as $x \to \infty$ is equivalent to the PNT.

(ii') $\sum_{n\leq x} \lambda(n) = O\Big(x^{\frac{1}{2}+\varepsilon}\Big)$ for every $\varepsilon > 0$ is equivalent to RH.

While elementary proofs of (i') have been found, most recently a lovely ergodic theoretic proof by Richter and McNamara (2021), these offer little towards the quantified (ii').

(ii') says that the $\lambda(n)$'s have ± 1 values mimicking random ones, but showing randomness features in something explicit is difficult.

- With the success of the linearization through cohomological methods in the function field, one might hope for something similar over $\mathbb{Q}$.
- A promising indication of the spectral nature of the zeros of zeta functions is the **Dyson–Montgomery (1973)**. The local spacing statistics

$$\text{(i.e. if } \rho = \frac{1}{2} + i\gamma, \quad \log T\,(\gamma - \gamma') \text{ corresponding to } \rho, \rho' \text{ with } \gamma \sim T)$$

  of the high zeros of (all) zeta functions follow the statistical laws of random Hermitian matrices : **GUE**.

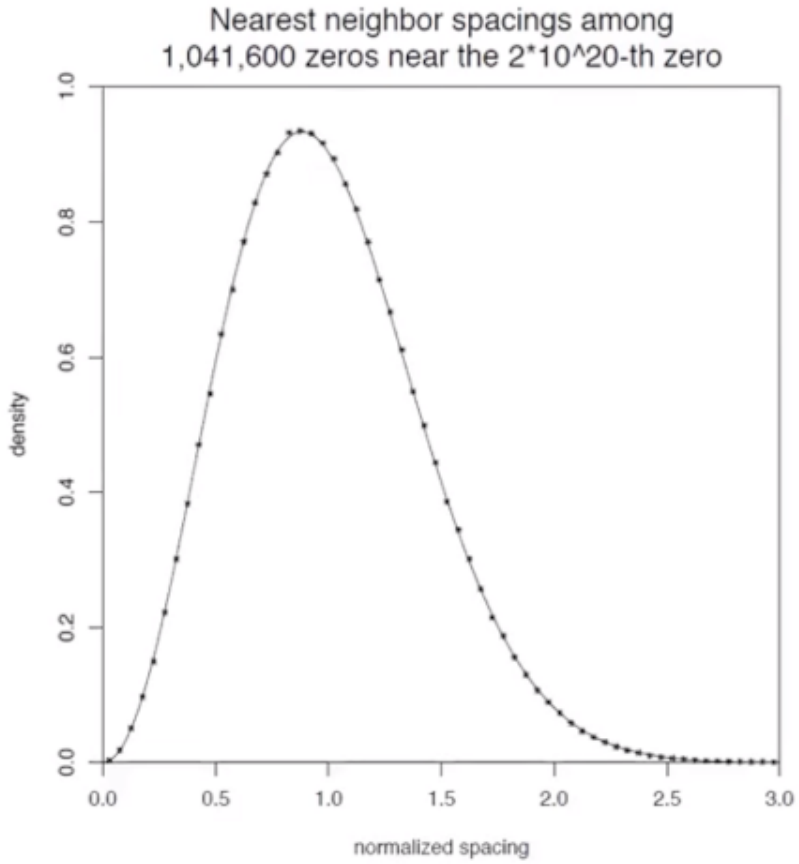

Figure : Nearest neighbor spacings among 1,041,600 zeros near the $2 \cdot 10^{20}$*-th zero.* (density vs normalized spacing)

- This has been verified numerically and theoretically in restricted ranges (assuming GRH). It is very convincing and has led to an active study of such statistics of zeta functions as well as families of such.
- The source in the function field of such a connection is understood (Katz–Sarnak, 1998) and arises through scaling limits of monodromy groups of families of zeta functions.
- One linearization that was suggested by Paul Cohen in 1977 and which has been advanced by Connes (1999) and Meyer (2005) is an action of multiplication over addition on certain adèle/idèle spaces.

Ring of Adèles of $\mathbb{Q}$ (additive) :

$$\mathbb{A} = \{(a_\infty, a_2, a_3, \dots) \mid a_p \in \mathbb{Z}_p \text{ for } p \neq 0\}.$$

Ring of Idèles :

$$\mathbb{A}^* = \left\{ (a_\infty^*, a_2^*, a_3^*, \dots) \mid a_p^* \in \mathbb{Q}_p^* \text{ for } p \neq 0 \right\}.$$

The standard spaces are (the "valuation vectors")

$$\mathbb{A}/\mathbb{Q}, \quad \mathbb{A}^*/\mathbb{Q}^*.$$

Consider $\mathbb{A}^*/\mathbb{Q}^*$, a very singular quotient. The action

$$x \mapsto xy, \quad y \in \mathbb{A}^*/\mathbb{Q}^*, \quad x \in \mathbb{A}^*/\mathbb{Q}^*$$

mixes the structures.

One can introduce suitable topological vector spaces of functions on $\mathbb{A}^*/\mathbb{Q}^*$ such that the trace of the above action yields the Riemann–Guinand–Weil explicit formula (Meyer). In particular, one has the zeros of L-functions interpreted as eigenvalues of this action.

Connes, arXiv :2602.04022 gives an update on how one may use such a spectral interpretation. Also included is an entertaining letter to Professor Bernhard Riemann.

No contradictions other than miscalculations have emerged from these.

There is another species of zeta (or L) functions that come from **transcendental automorphic forms** (cuspidal) on $\mathrm{GL}_n/\mathbb{Q}$ which are everywhere unramified (Maass forms). These elusive forms are the building blocks of the theory and are there only because of the "Archimedean place". They are eigenfunctions of differential operators on the locally symmetric spaces

$$X_n = \mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R}).$$

The "analytic conductors" of their zeta functions, that is the local densities of their zeros, involve their eigenvalues.

GRH $\Rightarrow$ very strong bounds for their eigenvalues as well as a paucity of such forms. In particular, in the $n \to \infty$ limit it implies very strong concentration properties for functions on the core of X_n as $n \to \infty$.

Large n : this limit has not been probed. It will be a real stress test, and I believe it is a test for GRH.

Given its very strong implications, GRH should be probed in extreme limits and for the different species of zeta functions.

Ramification in the $n \to \infty$ limit : The local count of the number of zeros of $\zeta_K(s)$ is controlled by

$$\log N_K = \text{"log of the conductor"}.$$

As such, together with GRH this yields very strong lower bounds for ramification conductor in terms of the degree n of K (Stark, Odlyzko, ...). These bounds are much sharper than those coming from the geometry of numbers (Minkowski). So much so that they come close to violating explicit constructions of towers with little ramification (Golod–Shafarevich).