Axiomatics as a strategy for complex proofs: the case of Riemann Hypothesis

Axiomatic Thinking
Académie Internationale de Philosophie des Sciences
Lisbon, October 11-14, 2017

Jean Petitot
CAMS (EHESS), Paris

Saturday, October 14, 2017

My purpose is to comment some claims of André Weil (1906-1998) in his celebrated letter written in prison to his sister Simone (March 26, 1940, *Collected Papers*, vol.1, 244-255, translated by Martin Krieger, *Notices of the AMS*, 52/3 (2005) 334-341).

Let us begin with the following quotation:

"It is hard for you to appreciate that modern mathematics has become so extensive and so complex that it is essential, if mathematics is to stay as a whole and not become a pile of little bits of research, to provide a unification, which absorbs in some simple and general theories all the common substrata of the diverse branches of the science, suppressing what is not so useful and necessary, and leaving intact what is truly the specific detail of each big problem. This is the good one can achieve with axiomatics (and this is no small achievement). This is what Bourbaki is up to." (p. 341)

I want to emphasize four points:

1. The *unity* of mathematics ("to stay as a whole").
2. The axiomatization of general structures, AND
3. The requirement of "leaving intact what is truly the specific detail of each big problem."
4. The insistance on "big problems".

For Weil (and Bourbaki) the dialectic balance between **general structures** and **specific details** is crucial.

A "big problem" needs a conceptually *complex* proof which is a very uneven, rough, rugged *multi*-theoretical route in a sort of "Himalayan chain" whose peaks seem inaccessible.

It cannot be understood without the key thesis of the *unity* of mathematics since its deductive parts are widely scattered in the global unity of the mathematical universe.

It is *holistic* and it is this holistic nature I am interested in.

As was emphasized by Israel Kleiner for the Shimura-Taniyama-Weil (STW) conjecture (Fermat theorem):

> *"What area does the proof come from? It is unlikely one could give a satisfactory answer, for the proof brings together many important areas – a characteristic of recent mathematics."*

As was also emphasized by Barry Mazur:

> *"The conjecture of Shimura-Taniyama-Weil is a profoundly unifying conjecture — its very statement hints that we may have to look to diverse mathematical fields for insights or tools that might leads to its resolution."*.

In his letter to Simone, Weil described in natural language his moves towards *Riemann Hypothesis* and he used a lot of military metaphors to emphasize the fact that finding a proof of a so highly complex conjecture is a problem of *strategy*:

> *"find an opening for an attack (please excuse the metaphor)", "open a breach which would permit one to enter this fort (please excuse the straining of the metaphor)", "it is necessary to inspect the available artillery and the means of tunneling under the fort (please excuse, etc.)".*

> *"It will not have escaped you (to take up the military metaphor again) that there is within all of this great problems of strategy".*

My purpose is not here to discuss philosophically Bourbaki's concept of structure as mere "simple and general" abstraction.

It has been done by many authors. See e.g. Leo Corry's "*Nicolas Bourbaki: Theory of Structures*" (Chapter 7 of *Modern Algebra and the Rise of Mathematical Structures*, 1996).

And many authors have criticized the very limited Bourbaki's conception of logic.

My purpose is rather to focus on the fact that, for these creative mathematicians, the concept of "structure" is a *functional* concept, which has always a "strategic" *creative* function.

One again, "*leaving intact what is truly the specific detail of each big problem*".

As Dieudonné always emphasized it, the "bourbakist choice " cannot be understood without references to "big problems".

It concerns the context of discovery rather than the context of justification.

There is a fundamental relation between the holistic and "organic" conception of the *unity* of mathematics and the thesis that some *analogies* can be creative and lead to essential discoveries.

It is a leitmotive since the 1948 Bourbaki (alias Dieudonné) Manifesto: "L'architecture des mathématiques" (*Les grands courant de la pensée mathématique*, F. Le Lionnais ed., Cahiers du Sud).

The constant insistance on the "immensity" of mathematics and on its "organic" unity, the claim that "to integrate the whole of mathematics into a coherent whole" (p. 222) is not a philosophical question as for Plato, Descartes, Leibniz or "logistics", the constant critique against the reduction of mathematics to a tower of Babel juxtaposing separed "corners", are not vanities of elitist mathematicians.

They have a very precise, strictly technical function: construct complex proofs in navigating into this holistic conceptually coherent world.

> *"The "structures" are tools for the mathematician." (p. 227)*

"Each structure carries with it its own language" and to discover a structure in a concrete problem

> *"illuminates with a new light the mathematical landscape" (p. 227)*

Leo Corry has well formulated the key point:

> *"In the "Architecture" manifesto, Dieudonné also echoed Hilbert's belief in the unity of mathematics, based both on its unified methodology and in the discovery of striking analogies between apparently far-removed mathematical disciplines." (p.304)*

And indeed, Dieudonné claimed that

> *"Where the superficial observer sees only two, or several, quite distinct theories, lending one another "unexpected support" through the intervention of mathematical genius, the axiomatic method teaches us to look for the deep-lying reasons for such a discovery".*

It is important to understand that structures are guides for *intuition* and to overcome

> *"the natural difficulty of the mind to admit, in dealing with a concrete problem, that a form of intuition, which is not suggested directly by the given elements, [...] can turn out to be equally fruitful." (p. 230)*

So

> *"more than ever does intuition dominate in the genesis of discovery" (p. 228)*

and **intuition is guided by structures**.

Navigating within the mathematical Hymalayan chain

A proof of RH would be highly complex and unfold in the labyrinth of many different theories.

As Connes explains in "*An essay on Riemann Hypothesis*" (p. 2) we would have (note the strategy metaphor as in Weil)

> *"to navigate between the many forms of the explicit formulas [see below] and possible strategies to attack the problem, stressing the value of the elaboration of new concepts rather than "problem solving"."*

For Connes "concepts" mean "structures".

In the history of RH we meet an incredible amount of deep and heterogeneous mathematics.

1. Riemann's use of complex analysis in arithmetics: ζ-function, the duality between the distribution of primes and the localization of the non trivial zeroes of $\zeta(s)$, RH.
2. The "algebraization" of Riemann's theory of complex algebraic (projective) curves (compact Riemann surfaces) by Dedekind and Weber.
3. The transfer of this algebraic framework to the arithmetics of algebraic number fields and the interpretation of integers n as "functions" on primes p. It is the archeology of the concept of spectrum (the scheme $\mathrm{Spec}(\mathbb{Z})$).

4. The move of André Weil introducing an intermediary third world (his "Rosetta stone") between arithmetics and the algebraic theory of compact Riemman surfaces: the world of projective curves over *finite* fields (characteristic $p \geq 2$). The translation of RH in this context and its far reaching proof using tools of algebraic geometry (divisors, Riemann-Roch theorem, intersection theory, Severi-Castelnuovo inequality) coupled with the action of Frobenius maps in characteristic $p \geq 2$.
5. The generalization of RH to higher dimensions in characteristic $p \geq 2$. The Weil's conjectures and the formal reconstruction of algebraic geometry achieved by Grothendieck: schemes, sites, topoï, etale cohomology, etc. Deligne's proof of Weil's conjectures. Alain Connes emphasized the fact that, through Weil's vision, Grothendieck's culminating discoveries proceeds from RH:

"It is a quite remarkable testimony to the unity of mathematics that the origin of this discovery [topos theory] lies in the greatest problem of analysis and arithmetic." (p. 3)

6. Connes' return to pure arithmetics and the original RH by translating algebraic geometry *à la* Grothendieck (topoï, etc.) and Weil's proof in characteristic $p \geq 2$ to the world of characteristic 1, that is the world of *tropical geometry* and *idempotent analysis*.

The distribution of primes

The story of RH begins with the enigma of the distribution of primes. The multiplicative structure of integers (divisibility) is awful.

For $x \geq 2$, let $\pi(x)$ be the number of primes $p \leq x$.

It is a *step function* increasing of 1 at every prime p (one takes $\pi(p) = \frac{1}{2}(\pi(p_-) + \pi(p_+))$ the mean value at the jump).

From Legendre (1788) and the young Gauss (1792) to Hadamard (1896) and de la Vallée Poussin (1896) it has been proved the asymptotic formula called the *prime number theorem*:

$$\pi(x) \sim \frac{x}{\log(x)}$$

The zeta function $\zeta(s)$ encodes arithmetic properties of $\pi(x)$ in analytic structures. Its initial definition is extremely simple and led to a lot of computations at Euler time:

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}$$

which is a series – now called a Dirichlet series – absolutely convergent for integral exponents $s > 1$.

Euler already proved $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$.

A trivial expansion and the existence of a *unique* decomposition of any integer in a product of primes show that, in the convergence domain, the sum is equal to an infinite Euler product (Euler 1748) containing a factor for each prime p (we note $\mathcal{P}$ the set of primes):

$$\zeta(s) = \prod_{p \in \mathcal{P}} \left(1 + \frac{1}{p^s} + \dots \frac{1}{p^{ms}} + \dots\right) = \prod_{p \in \mathcal{P}} \frac{1}{1 - \frac{1}{p^s}}$$

The *local* ζ-functions $\zeta_p(s) = \sum_{k \geq 0} \frac{1}{p^{ks}} = \frac{1}{1 - \frac{1}{p^s}}$ are the ζ-functions of the local rings $\mathbb{Z}_p$ of p-adic integers.

The ζ-function is a symbolic expression associated to the distribution of primes, which is well known to have a very mysterious structure.

Its fantastic strength as a tool comes from the fact that *it can be extended by analytic continuation to the complex plane*.

It has a simple pole at $s = 1$ with residue 1.

Mellin transform, theta function and functional equation

It was discovered by Riemann in his celebrated 1859 paper "Über die Anzahl der Primzahlen unter einer gegeben Grösse" ("On the number of prime numbers less than a given quantity"), that $\zeta(s)$ has also beautiful properties of symmetry.

This can be made explicit noting that $\zeta(s)$ is related by a *Mellin transform* (a sort of Fourier transform), to the *theta function* which possesses beautiful properties of automorphy.

Let $\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx$ be the *gamma function*, which is the analytic meromorphic continuation of the factorial function $(\Gamma(n+1) = n!)$ to the entire complex plane $\mathbb{C}$.

Γ satisfies the functional equation:

$$\Gamma(s+1) = s\Gamma(s)$$

Γ has poles at $s \in -\mathbb{N}$. The figure 1 shows it on the real axis.

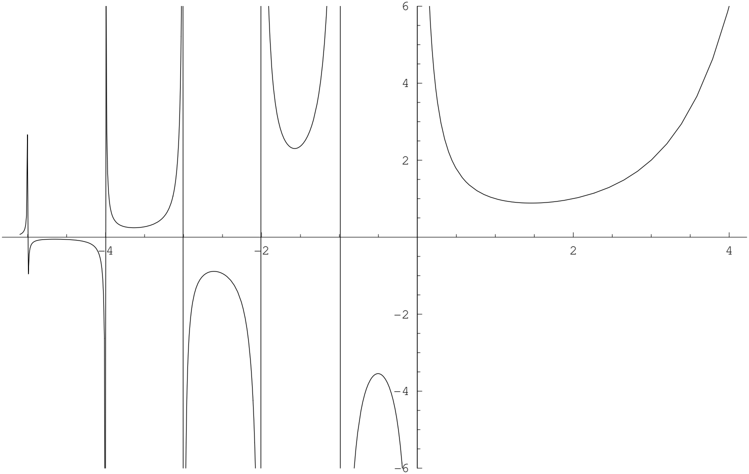

Figure: The Γ function on the real axis.

Let

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

$\zeta^*(s)$ is the *total* ζ-function. Due to the automorphic symmetries of the theta function it satisfies a *functional equation* (symmetry w.r.t. the critical line $\Re(s) = \frac{1}{2}$)

$$\zeta^*(s) = \zeta^*(1-s)$$

As an Euler product

$$\zeta^*(s) = \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\prod_{p\in\mathcal{P}}\frac{1}{1-\frac{1}{p^s}}$$

The factor $\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)$ corresponds to the place at infinity ∞ of $\mathbb{Q}$ (see below) and $\zeta^*(s)$ is a product of factors associated to all the places of $\mathbb{Q}$:

$$\zeta^*(s) = \prod_{p\in\mathcal{P}\cup\{\infty\}}\zeta_p^*(s)$$

with $\zeta_p(s) = \frac{1}{1-\frac{1}{p^s}}$ for $p \in \mathcal{P}$ and $\zeta_\infty(s) = \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)$

Zeroes of $\zeta(s)$

As $\zeta(s)$ is well defined for $\Re(s) > 1$, it is also well defined, via the functional equation of ζ^*, for $\Re(s) < 0$, and the difference between the two domains comes from the difference of behavior of the gamma function Γ.

As ζ^* is without poles on $]1, \infty[$ (since ζ and Γ are without poles), ζ^* is also, by symmetry, without poles on $]-\infty, 0[$. So, as the $s = -2k$ are poles of $\Gamma\left(\frac{s}{2}\right)$, they must be zeroes of ζ.

See figure 2

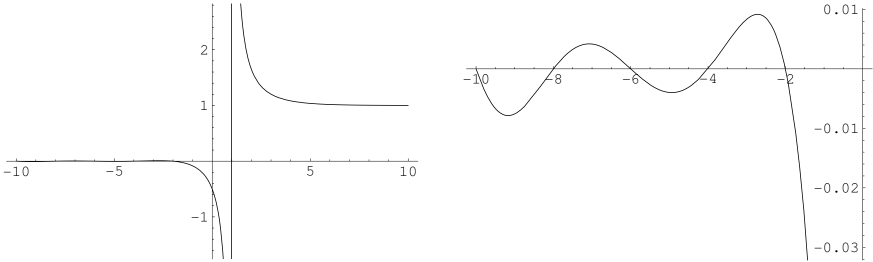

Figure: The graph of the zeta function along the real axis showing the pole at 1 (left). A zoom shows the trivial zeroes at even negative integers (right).

But $\zeta(s)$ has non trivial zeroes **outside** the domain $\Re(s) > 1$ where it is *explicitely defined* by the Euler product.

Due to the functional equation they are symmetric w.r.t. the critical line $\Re(s) = \frac{1}{2}$.

Their distribution reflects the distribution of primes and the *localization* of these zeroes is one of the main tools for understanding the mysterious distribution of primes.

A pedagogical way for seeing the (non trivial) zeroes (J. Arias-de-Reyna) is to plot in the s plane the curves $\Re(\zeta(s)) = 0$ and $\Im(\zeta(s)) = 0$ and to look at their crossings (see figure 3).

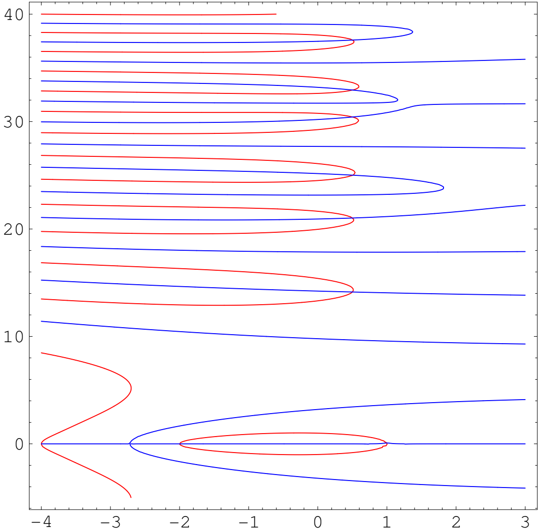

Figure: The null-lines of the real part (red) and the imaginary part (blue) of the zeta function.

It is traditional to write the non trivial zeroes $\rho = \frac{1}{2} + it$ with $t \in \mathbb{C}$. As they code for the irregularity of the distribution of primes, they must be irregularly distributed. But the irregularity can concern $\Re(t)$ and/or $\Im(t)$. When $\Im(t) \neq 0$ we get pairs of symmetric zeroes whose horizontal distance can fluctuate.

An enormous amount of computations from Riemann time to actual supercomputers (ZetaGrid: more than 10^{12} zeroes in 2005) via Gram, Backlund, Titchmarsh, Turing, Lehmer, Lehman, Brent, van de Lune, Wedeniwski, Odlyzko, Gourdon, and others, shows that all computed zeroes lie on the critical line $\Re(s) = \frac{1}{2}$.

Riemann hypothesis

The Riemann hypothesis (part of 8th Hilbert problem) conjectures that **all the non trivial zeroes of** $\zeta(s)$ **are exactly on the critical line**, that is are of the form $\rho = \frac{1}{2} + it$ with $t \in \mathbb{R}$ (i.e. $\Im(t) = 0$).

It is an incredibly strong still open conjecture and an enormous part of modern mathematics has been created to solve it.

Speiser proved that RH is equivalent to the fact that all folded blue lines cross the critical line. One point of intersection (crossing with a red line) is a non trivial zero and the other is called a *Gram point*.

Gram points seem to *separate* the non trivial zeroes (Gram's law), but it is not always the case.

We meet a lot of strange configurations.

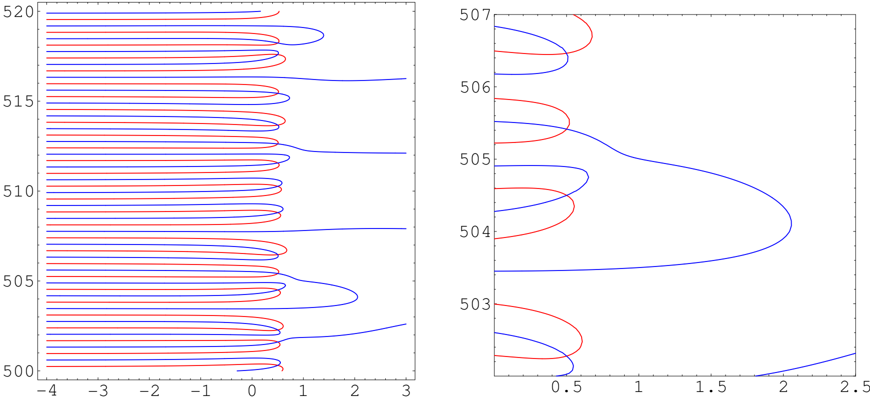

Figure: Configuration where a zero (crossing of folded red and blue lines) is nested. Alternating Gram→zero→Gram→zero.

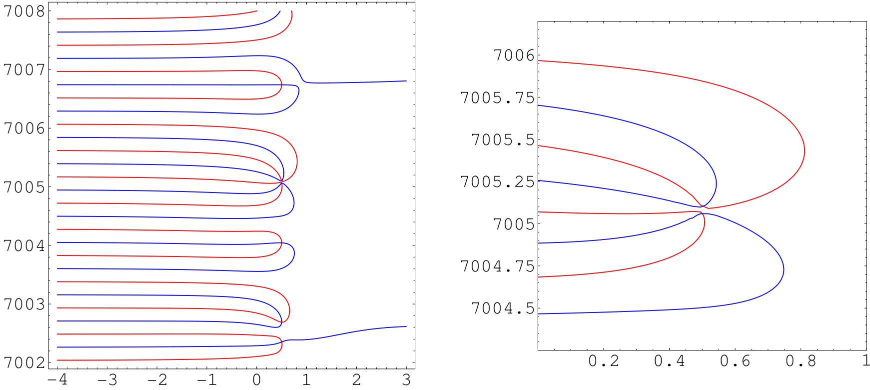

Figure: Lehmer's example of two extremely close consecutive zeroes *between* two Gram points. (We are at the height of the 26 830-th line)

So RH is really not evident.

As noted Pierre Cartier, the risk would be to see a pair of very close “good” zeroes bifurcate into a pair of very close symmetric “bad” zeroes.

The problem of localizing zeroes

The problem is, given the explicit definition of $\zeta(s)$, to find some informations on the *localization* of its zeroes.

As was emphasized by Alain Connes, it is a great generalization of the problem solved by Galois for polynomials (of one variable).

Riemann's explicit formula

One of the most "magical" results of Riemann is the *explicit and exact* formula linking explicitely and exactly the distribution of primes and the (non trivial) zeroes of $\zeta(s)$.

We have seen that, for $x \geq 2$, $\pi(x)$, the number of primes $p \leq x$ satisfies the asymptotic formula (prime number theorem) $\pi(x) \sim \frac{x}{\log(x)}$.

A better approximation, due to Gauss (1849), is $\pi(x) \sim \mathrm{Li}(x)$ where the logarithmic integral is $\mathrm{Li}(x) = \int_2^x \frac{dt}{\log(t)}$ (for small n, $\pi(x) < \mathrm{Li}(x)$, but Littelwood proved in 1914 that the inequality reverses an infinite number of times).

A still better approximation was given by a Riemann formula $R(x)$.

Figure 6 shows the step function $\pi(x)$ and its two approximations $\frac{x}{\log(x)}$ (in gray) and $R(x)$ (in blue).

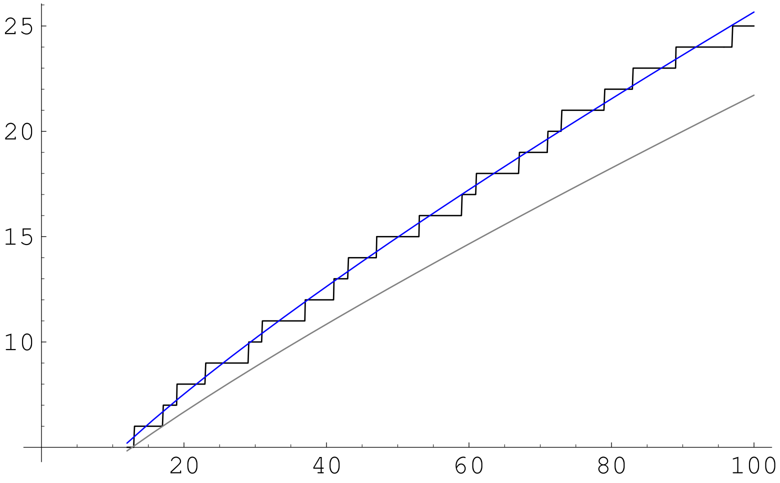

Figure: Two classical approximations of the distribution of primes: $\frac{x}{\log(x)}$ (in gray) and Riemann's $R(x)$ (in blue).

Let

$$f(x) = \sum_{k=1}^{k=\infty} \frac{1}{k} \pi \left(x^{\frac{1}{k}} \right)$$

($\pi(x)$ can be retrieved from $f(x)$ by an inverse transformation).

In his 1859 paper, Riemann proved the following (fantastic) *explicit formula*:

$$f(x) = \mathrm{Li}(x) - \sum_{\rho} \mathrm{Li}\left(x^{\rho}\right) + \int_{x}^{\infty} \frac{1}{t^2 - 1} \frac{dt}{t \log t} - \log 2$$

The approximation of $\pi(x)$ using Riemann's explicit formula up to the twentieth zero of $\zeta(s)$ is shown in figure 7.

We see that the red curve departs from the approximation $R(x)$ and that its oscillations draw near $\pi(x)$.

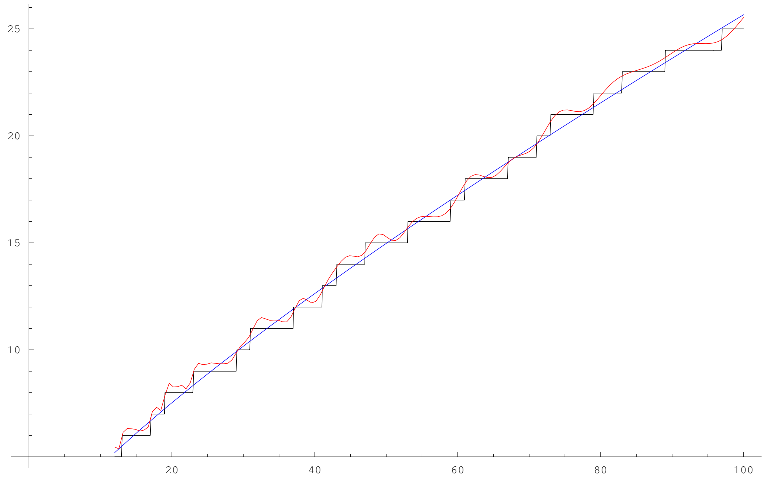

Figure: The approximation (red curve) of $\pi(x)$ by Riemann's explicit formula up to the twentieth zero of $\zeta(s)$.

Riemann's EF concerns only $\zeta(s)$ and therefore only the *p*-adic places of $\mathbb{Q}$ (with completions $\mathbb{Q}_p$).

But we know that the *structural* formulas concern $\zeta^*(s)$ with its Γ-factor and must take into account the Archimedean real place ∞ (with completion $\mathbb{R}$).

This step was accomplished by Weil.

Local/global in arithmetics

Let us go now to the deep analogies discovered between arithmetics and geometry.

One of the main idea, introduced by Dedekind and Weber in their celebrated 1882 paper "Theorie der algebraischen Funktionen einer Veränderlichen" (*J. Reine Angew. Math*, 92 (1882) 181-290. Theory of algebraic functions of one complex variable), was to consider integers n as kinds of "functions" over the sets $\mathcal{P}$ of primes p, "functions" having a value and an order at every "point" $p \in \mathcal{P}$.

These values and orders being local concepts, Dedekind and Weber had to define the concept of localization in a purely algebraic manner.

If p is prime, the ideal $(p) = p\mathbb{Z}$ of p in $\mathbb{Z}$ is a prime (and even maximal) ideal. To localize the ring $\mathbb{Z}$ at p means to delete all the ideals $\mathfrak{a}$ that are not included into (p) and to reduce the arithmetic of $\mathbb{Z}$ to the ideals $\mathfrak{a} \subseteq (p)$.

For that, we add the inverses of the elements of the complementary multiplicative subset S of (p), $S = \mathbb{Z} - (p)$.

We get a *local ring* $\mathbb{Z}_{(p)}$ (that is with a *unique* maximal ideal) intermediary between $\mathbb{Z}$ and $\mathbb{Q}$.

$\mathbb{Z}_{(p)}$ is arithmetically much simpler than the *global* ring $\mathbb{Z}$ but more complex than the global fraction field $\mathbb{Q}$ since it preserves all the arithmetic inside (p).

The maximal ideal of $\mathbb{Z}_{(p)}$ is $\mathfrak{m}_{(p)} = p\mathbb{Z}_{(p)}$ and the residue field is $\mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)} = \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$.

In the local ring $\mathbb{Z}_{(p)}$ every ideal $\mathfrak{a}$ is equal to some power $(p)^k$ of (p).

As $(p)^k \supset (p)^{k+1}$ we get a decreasing sequence – what is called a *filtration* – of ideals which exhausts the arithmetic of $\mathbb{Z}_{(p)}$.

The successive quotients $\mathbb{Z}_{(p)}/p^{k+1}\mathbb{Z}_{(p)}$ correspond to the *expansion* of natural integers n in base p. Indeed, to make $p^{k+1} = 0$ is to approximate n by a sum $\sum_{i=0}^{i=k} n_i p^i$ with all $n_i \in \mathbb{F}_p$.

They constitute a projective system and their projective limit yields the ring $\mathbb{Z}_p$ of p-adic integers:

$$\mathbb{Z}_p = \varprojlim \frac{\mathbb{Z}}{p^k\mathbb{Z}}$$

If $n \in \mathbb{Z}$, to look at n "locally" at p is to look at n in $\mathbb{Z}_{(p)}$.

The "value" of n at p is its class in $\mathbb{F}_p$, i.e. $n \bmod p$ and the local structure of n at p can be read in $\mathbb{Z}_{(p)}$.

It is the origin of the modern concept of *spectrum* in algebraic geometry.

Then $\mathbb{Q}$ becomes the "global" field of "rational functions" on this "space".

Weil's description of Dedekind's analogy

In his letter to Simone, Weil describes very well Dedekind's analogy:

> *"[Dedekind] discovered that an analogous principle permitted one to establish, by purely algebraic means, the principal results, called "elementary", of the theory of algebraic functions of one variable, which were obtained by Riemann by transcendental [analytic] means."*

Since Dedekind's analogy is algebraic it can be applied to other fields than $\mathbb{C}$.

integers	$\longleftrightarrow$	polynomials
divisibility of integers	$\longleftrightarrow$	divisibility of polynomials
rationals	$\longleftrightarrow$	rational functions
algebraic numbers	$\longleftrightarrow$	algebraic functions
Dedekind's "different" ideal	$\underset{\text{Hilbert}}{\longleftrightarrow}$	Riemann-Roch theorem
Abelian extensions	$\longleftrightarrow$	Abelian functions
classes of ideals	$\longleftrightarrow$	divisors

And Weil adds

> *"At first glance, the analogy seems superficial. [... But] Hilbert went further in figuring out these matters."*

Valuations and ultrametricity

Dedekind and Weber defined the *order* of $n \in \mathbb{N}$ at p using the decomposition of n into primes.

Let $n = \prod_{i=1}^{i=r} p_i^{v_i}$. v_i is the valuation of n at p_i: $v_{p_i}(n)$. It is trivial to generalize the definition to $\mathbb{Z}$ and $\mathbb{Q}$. So the valuation $v_p(x)$ of $x \in \mathbb{Q}$ is the power of p in the decomposition of x in prime factors.

It satisfies "good properties" in the sense that $|x|_p = p^{-v_p(x)}$ is a *norm* on $\mathbb{Q}$ defining a *non-Archimedean metric* $d_p(x,y) = |x-y|_p$ which satisfies the *ultrametric* property

$$|x+y|_p \leq \mathrm{Max}\left(|x|_p, |y|_p\right)$$

This inequality is much stronger than the triangular inequality of classical metrics.

p-adic numbers

The idea of expanding natural integers along the base p with a metric such that $\left|p^k\right|_p \underset{k \to \infty}{\longrightarrow} 0$ leads naturally to add a “point at infinity” to the localization $\mathbb{Z}_{(p)}$.

This operation is a *completion* procedure for the metric $|\bullet|_p$ associated to the valuation v_p and yields the ring $\mathbb{Z}_p$ of *p-adic integers* (Hensel).

Hensel's geometric analogy

In Bourbaki's *Manifesto*, Dieudonné emphasizes Hensel's unifying analogy:

> *"where, in a still more astounding way, topology invades a region which had been until then the domain par excellence of the discrete, of the discontinuous, viz. the set of whole numbers." (p. 228)*

As we have already noted, the geometrical lexicon of Hensel's analogy can be rigourously justified using the concept of *scheme*:

1. primes p are the points of the *spectrum* of $\mathbb{Z}$,
2. the finite prime fields $\mathbb{F}_p$ are the fibers of the structural sheaf $\mathcal{O}$ of $\mathbb{Z}$,
3. integers n are global sections of $\mathcal{O}$,
4. and $\mathbb{Q}$ is the field of global sections of the sheaf of fractions of $\mathcal{O}$.

In this context, $\mathbb{Z}_p$ and $\mathbb{Q}_p$ correspond to the localization of global sections, analog to what are called *germs* of sections in classical geometry.

Places

On $\mathbb{Q}$ there exist not only the p-adic valuations of the "finite" primes p but also the real absolute value $|x|$, which can be interpreted as an "infinite" prime and is conventionally written $|x|_\infty$.

To emphasize the geometrical intuition of a "point", the "finite" and "infinite" primes are also called *places*.

To work in arithmetics with *all* places is a necessity if we want to continue the analogy with projective (birational) algebraic geometry (Riemann surfaces) and transfer some of its results (as those of the Italian school of Severi, Castelnuovo, etc.) to arithmetics.

Indeed, in projective geometry the point ∞ is on a par with the other points.

Weil emphasized strongly this point from the start.

Already in his 1938 paper "*Zur algebraischen Theorie des algebraischen Funktionen*" (*Journal de Crelle*, 179 (1938) 129-138) he explains that he wants to reformulate Dedekind-Weber in a *birationally invariant* way.

In his letter to Simone, he says

> *"In order to reestablish the analogy [lost by the singular role of ∞ in Dedekind-Weber], it is necessary to introduce, into the theory of algebraic numbers, something that corresponds to the point at infinity in the theory of functions."*

It is achieved by valuations, places and Hensel's p-adic numbers (plus Hasse, Artin, etc.).

So, Weil strongly stressed the use of analogies as a discovery method:

> *"If one follows it in all of its consequences, the theory alone permits us to reestablish the analogy at many points where it once seemed defective: it even permits us to discover in the number field simple and elementary facts which however were not yet seen."*

Local and global fields

All the knowledge gathered during the extraordinay period initiated by Kummer in arithmetics and Riemann in geometry, led to the recognition of two great classes of fields, local fields and global fields.

In characteristic 0, local fields are $\mathbb{R}$, $\mathbb{C}$ and finite extensions of $\mathbb{Q}_p$.

In characteristic p, local fields are the fields of Laurent series over a finite field $\mathbb{F}_{p^n}$ and their ring of integers are those of the corresponding power series.

Local fields possess a discrete valuation v and are complete for the associated metric. Their ring of integers is local. Finite extensions of local fields are themselves local.

In characteristic 0, global fields are finite extensions $\mathbb{K}$ of $\mathbb{Q}$, i.e. algebraic number fields.

In characteristic p, global fields are the fields of rational functions of algebraic curves over a finite field $\mathbb{F}_{p^n}$. The completions of global fields at their different places are local fields.

The RH for elliptic curves over $\mathbb{F}_q$ (Hasse)

One of the greatest achievements of Weil has been the *proof* of RH for the global fields other than $\mathbb{Q}$, namely the global fields $\mathbb{K}/\mathbb{F}_q(T)$ of rational functions on an algebraic curve defined over a finite field $\mathbb{F}_q$ of characteristic p ($q = p^n$), that is finite algebraic extensions of $\mathbb{F}_q(T)$.

The “Rosetta stone”

The main difficulty was that in Dedekind-Weber’s analogy between arithmetics and the theory of Riemann surfaces, the latter is “too rich” and “too far from the theory of numbers”. So

> *“One would be totally obstructed if there were not a bridge between the two.” (p. 340)*

Hence the celebrated metaphor of the “Rosetta stone”:

> *“my work consists in deciphering a trilingual text; of each of the three columns I have only disparate fragments; I have some ideas about each of the three languages: but I know as well there are great differences in meaning from one column to another, for which nothing has prepared me in advance. In the several years I have worked at it, I have found little pieces of the dictionary.” (p. 340)*

From the algebraic number theory side, one can transfer the Riemann-Dirichlet-Dedekind ζ and L-functions (Artin, Schmidt, Hasse) to the algebraic curves over $\mathbb{F}_q$.

In this third world they become *polynomials*, which simplifies tremendously the situation.

The Hasse-Weil function

For the history of the ζ-function of curves over $\mathbb{F}_q$, see Cartier's 1993 paper "*Des nombres premiers à la géométrie algébrique (une brève histoire de la fonction zeta)*" (Cahiers du Séminaire d'Histoire des mathématiques (2ème série), tome 3 (1993) 51-77).

1. On the *arithmetic* side (spec($\mathbb{Z}$), $\mathbb{Q}_p$, etc.), we have RH.
2. On the geometric side, we have the theory of compact Riemann surfaces (projective algebraic curves on $\mathbb{C}$).

On the *intermediary* level, at the beginning of the XX-th century Emil Artin (thesis, 1921 published in 1924) and Friedrich Karl Schmidt (1931) formulated the RH no longer for global number fields $\mathbb{K}/\mathbb{Q}$ but for global fields of functions $\mathbb{K}/\mathbb{F}_q(T)$. As Cartier says,

> *"La théorie d'Artin-Schmidt se développe donc en parallèle avec celle de Dirichlet-Dedekind, et elle s'efforce de calquer les résultats acquis : définition par série de Dirichlet et produit eulérien, équation fonctionnelle, prolongement analytique." (p. 61)*

The main challenge was to interpret *geometrically* the zeta-function $\zeta_C(s)$ for algebraic curves C defined over $\mathbb{F}_q$.

It took a long time to understand that ζ_C was a *counting function*, counting the (finite) number $N(q^r)$ of points of C rational over the successive extensions $\mathbb{F}_{q^r}$ of $\mathbb{F}_q$:

C being defined over $\mathbb{F}_q$, all its points are with coordinates in $\overline{\mathbb{F}_q}$, and we can look at its points with coordinates in intermediary extensions $\mathbb{F}_q \subset \mathbb{F}_{q^r} \subset \overline{\mathbb{F}_q}$.

The generating function of the $N(q^r)$ is by definition

$$Z_C(T) := \exp\left(\sum_{r\geq 1} N(q^r)\frac{T^r}{r}\right)$$

The Hasse-Weil function $\zeta_C(s)$ of C is defined as

$$\zeta_C(s) := Z_C\left(q^{-s}\right)$$

It corresponds to the two expressions of the classical Riemann's ζ-function (Dirichlet series and Euler product) if one introduces the concept of a *divisor* D on C as a finite $\mathbb{Z}$-linear combination of points of C: $D = \sum_j a_j x_j$.

The *degree* of D is $\deg(D) = \sum_j a_j$.

D is said to be *positive* ($D \geq 0$) if all $a_j \geq 0$.

Then

$$\zeta_C(s) = \sum_{D>0} \frac{1}{N(D)^{-s}} = \prod_{P>0} \left(1 - N(D)^{-s}\right)^{-1}$$

where D are positive divisors on $\mathbb{F}_q$-points, P prime positive divisors (i.e. P is not the sum of two smaller positive divisors) and $N(D) = q^{\deg(D)}$.

The key problem is, as before, the *localization* of the zeroes of $\zeta_C(s)$.

If ρ is a zero, $q^{-\rho}$ is a zero of Z_C.

Conversely, if $q^{-\rho}$ is a zero and if $\rho' = \rho + k\frac{2\pi i}{\log(q)}$, then $q^{-\rho'} = q^{-\rho}$ is also a zero.

So the zeroes of the Hasse-Weil function $\zeta_C(s)$ come in *arithmetic progressions*. It is a fundamental new phenomenon.

Divisors and classical Riemann-Roch (curves)

On the other direction, one try to transfer to curves over $\mathbb{F}_q$ the results of the theory of Riemann compact surfaces, and in particular the *Riemann-Roch theorem*.

If C is a compact Riemann surface of genus g, to deal with the distribution and the orders of zeroes and poles of meromorphic functions on C, one introduces the concept of a *divisor* D on C as a $\mathbb{Z}$-linear combination of points of C:

$D = \sum_{x \in C} \mathrm{ord}_x(D)x$ with $\mathrm{ord}_x(D) \in \mathbb{Z}$ the order of D at x.

All the terms vanish except a finite number of them.

The *degree* of D is then defined as $\deg(D) = \sum_{x \in C} \mathrm{ord}_x(D)$. It is additive.

D is said to be *positive* ($D \geq 0$) if $\mathrm{ord}_x(D) \geq 0$ at every point x.

If f is a meromorphic function on C, poles of order k can be considered as zeroes of order $-k$ and the divisor $(f) = \sum_{x \in C} \mathrm{ord}_x(f)x$ is called *principal* and its degree vanishes: $\deg(f) = \sum_{x \in C} \mathrm{ord}_x(f) = 0$.

By construction, divisors form an additive group $\mathrm{Div}(C)$ and, as the meromorphic functions contitute a field $K(C)$ having the property that the order of a product is the sum of the orders, principal divisors constitute a subgroup $\mathrm{Div}_0(C)$.

The quotient group $\mathrm{Pic}(C) = \mathrm{Div}(C)/\mathrm{Div}_0(C)$, that is the group of classes of divisors modulo principal divisors, is called the *Picard group* of C.

If ω and ω' are two meromorphic differential 1-forms on C, $\omega' = f\omega$ for some $f \in K(C)^*$ (the set of invertible elements of $K(C)$), $\mathrm{div}(\omega') = \mathrm{div}(\omega) + (f)$ and therefore the class of $\mathrm{div}(\omega)$ $\mathrm{mod}\,(\mathrm{Div}_0(C))$ is unique: it is called the *canonical class* of C and one can show that its degree is $\deg(\omega) = 2g - 2$.

For instance, if $g = 0$, C is the Riemann sphere $\widehat{\mathbb{C}}$ and the standard 1-form is $\omega = dz$ on the open subset $\mathbb{C}$.

Since to have a local chart at infinity we must use the change of coordinate $\xi = \frac{1}{z}$ and since $d\xi = -\frac{dz}{z^2}$, we see that, on $\widehat{\mathbb{C}}$, ω possesses no zero and a single double pole at infinity.

Hence $\deg(\omega) = -2 = 2g - 2$.

For $g = 1$ (elliptic case) $\deg(\omega) = 0$ and there exists holomorphic nowhere vanishing 1-forms. As $C \simeq \mathbb{C}/\Lambda$ (Λ a lattice), one can take $\omega = dz$.

To any divisor D one can associate what is called a *linear system*, that is the set of meromorphic functions on C whose divisor (f) is greater than $-D$:

$$L(D) = \{f \in K(C)^* : (f) + D \geq 0\} \cup \{0\}$$

Since a holomorphic function on C is necessarilly constant (Liouville theorem), we have $L(0) = \mathbb{C}$. One of the most fundamental theorem of Riemann's theory is the theorem due to himself and his disciple Gustav Roch:

Riemann-Roch theorem.
$\dim L(D) = \deg(D) + \dim L(\omega - D) - g + 1$.

If $\dim L(D)$ is noted $\ell(D)$, we get

$$\ell(D) - \ell(\omega - D) = \deg(D) - g + 1$$

Corollary. $\ell(\omega) = 2g - 2 + 1 - g + 1 = g$ (since $\ell(0) = 1$).

A very important conceptual improvement of RR is due to Pierre Cartier in the 1960s using the new tools of *sheaf theory* and *cohomology*.

Let $\mathcal{O} = \mathcal{O}_C$ be the structural sheaf of rings $\mathcal{O}(U)$ of holomorphic functions on the open subsets U of C, and $\mathcal{K} = \mathcal{K}_C$ the sheaf of fields $\mathcal{K}(U)$ of meromorphic functions.

To any divisor D, Cartier was able to associate a line bundle on C with a sheaf of sections $\mathcal{O}(D)$.

Then he has shown that the $\mathbb{C}$-vector space of global sections of $\mathcal{O}(D)$ can be identified with $L(D)$, i.e. $L(D) = H^0(C, \mathcal{O}(D))$.

This *cohomological* interpretation is fundamental and allows a deep "conceptual" cohomological interpretation of RR using the fact that $\dim L(D) = \dim H^0(C, \mathcal{O}(D))$.

Divisors and classical Riemann-Roch (surfaces)

For surfaces S over $\mathbb{C}$, RR is more involved. Divisors are now $\mathbb{Z}$-linear combinations no longer of points but of curves C_i.

One has to use what is called the *intersection number* of two curves $C_1 \bullet C_2$ (and of divisors $D_1 \bullet D_2$).

For two curves *in general position* one defines $C_1 \bullet C_2$ in an intuitive way as the sum of the points of intersection.

One shows that, as the base field $\mathbb{C}$ is algebraically closed, this number is invariant by linear equivalence.

One shows also that for any divisors D_1 and D_2, even when $D_1 = D_2$, there exist $D_1' \sim D_1$ and $D_2' \sim D_2$ which are in general position.

One then defines $D_1 \bullet D_2 = D_1' \bullet D_2'$.

The RR theorem is then

$$\sum_{j=0}^{j=2} (-1)^j \dim H^j(S, \mathcal{O}(D)) = \frac{1}{2} D \bullet (D - K_S) + \chi(S)$$

with $\chi(S) = 1 + p_a$, p_a being the "arithmetic genus".

What is called *Serre duality* says that

$$\dim H^2(S, \mathcal{O}(D)) = \dim H^0(S, \mathcal{O}(K_S - D))$$

Now, $\dim H^0$ and $\dim H^2$ are ≥ 0 while $-\dim H^1$ is ≤ 0, so one gets the RR *inequality* :

$$\ell(D) + \ell(K_S - D) \geq \frac{1}{2} D \bullet (D - K_S) + \chi(S)$$

RR for curves over $\mathbb{F}_q$

From Artin to Weil, the theory of compact Riemann surfaces has been transfered to the intermediary case of the curves $C/\mathbb{F}_q$. Schmidt and Hasse transfered the RR theorem.

A fundamental consequence was that $Z_C(T)$ not only satisfies a *functional equation* but is a *rational function* of T.

For instance, let us consider the simplest case $\mathbb{K} = \mathbb{F}_q(T)$ (analogous to the simplest number field $\mathbb{Q}$).

Each unitary polynomial $P(T) = T^m + a_1 T^{m-1} + \ldots + a_m$ of degree m gives a contribution $(q^m)^{-s}$ to the additive (Dirichlet) formulation of $Z_{\mathbb{K}}(T)$ since the norm of its ideal is q^m.

But there are q^m such polynomials since the m coefficients a_j belong to $\mathbb{F}_q$. So

$$\begin{cases} \zeta_{\mathbb{K}}(s) = \sum_{m=0}^{m=\infty} q^m (q^m)^{-s} = \frac{1}{1-q^{1-s}} \\ Z_C(T) = \frac{1}{1-qT} \end{cases}$$

Hence, as $Z_C(T)$ is a rational function of T, it has a *finite* number of zeroes $t_1, \dots, t_M$ and therefore, the zeroes of $\zeta_C(s)$ are organized in a *finite number of arithmetic progressions* $\rho_j + k\frac{2\pi i}{\log(q)}$ with $q^{-\rho_j} = t_j$.

This is a fundamental difference with the arithmetic case, which makes the proof of RH easier.

The Frobenius

But in the $\mathbb{F}_q$ case, a completely original phenomenon appears.

A fundamental property of any finite field $\mathbb{F}_q$ is that $x^q = x$ for every x. So, one can consider the *automorphism* φ_q of $\overline{\mathbb{F}_q}$, $\varphi_q : x \mapsto x^q$ (it is an automorphism) and retrieve $\mathbb{F}_q$ as the field of *fixed points* of φ_q.

φ_q is called the *Frobenius* morphism.

For a curve $C/\mathbb{F}_q$, the Frobenius φ_q acts, for every r, on the set of points $C(\mathbb{F}_{q^r})$ with coordinates in $\mathbb{F}_{q^r}$, and the number $N_r = N(q^r)$ of points of C rational over $\mathbb{F}_{q^r}$ is *the number of fixed points* of the Frobenius φ_{q^r}.

So the generating counting function $Z_C(T)$ counts fixed points and has to do with the world of *trace formulas* counting fixed points of maps.

In particular, $N_1 = C(\mathbb{F}_q) = \#\varphi_q^{\text{Fix}} = \left|Ker\left(\varphi_q - Id\right)\right|$. It is like a "norm".

RH for elliptic curves (Schmidt and Hasse)

Schmidt was the first to add the point at infinity (projective curves) and to understand that, in the case of $\mathbb{K}/\mathbb{F}_q(T)$, the functional equation of ζ_C was correlated to the duality between divisors D and $D - K$ in Riemann's theory. As Cartier says

> *"on voit se manifester ici l'une des premières apparitions de la tendance à la géométrisation dans l'étude de la fonction ζ." (p. 69)*

Schmidt proved that

$$Z_C(T) = \frac{L(T)}{(1-T)(1-qT)} \text{ with } L(T) \text{ a polynomial of degree } 2g$$

He showed also that $L(T)$ is in fact the *characteristic polynomial* of the Frobenius φ_q, i.e. the "norm" (the determinant) of $Id - T\varphi_q$.

So

$$Z_C(T) = \frac{\det\left(Id - T\varphi_q\right)}{(1-T)(1-qT)}$$

and $Z_C(T)$ satisfies the functional equation

$$Z_C\left(\frac{1}{qT}\right) = q^{1-g}T^{2-2g}Z_C(T)$$

For ζ_C the symmetric functional equation is

$$q^{(g-1)s}\zeta_C(s) = q^{(g-1)(1-s)}\zeta_C(1-s)$$

Then, in three fundamental papers of 1936, Hasse proved RH for *elliptic curves*.

As $g = 1$, $L(T)$ is a polynomial of degree 2. And as C is elliptic, it has a *group* structure (C is isomorphic to its Jacobian $J(C)$), which is used as a crucial feature in the proof.

Indeed, one can consider the *group* endomorphisms $\psi : C \rightarrow C$ and their graphs Ψ in $C \times C$, what Hasse called *correspondences*.

For $g = 1$, $Z_C(T)$ satisfies the functional equation

$$Z_C\left(\frac{1}{qT}\right) = Z_C(T)$$

and ζ_C the symmetric functional equation

$$\zeta_C(s) = \zeta_C(1-s)$$

as Riemann's ζ .

Then, Hasse proved that, due to the functional equation, $L(T)$ is the polynomial $L(T) = 1 - c_1 T + qT^2$ with

$$L(1) = 1 - c_1 + q = N_1 = |C(\mathbb{F}_q)|$$

So

$$L(T) = (1 - \omega T)(1 - \overline{\omega} T)$$

with $\omega\overline{\omega} = q$ and $\omega + \overline{\omega} = c_1$ the *inverses* of the zeroes since

$$L(T) = \omega\overline{\omega}\left(T - \frac{1}{\omega}\right)\left(T - \frac{1}{\overline{\omega}}\right)$$

As $|\omega| = |\overline{\omega}|$, we have $|\omega| = \sqrt{q}$. But, since $\zeta_C(s) = Z_C(q^{-s})$, the zeroes of $\zeta_C(s)$ correspond to $q^{-s_j} = (\omega_j)^{-1}$. So we must have

$$\left|q^{-s_j}\right| == |q|^{-\Re(s_j)} = q^{-\Re(s_j)} = \frac{1}{|\omega_j|} = \frac{1}{\sqrt{q}} = q^{-\frac{1}{2}}$$

and $\Re(s) = \frac{1}{2}$.

Hence, the RH for elliptic curves over $\mathbb{F}_q$.

We can rewrite RH in a way easier to generalize.

One has $|C(\mathbb{F}_q)| - q - 1 = -c_1$ with $c_1 = \omega + \overline{\omega} = 2\Re(\omega)$. But $\omega = \sqrt{q}e^{i\alpha}$ and therefore $\Re(\omega) = \sqrt{q}\cos(\alpha)$. So $c_1 = 2\sqrt{q}\cos(\alpha)$ and RH is equivalent to

$$||C(\mathbb{F}_q)| - q - 1| \leq 2q^{\frac{1}{2}}$$

Weil's "conceptual" proof of RH

To tackle the case $g > 1$, Weil had to take into account that C is no longer isomorphic to its Jacobian.

He worked on $\overline{\mathbb{F}_q}$ (to have a good intersection theory) and in the *square* $S = \overline{C} \times \overline{C}$ of the curve C extended to $\overline{\mathbb{F}_q}$.

For a description of Weil's proof, see e.g. James Milne's paper "*The Riemann Hypothesis over finite fields from Weil to the present day*" (2015).

He used the graph Φ_q of the Frobenius φ_q on $\overline{\mathbb{F}_q}$. It is a divisor of the surface $S = \overline{C} \times \overline{C}$.

As the $\mathbb{F}_q$-points of C, i.e. $C(\mathbb{F}_q)$, are the *fixed* points of φ_q, their number is the intersection number: $\Phi_q \bullet \Delta$ where Δ is the *diagonal* of $S = \overline{C} \times \overline{C}$.

Then Weil applied Hurwitz trace formula (1887), which implies that:

$$\begin{aligned} \Phi_q \bullet \Delta &= \Phi_q \bullet \xi_1 - \mathrm{Tr}\left(\varphi_q \mid H_1\left(\overline{C}\right)\right) + \Phi_q \bullet \xi_2 \\ &= 1 - \mathrm{Tr}\left(\varphi_q \mid H_1\left(\overline{C}\right)\right) + q \end{aligned}$$

with $\xi_1 = e_1 \times \overline{C}$ and $\xi_2 = \overline{C} \times e_2$ (e_j points of $\overline{C}$)

The key point is that, in this geometric context, RH for curves over $\mathbb{F}_q$ is equivalent to the *negativity condition* $D \bullet D \leq 0$ for all D of degree $= 0$.

This is equivalent to the *Castelnuovo-Severi inequality* for *every* divisor D :

$$D \bullet D \leq 2\,(D \bullet \xi_1)\,(D \bullet \xi_2)$$

If we apply this to the Frobenius divisor Φ_q when $\overline{C}$ has genus g we can deduce RH.

It is to prove Castelnuovo-Severi inequality that RR enters the stage with the inequality

$$\ell(D) - \ell(K_S - D) \geq \frac{1}{2} D \bullet (D - K_S) + \chi(S)$$

Connes' strategy : "a universal object for the localization of L functions"

To summarize: Weil introduced an intermediate world, the world of curves over finite fields $\mathbb{F}_q$. He reformulated the RH in this new framework and used tools inspired by algebraic geometry and cohomology over $\mathbb{C}$ to prove it.

It is well known that the generalization of this result to *higher dimensions* led to his celebrated conjectures and that the strategy for proving them has been at the origin of the monumental programme of Grothendieck (schemes, sites, topoï, etale cohomology).

But after Deligne's proof of Weil's conjectures in 1973 the original RH remained unbroken.

Some years ago, Alain Connes proposed a new strategy consisting in constructing a new *geometric* framework for arithmetics where Weil's proof could be transfered by analogy.

The fundamental discovery is that for finding a strategy, one needs to work in the world of "*tropical algebraic geometry in characteristic* 1", and apply it to the noncommutative space of the classes of adeles.

In his 2014 Lectures at the Collège de France he said that he was looking since 18 years for a geometric interpretation of adeles and ideles in terms of algebraic geometry *à la* Grothendieck.

In his essay he explains:

> *"It is highly desirable to find a geometric framework for the Riemann zeta function itself, in which the Hasse-Weil formula, the geometric interpretation of the explicit formulas, the Frobenius correspondences, the divisors, principal divisors, Riemann-Roch problem on the curve and the square of the curve all make sense. (p.8)"*

But this is another story.