Some analogies

Maksym Radziwill

June 6, 2018

Dictionary

L-functions	$\leftrightarrow$	"Sieves"
$\log\lvert\zeta(\frac{1}{2}+it)\rvert$		$w(n)=\sum_{p\mid n} 1$
$\lvert\zeta(\frac{1}{2}+it)\rvert^2$		$d(n)=\sum_{d\mid n} 1 \approx 2^{\omega(n)}$
$\lvert\zeta(\sigma+it)\rvert^2$		$\sigma_{1-2\sigma}(n)=\sum_{d\mid n} d^{1-2\sigma}$
$L(\frac{1}{2}, f\otimes\chi_{-n})$		$\lvert\lambda_f(n)\rvert^2$

First analogies

1. The first analogy appears in work of Selberg on the proportion of zeros of $\zeta(s)$ on the half-line.
2. Consider,

$$Z(t) := e^{-i\theta(t)}\zeta(\tfrac{1}{2} + it) \in \mathbb{R} \ , \ e^{2i\theta(t)} := \frac{\Gamma(\frac{\frac{1}{2}-it}{2})}{\Gamma(\frac{\frac{1}{2}+it}{2})}$$

   Counting sign changes of $Z(t)$ corresponds to counting zeros of $\zeta(s)$ on the half-line.
3. Hardy-Littlewood: Produce sign changes by finding disjoint intervals $I_1, I_2, \ldots, I_R$ with,

$$\Big| \int_{I_j} Z(u)du \Big| < \int_{I_j} |Z(u)|du.$$

First analogies

1. Up to height T there are $\asymp T \log T$ zeros, so to obtain a positive proportion we would need to find $\approx T \log T$ disjoint intervals $I_1, \ldots, I_R$ each of length $\approx (\log T)^{-1}$ on which (**??**) holds.
2. Unfortunately $Z(t)$ fluctuates quite wildly and is most of the times either small or large, so the mean-value of $Z(t)$ over an interval of length $(\log t)^{-1}$ is dominated by the peak of $Z(t)$, making the event

$$\Big| \int_I Z(u)du \Big| < \int_I |Z(u)|du \qquad (1)$$

   hard to detect analytically if $|I| \asymp (\log T)^{-1}$.
3. One needs to choose I longer so that at least several peaks occur and their signs cancel out. For this reason Hardy-Littlewood were only able to produce $\gg T$ zeros up to height T by taking each interval I of length about 1.

First analogies

1. To obtain a positive proportion of sign changes on the half-line one needs to find a way to dampen the size of $Z(t)$ so to diminish the importance of peaks, and allow for more sign cancellations.
2. For this purpose Selberg introduces a *mollifier*,

$$M(s) := \sum_{n \leq z} \frac{d_{1/2}(n)\mu(n)}{n^s} \cdot \Big(1 - \frac{\log n}{\log z}\Big).$$

3. This has the property that $|Z(t)M(\frac{1}{2} + it)^2| \approx 1$ for most t if $z > t^{\varepsilon}$.
4. Furthermore if z is not too large then,

$$\int_T^{2T} |Z(t)M(\tfrac{1}{2} + it)^2|^2 dt \asymp \int_T^{2T} |Z(t)M(\tfrac{1}{2} + it)^2|^4 dt \asymp T.$$

   are computable.
5. This allows one to conclude that there are $\asymp T \log T$ disjoint intervals $I_1, \ldots, I_R$ with

$$\Big| \int_{I_j} Z(u)|M(\tfrac{1}{2} + iu)|^2 du \Big| < \int_{I_j} |Z(u)| \cdot |M(\tfrac{1}{2} + iu)|^2 du$$

   and this gives $\asymp T \log T$ sign changes of $Z(t)$ hence also that many zeros on the critical line.

First analogies

1. The proportion was subsequently improved to $\frac{1}{3}$ by Levinson using some highly non-trivial function theoretic methods
2. The next improvement to 36% is due to Conrey who appealed among other to Weil's bounds for Kloosterman sums, and thus to the Riemann Hypothesis for Curves.
3. The further improvement to 40% is also due to Conrey, using the spectral theory of the hyperbolic Laplacian (through work of Deshouillers-Iwaniec), Vaughan's identity for the Moebius function, and using an improvement of Levinson's method. This is in my opinion one of the strongest results towards RH to date.
4. There have been further very technical improvements, using various innovative mollifiers, in particular in work of Bui. There are too many people who made small incremental improvements here, so unfortunately I cannot quote them all (Robles, Zaharescu, Pratt, Zeindler, ...)

First analogies

1. Soon after his work on zeros of $\zeta(s)$, proceeding by analogy Selberg developed the Selberg sieve. The idea is that for any $n > z$ and any coefficients λ_d with $\lambda_1 = 1$,

$$\mathbf{1}_{n \text{ is prime}} \leq \Big(\sum_{\substack{d \mid n \\ d \leq z}} \lambda_d \Big)^2$$

   The usefulness of this inequality that the analytic complexity of the RHS is low, where-as the analytic complexity of the LHS is high.
2. Thus it remains to find λ_d that will minimize the above inequality in an appropriate average sense. For instance, find λ_d that minimize,

$$\sum_{n \leq x} \Big(\sum_{\substack{d \mid n \\ d \leq z}} \lambda_d \Big)^2$$

3. This is very similar to the problem of finding λ_d that minimze,

$$\int_{T}^{2T} \Big| \zeta(\tfrac{1}{2} + it) \sum_{d \leq z} \frac{\lambda_d}{d^{\frac{1}{2} + it}} \Big|^2 dt.$$

4. In both cases the optimal choice is $\lambda_d = \mu(d) \cdot \Big(1 - \frac{\log d}{\log z} \Big)$.

First analogies

The Selberg sieve in its multidimensional incarnation had striking success with bounded gaps between primes, in works of Goldston-Pintz-Yildrim, Maynard, Tao and the Polymath team.

Second analogy

1. In modern work on L-functions the sieve returns in work of Soundararajan on moments of L-functions.
2. The objective is to obtain conditionally on RH, upper bounds for moments such as,

$$\int_T^{2T} |\zeta(\tfrac{1}{2}+it)|^{2k} dt \ll T(\log T)^{k^2}.$$

3. In order to achieve this Soundararajan shows that on RH, uniformly in $2 \leq X \leq T$ and $t \in [T, 2T]$,

$$\log|\zeta(\tfrac{1}{2}+it)| \leq \Re \sum_{p \leq X} \frac{1}{p^{\frac{1}{2}+it}} + \frac{\log T}{\log X}$$

4. The sieve theoretic analogue of this inequality is the elementary inequality for $n \in [N, 2N]$,

$$w(n) := \sum_{p|n} 1 \leq \sum_{\substack{p|n \\ p \leq y}} 1 + \frac{\log 2N}{\log y}$$

Second analogy

1. Both inequalities

$$\log|\zeta(\tfrac{1}{2}+it)| \leq \Re \sum_{p \leq X} \frac{1}{p^{1/2+it}} + \frac{\log T}{\log x} \ , \ w(n) \leq \sum_{\substack{p|n \\ p \leq y}} 1 + \frac{\log 2N}{\log y}$$

achieve the same aim: they allow for distributional control over $\log|\zeta|$, respectively $w(n)$. Again the RHS is of low analytic complexity where-as the LHS is difficult.

2. This then leads to upper bounds for moments, respectively,

$$\int_T^{2T} |\zeta(\tfrac{1}{2}+it)|^{2k} dt \ , \ \sum_{n \leq x} e^{\alpha w(n)}.$$

3. In the first incarnation one obtains sharp upper bounds for moments due to Soundararajan, and then refined by Harper, and in the second case one obtains sharp upper bounds for non-negative multiplicative functions, originally due to Shiu.

Third analogy

In recent work with Soundararajan we have been able to develop a third analogy, transporting the pure Brun sieve into the realms of L-functions. This allows to make progress on various distributional questions of which I will detail only one.

Third analogy

1. An important problem in the theory of L-functions is to establish non-vanishing at the central point. There are analytic and algebraic methods for this.
2. The algebraic methods typically rely on the algebraicity of the central values. For instance for f an integral weight cusp form, one uses Waldspurger in the form $L(\frac{1}{2}, f \otimes \chi_{-n}) \asymp |\lambda_g(n)|^2$ for some half-integral g, and then one establishes a congruence for λ_g that implies $\lambda_g(n) \neq 0$ (i.e Ono-Skinner, ...).
3. Alternatively for E an elliptic curve one could use that for any Galois automorphism σ we have $L(\frac{1}{2}, E \otimes \chi)^\sigma = L(\frac{1}{2}, E \otimes \chi^\sigma)$. Thus showing $L(\frac{1}{2}, E \otimes \chi) \neq 0$ for all χ in a given Galois orbit amounts to showing that there is at least one χ in each Galois orbit for which $L(\frac{1}{2}, E \otimes \chi) \neq 0$. If the Galois orbits are large, this can be easily achieved by considering a first moment of $L(\frac{1}{2}, E \otimes \chi)$ within the same Galois orbit (i.e Rohrlich, Chinta, ...).

Third analogy

1. The analytic methods for non-vanishing at the central point rely almost exclusively on the "second moment method" and mollification.
2. For instance,

$$\sum L(\tfrac{1}{2}, \chi_{-8d}) \leq \Big(\sum_{L(\frac{1}{2}, \chi_{-8d}) \neq 0} 1 \Big)^{1/2} \cdot \Big(\sum L(\tfrac{1}{2}, \chi_{-8d})^2 \Big)^{1/2}$$

3. Computing the two moments leads to

$$\sum_{\substack{8d \leq X \\ (d,2)=1 \\ d \text{ square-free} \\ L(\frac{1}{2}, \chi_{-8d}) \neq 0}} 1 \gg \frac{X}{\log X}.$$

Third analogy

1. The inefficiencies are again due to the fluctuation in size of $L(\frac{1}{2}, \chi_{-8d})$ and one can reduce this by introducing a mollifier,

$$M(\chi_{-8d}) := \sum_{n \leq z} \frac{\chi_{-8d}(n)\mu(n)}{\sqrt{n}} \cdot \Big(1 - \frac{\log n}{\log z}\Big).$$

2. The point of this mollifier is that for most d we have $L(\frac{1}{2}, \chi_{-8d})M(\chi_{-8d}) \asymp 1$.
3. Since,

$$\sum_{8d \leq X} L(\tfrac{1}{2}, \chi_{-8d})M(\chi_{-8d}) \asymp \sum_{8d \leq X} (L(\tfrac{1}{2}, \chi_{-8d})M(\chi_{-8d}))^2 \asymp X$$

applying Cauchy-Schwarz leads to a positive proportion of non-vanishing for $L(\frac{1}{2}, \chi_{-8d})$.

Third analogy

1. The algebraic and analytic method are mostly disjoint in their domain of applicability (one methods works very well when the other works poorly or not at all, and vice-versa).
2. One consistent advantage of the analytic method is that it not only produces non-vanshing but also shows that the non-zero values are not too small.
3. For instance the analytic method for non-vanishing at the central point not only shows that $L(\frac{1}{2}, \chi) \neq 0$ for $\frac{3}{8}$ of characters $\chi \pmod{q}$ (Iwaniec-Sarnak, Khan-Ngo), but also shows that $|L(\frac{1}{2}, \chi)| \gg (\log q)^{-10}$ for $\frac{3}{8}$ of characters $\chi \pmod{q}$.

Third analogy

This leads us to the main question:
Can one refine the mollifier method so that it also produces tight results in terms of the size of L-functions?

Conjecture (Keating-Snaith)

Let $\mathcal{F}$ be the set of odd, square-free integers. Then, as $x \to \infty$,

$$\frac{1}{|\mathcal{F} \cap [1,x]|} \# \left\{ d \in \mathcal{F} \cap [1,x] : \frac{\log |L(\frac{1}{2}, \chi_{-8d})|}{\sqrt{\log \log d}} \in (\alpha, \beta) \right\}$$

is in the limit equal to

$$\int_{\alpha}^{\beta} e^{-u^2/2} \cdot \frac{du}{\sqrt{2\pi}}.$$

This conjecture is thus our goal.

Third analogy

1. Sound mentioned some of our joint results in the lecture yesterday and I would like to explain the ideas behind the following result.

Theorem

Let $\mathcal{F}$ be the set of odd, square-free integers. Then, for any $\alpha < \beta$ as $x \to \infty$, the quantity,

$$\frac{1}{|\mathcal{F} \cap [1,x]|} \# \left\{ d \in \mathcal{F} \cap [1,x] : \frac{\log |L(\frac{1}{2}, \chi_{-8d})|}{\sqrt{\log \log d}} \in (\alpha, \beta) \right\}$$

is at least

$$> \left(\frac{7}{8} + o(1) \right) \int_{\alpha}^{\beta} e^{-u^2/2} \cdot \frac{du}{\sqrt{2\pi}}.$$

Third analogy

1. Our result is based on the introduction of a new mollifier based on the pure Brun sieve. So let us first explain what is the Brun and the pure Brun sieve.
2. The Brun sieve is a consequence of Bonferoni's inequality, accordingly it gives for $n \in [X, 2X]$,

$$\mathbf{1}_{n \text{ is prime}} \leq \sum_{\substack{d|n \\ p|d \implies p \leq Y_1 \\ \Omega(d) \leq W_1}} \mu(d). \tag{2}$$

   where W_1 is even.
3. It is important that the "analytic complexity" of the RHS remains low. To achieve this we need to choose Y_1 and W_1 so that $Y_1^{W_1} \leq X$.
4. Moreover for the sieve to be any good it needs to sift typical divisors. Thus we need $\Omega(d) > \log\log d$ and therefore $W_1 > \log\log X$.
5. So one is limited in the choice of Y_1 to $Y_1 \leq \exp(\log X / \log\log X)$. We choose for instance $Y_1 = \exp(\log X / (\log\log X)^2)$.

Third analogy

1. Unfortunately if one averages the resulting sieve,

$$\mathbf{1}_{n \text{ is prime}} \leq \sum_{\substack{d|n \\ p|d \implies p \leq Y_1 \\ \Omega(d) \leq W_1}} \mu(d)$$

then we obtain an upper bound for the number of primes $\leq X$ that is off by $\log\log X$, i.e

$$\ll \frac{X \log\log X}{\log X}$$

2. To correct this defect one can multiply the Brun sieve, by another Brun sieve that corrects the defects on the larger primes. One is faced with similar limitations as before, so we choose our second piece to be

$$h_2(n) := \sum_{\substack{d|n \\ p|d \implies Y_1 \leq p \leq Y_2 \\ \Omega(d) \leq W_2}} \mu(d)$$

where W_2 is even and the parameters choosen so that $Y_2^{W_2} \leq X$ and we need also that $W_2 > \log\log\log X$. Therefore $Y_2 = \exp(\log X/(\log\log\log X)^2)$ is a reasonable choice.

Third analogy

1. Thus consider,

$$\mathbf{1}_{n \text{ is prime}} \leq h_1(n)h_2(n)$$

where

$$h_i(n) := \sum_{\substack{d|n \\ Y_{i-1} \leq p \leq Y_i \\ \Omega(d) \leq W_i}} \mu(d)$$

with parameters choosen as before.

2. Averaging this establishes a bound for the number of primes which is now off by $\log\log\log X$. So we are doing better.
3. The optimal sieve is then constructed by iterating and writting,

$$\mathbf{1}_{n \text{ is prime}} \leq h_1(n)h_2(n)\cdots h_R(n)$$

for some R going to infinity very slowly.

Third analogy

1. We introduce a mollifier based on the pure Brun sieve. The main point is that we would like to construct a mollifier that provably behaves like the inverse of an Euler product.
2. Let $Y_0 = 1$, $Y_1 = X^{1/(\log\log X)^2}$, $Y_2 = X^{1/(\log\log\log X)^2}$, ...
3. Consider for each i,

$$M_i(s) := \sum_{\substack{p|n \implies Y_{i-1} \leq p \leq Y_i \\ \Omega(n) \leq 10\sqrt{\frac{\log X}{\log Y_i}}}} \frac{\mu(n)}{n^s} \ , \ \mathcal{Q}_i(s) := \sum_{Y_{i-1} \leq p \leq Y_i} p^{-s}.$$

   Then, whenever $|\mathcal{Q}_i(\frac{1}{2} + it)| \leq 10\sqrt{\frac{\log X}{\log Y_i}}$ we have by a Taylor expansion,

$$M_i(s) \approx e^{-\mathcal{Q}_i(s)}.$$

4. Importantly $|\mathcal{Q}_i| \leq 10\sqrt{\frac{\log X}{\log Y_i}}$ is a *typical* event.
5. Our Brun sieve mollifier is $M(s) := M_1(s)M_2(s) \ldots\ldots M_R(s)$ where R is choosen so that the total length of $M(s)$ is in $[X^{\varepsilon^{10}}, X^{\varepsilon}]$

Third analogy

1. Our strategy is as follows : Using the fact that M is a mollifier, we show that for $\frac{7}{8} + O(\varepsilon)$ of quadratic characters we have,

$$\varepsilon < |L(\tfrac{1}{2}, \chi_{-8d})M(\chi_{-8d})| \leq \frac{1}{\varepsilon}$$

   for any fixed $\varepsilon > 0$.

2. However for almost all χ_{-8d} with $d \in [X, 2X]$ we also have by construction

$$M(\chi_{-8d}) \approx e^{-\mathcal{Q}_1(\chi_{-8d}) - \mathcal{Q}_2(\chi_{-8d}) + O(\log\log\log X)}$$

3. So it follows that for $\frac{7}{8}$ of characters χ_{-8d} with $d \in [X, 2X]$, we also have,

$$\log|L(\tfrac{1}{2}, \chi_{-8d})| \approx \sum_{p \leq X^{1/(\log\log X)^2}} \frac{\chi_{-8d}(p)}{\sqrt{p}} + O(\log\log\log X).$$

   It thus remains to run through this argument by also imposing the condition that the sum over primes (i.e $\mathcal{Q}_1 + \mathcal{Q}_2$) behave in an appropriate Gaussian way. This can be achieved with moments.

Third analogy

Rigorously the argument runs as follows.

1. Consider the first moment,

$$\sum_{\frac{(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})-\frac{1}{2}\log\log d}{\sqrt{\log\log d}}\in(\alpha,\beta)} L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})$$

2. Notice that "for free" at any time, we can add the condition that $\mathcal{Q}_1, \mathcal{Q}_2$ are not too large, and that $\varepsilon < |L(\frac{1}{2},\chi_{-8d})M(\chi_{-8d})| \leq \frac{1}{\varepsilon}$.
3. The condition that

$$\frac{(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})-\frac{1}{2}\log\log d}{\sqrt{\log\log d}} \in (\alpha,\beta)$$

   can be controlled through moments, by considering,

$$\sum L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})\Big(\frac{(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})-\frac{1}{2}\log\log d}{\sqrt{\log\log d}}\Big)^{2k}$$

4. Notice that $\mathcal{Q}_1, \mathcal{Q}_2$ is supported on small primes, where-as $L(\chi_{-8d})M(\chi_{-8d})$ is supported on large primes. So we expect independence.

Third analogy

1. In any case we expect that the first moment

$$\sum_{\frac{(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})-\frac{1}{2}\log\log d}{\sqrt{\log\log d}}\in(\alpha,\beta)} L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})$$

$$\approx \int_{\alpha}^{\beta} e^{-u^2/2}\cdot\frac{du}{\sqrt{2\pi}}\sum L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})$$

2. And a similar computation also works for the second moment.
3. Finally recall that we can always add for free in the first moment, the conditions that,

$$\varepsilon < |L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})| \leq \frac{1}{\varepsilon}$$

   and that $|\mathcal{Q}_1|, |\mathcal{Q}_2|$ are not large so that, also

$$M(\chi_{-8d}) \approx e^{-\mathcal{Q}_1(\chi_{-8d})-\mathcal{Q}_2(\chi_{-8d})+O(\log\log\log X)}.$$

Third analogy

1. Therefore applying Cauchy-Schwarz we end up with,

$$\int_\alpha^\beta e^{-u^2/2}\cdot\frac{du}{\sqrt{2\pi}}\sum L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})$$

$$\leq\Bigg(\sum_{\substack{\frac{(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})-\frac{1}{2}\log\log d}{\sqrt{\log\log d}}\in(\alpha,\beta)\\ \varepsilon<|L(\frac{1}{2},\chi_{-8d})M(\chi_{-8d})|\leq\frac{1}{\varepsilon}\\ M(\chi_{-8d})\approx e^{-(\mathcal{Q}_1+\mathcal{Q}_2)(\chi_{-8d})}}} 1\Bigg)^{1/2}$$

$$\times\Bigg(\int_\alpha^\beta e^{-u^2/2}\cdot\frac{du}{\sqrt{2\pi}}\sum\big(L(\tfrac{1}{2},\chi_{-8d})M(\chi_{-8d})\big)^2\Bigg)^{1/2}$$

2. Upon dividing this gives the desidered LHS. On the other hand we can re-arrange the RHS into an lower bound for exactly what we want, that is

$$\#\left\{d\in\mathcal{F}\cap[1,x]:\frac{\log|L(\frac{1}{2},\chi_{-8d})|-\frac{1}{2}\log\log d}{\sqrt{\log\log d}}\in(\alpha,\beta)\right\}$$

Third analogy

Consequently one gets the corollary that not only $\frac{7}{8}$ of the χ_{-8d} are such that $L(\frac{1}{2}, \chi_{-8d}) \neq 0$ but in fact $\frac{7}{8}$ of the χ_{-8d} are such that,

$$L(\tfrac{1}{2}, \chi_{-8d}) \asymp (\log d)^{\frac{1}{2}+o(1)}.$$

In some sense it is no surprise that we get a good proportion of non-vanishing in the case of $L(\frac{1}{2}, \chi_{-8d})$ because these values are typically large, so one would expect it should be easier to show that they are non-zero!!
The proof on RH is rather similar but easier using the explicit formula and conditioning it.

Third analogy

The final technical remark is that our actual mollifier is a hybrid Brun pure sieve / Selberg mollifier, and takes the shape of

$$M(s) = M_0(s) \sum_{\substack{p|n \implies p>X^{\varepsilon} \\ n<\sqrt{X}}} \frac{\mu(n) w_n \chi_{-8d}(n)}{\sqrt{n}}$$

where M_0 is the Brun pure sieve mollifier supported on $n \leq X^{\varepsilon}$ and the weights w_n are the weights of the optimal Selberg mollifier that minimizes the ratio of

$$\frac{\sum (L(\frac{1}{2}, \chi_{-8d}) \sum \mu(n) w_n \chi_{-8d}(n) n^{-1/2})^2}{\sum L(\frac{1}{2}, \chi_{-8d}) \sum \mu(n) w_n \chi_{-8d}(n) n^{-1/2}}$$

A final nice analogy

A final nice analogy that I didn't have the opportunity to mention is the following approximate functional equation,

$$\zeta(\sigma+it) = \sum_{n\leq\sqrt{t/2\pi}} \frac{1}{n^{\sigma+it}} + \chi(\sigma+it) \sum_{n\leq\sqrt{t/2\pi}} \frac{1}{n^{1-\sigma-it}} \ , \ \chi(s) = \frac{\Gamma(1-s)}{\Gamma(s)}$$

"Sieve" version:

$$\sigma_{1-2\sigma}(n) = \sum_{\ell} \frac{c_\ell(n)}{n^{2\sigma}} f_{1-2\sigma}\left(\frac{\ell}{\sqrt{n}}\right) + n^{1-2\sigma} \sum_{\ell} \frac{c_\ell(n)}{n^{2-2\sigma}} f_{2\sigma-1}\left(\frac{\ell}{\sqrt{n}}\right)$$

where $f_\sigma(x)$ are some smooth functions concentrated in $x \ll 1$ and $c_\ell(n)$ are Ramanujan sums.