Letter to: Barry Mazur on "Chebyshev's bias" for $\tau(p)$

November, 2007

from

Peter Sarnak

Dear Barry,

This concerns figures 1.6, 2.2, 2.3, 2.4 and 2.5 in your paper [1], that is the fluctuations of $\#\{p \le x : \tau(p) > 0\} - \#p\{\le x : \tau(p) < 0\}$. Here $\tau(p)$ are the Fourier coefficients of Ramanujan's form $\triangle(z)$ and p always denotes a prime. Let

$$\lambda(p) = \tau(p)/p^{11/2} = \alpha_p + \beta_p, \tag{1}$$

with

$$\alpha_p = e^{i\theta_p},\, \beta_p = e^{-i\theta_p}$$

and

$$\theta_p \in [0, \pi]\,. \tag{2}$$

For $n \ge 1$ the symmetric power L-function for $\pi = \triangle$ is given by

$$L(s, \pi, \mathsf{sym}^n) := \prod_p \prod_{j=0}^{n} (1 - \alpha_p^{n-j}\, \beta_p^j\, p^{-s})^{-1}\,. \tag{3}$$

Hence

$$-\tfrac{L'}{L}\,(s, \pi, \mathsf{sym}^n) =$$

$$\sum_p \frac{\log p}{p^s}\left(\sum_{j=0}^{n} \alpha_p^{n-j}\, \beta_p^j\right) + \sum_p \frac{\log p}{p^{2s}}\left(\sum_{j=0}^{n} \alpha_p^{2n-2j}\, \beta_j^{2j}\right) + \cdots \tag{4}$$

$$= \sum_p \frac{\log p}{p^s}\, U_n(\theta_p) + \sum_p \frac{\log p}{p^{2s}}\, U_n(2\theta_p) + \cdots \tag{5}$$

where

$$U_n(\theta) = \frac{\sin(n+1)\theta}{\sin\theta}\,. \tag{6}$$

Hence by Riemann-Von-Mangolt;

$$\begin{aligned} &\sum_{p \le x} \log p\, U_n(\theta_p) + \sum_{p^2 \le x} \log p\, U_n(2\theta_p) \\ &= -\sum_{\rho_{j,n}} \frac{x^{\rho_{j,n}}}{\rho_{j,n}} + O_\epsilon\left(x^{1/3+\epsilon}\right)\,, \end{aligned} \tag{7}$$

where $\rho_{j,n}$ range over the nontrivial zeros of $L(s, \pi, \mathsf{sym}^n)$. We assume further that

(a) $L(s, \pi, \mathsf{sym}^n)$ extends to an entire function in s and satisfies an expected functional equation (see below). Actually this and some parts of what we claim below would follow from the meromorphicity (such as is known by Taylor if $\pi = \pi_E$ corresponds to an elliptic curve over $\mathbb{Q}$), in that case the sum is also over the possible poles of $L(s, \pi, \mathsf{sym}^n)$.

(b) $L(s,\pi,\mathsf{sym}^n)$ satisfies the Riemann Hypothesis so that $\rho_{j,n} = \frac{1}{2} + i\,\gamma_{j,n}$ with $\gamma_{j,n} \in \mathbb{R}$.

First we examine the second term in (7) which is one source of the "bias". For $V(\theta)$ a smooth function on $[0,\pi]$ consider

$$\sum_{p\le x} \log p\, V(\theta_p)\,. \tag{8}$$

We have that

$$\sum_{p\le x} \log p\, V(\theta_p) \;=\; \sum_{p\le x} \log p \sum_{n=0}^{\infty} \langle V, U_n\rangle\, U_n(\theta_p)\,. \tag{9}$$

Here

$$\langle V_1, V_2\rangle \;=\; \frac{2}{\pi}\int_0^{\pi} V_1(\theta)V_2(\theta)\, \sin^2\theta d\theta \tag{10}$$

and note that $U_n, n = 0,1,2,\ldots$ are an orthonomal basis for $L^2[0,\pi]$ with this inner product. Hence

$$\sum_{p\le x} \log p\, V(\theta_p) \;=\; \sum_{n=0}^{\infty} \langle V, U_n\rangle \sum_{p\le x} \log p\, U_n(\theta_p)$$

and by our assumption that $L(s,\pi,\mathsf{sym}^n)$ has no pole at $s = 1$ for $n \ge 1$, we deduce that

$$\sum_{p\le x} \log p\, V(\theta_p) \;\sim\; \langle V, V_0\rangle x\,, \quad \text{as} \quad x \to \infty\,. \tag{11}$$

We apply this to $V(\theta) = U_m(2\theta)$ for which

$$\langle V, U_0\rangle \;=\; \frac{2}{\pi}\int_0^{\pi} U_m(2\theta)\, \sin^2\theta d\theta \;=\; (-1)^m\,. \tag{12}$$

Hence from (7) we have

$$\sum_{p\le x} \log p\, U_n(\theta_p) \;=\; (-1)^{n+1}\, x^{1/2} \; - \sum_{\rho_{j,n}} \frac{x^{\rho_{j,n}}}{\rho_{j,n}} \;+\; \text{small.} \tag{13}$$

We separate out the zeros $\rho_{j,n} = \frac{1}{2}$ which occur with multiplicity denoted by $M_n\left(\frac{1}{2}\right)$, to get

$$\sum_{p\le x} \log p\, U_n(\theta_p) \;=\; \left(-2M_n\left(\frac{1}{2}\right) \;+\; (-1)^{n+1}\right)\, x^{1/2} \;-\; \sum_{\gamma_{j,n}\ne 0} \frac{x^{\rho_{j,n}}}{\rho_{j,n}}\,. \tag{14}$$

Again if V is a smooth function on $[0,\pi]$ for which $\langle V, U_0\rangle = 0$, we have

$$V(\theta) \;=\; \sum_{n=1}^{\infty} A_n\, U_n(\theta) \tag{15}$$

$$A_n \;=\; \langle V, U_n \rangle \tag{16}$$

as a rapidly convergent expansion. Then

$$\frac{1}{\sqrt{x}} \sum_{p \le x} \log p \, V(\theta_p) \;=\; \sum_{n=1}^{\infty} A_n \left(-2 M_n \left(\frac{1}{2} \right) \;+\; (-1)^{n+1} \right)$$
$$- \sum_{n=1}^{\infty} A_n \sum_{\gamma_{j,n} \neq 0} \frac{x^{i\gamma_{j,n}}}{\frac{1}{2} \;+\; i\gamma_{j,n}} \tag{17}$$

It follows as in the analysis in [2] that

$$S_V(x) \;:=\; \frac{\log x}{\sqrt{x}} \sum_{p \le x} V(\theta_p) \tag{18}$$

has a limiting distribution μ_V, w.r.t. dx/x. That is, for $f \in C(\mathbb{R})$ and bounded

$$\frac{1}{\log X} \int_2^X f(S_V(x)) \, \frac{dx}{x} \;\to\; \int_{\mathbb{R}} f(x) d_{\mu_V}(x) \,, \quad \text{as} \quad X \to \infty \,. \tag{19}$$

The measure μ_V contains all the information about the fluctuations of S_V and of any bias. The mean $E(S_V)$ of μ_V is given by

$$E(S_V) \;=\; \sum_{n=1}^{\infty} A_n \left(-2 M_n \left(\frac{1}{2} \right) \;+\; (-1)^{n+1} \right) \,. \tag{20}$$

To examine μ_V further we make the working hypothesis:

(c) The numbers $\gamma_{n,j} \;>\; 0$, $n \;\geq\; 1$ and $j \;\geq\; 1$ are linearly independent over $\mathbb{Q}$.

While we have no means of verifying such an assumption, computing a large number of these $\gamma_{n,j} > 0$ will lead to a verification (under (a) and (b)) to any order of accuracy, of the following conclusion:

Assuming (a), (b) and (c) the Fourier transform of μ_V is given by

$$\hat{\mu}_V(\xi) \;=\; e^{-i \, E(S_V) \xi} \prod_{n=0}^{\infty} \prod_{\gamma_{j,n} > 0} J_0 \left(\frac{2 |A_n| \xi}{\sqrt{\frac{1}{4} \;+\; \gamma_{n,j}^2}} \right) \tag{21}$$

where

$$J_0(z) \;=\; \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{2m}}{(m!)} \,. \tag{22}$$

The product in (21) converges absolutely and one concludes that the measure μ_V is symmetric about its mean $E(S_V)$, it is smooth and has support all of $(-\infty, \infty)$. Moreover, its variance is given by

$$V(S_V) \;=\; \sum_{n=1}^{\infty} \sum_{\gamma_{j,n} > 0} \frac{|A_n|^2}{\frac{1}{4} \,+\, \gamma_{j,n}^2}\,. \tag{23}$$

Applying this to $V(\theta) \;=\; U$, yields a limiting distribution, μ_1 for

$$S_1(x) \;:=\; \frac{\log x}{\sqrt{x}} \sum_{p \leq x} \lambda_p\,. \tag{24}$$

Now $m_1\, \left(\frac{1}{2}\right) \;=\; 0$ so

$$E(S_1) \;=\; 1\,. \tag{25}$$

Thus the weighted sum over $\tau(p)/p^{11/2}$ has a definite bias to being positive.

One can apply the above analysis to $\pi \;=\; \pi_E$, the automorphic cusp form of weight 2 corresponding to an elliptic curve $E/\mathbb{Q}$. If $S_{1,E}(x)$ are the sums

$$S_{1,E}(x) \;=\; \frac{\log x}{\sqrt{x}} \sum_{p \leq x} \frac{a_E(p)}{\sqrt{p}}\,, \tag{26}$$

then $S_{1,E}(x)$ has a limiting distribution $\mu_{1,E}$ with mean

$$E(S_{1,E}) \;=\; -2r(E) \,+\, 1 \tag{27}$$

where $r(E)$ is the rank of E, and variance

$$V(S_{1,E}) \;=\; \sum_{\gamma_{j,E} > 0} \frac{1}{\frac{1}{4} \,+\, \gamma_{j,E}^2}\,. \tag{28}$$

Here $\frac{1}{2} + i\gamma_{j,E}$ are the nontrivial zeros of $L(S, \pi_E)$. So if $r(E) > 0$ then $S_{1,E}$ has a definite bias to being negative, while if $r(E) = 0$ it has a bias to being positive. If $N(E)$ is the conductor of E then standard winding number arguments show that

$$V(S_{1,E}) \;\sim\; c\, \log\, N(E) \tag{29}$$

for a fixed positive constant c. So while most of us believe that $r(E)$ can be made as large as we like by varying E, this rank will be much smaller than $\log N(E)$ as $N(E) \to \infty$. That is to say that while for any E for which $r(E) > 0$, there is a bias for $S_{1,E}(x)$ to be negative precisely $\delta\left(\{x :\, S_{1,E}(x) \,<\, 0\}\right) \,>\, \frac{1}{2}$. where δ is the logarithmic density of the set. This bias will dissolve (i.e. $\delta \to \frac{1}{2}$) as $N(E) \to \infty$.

We return to $\lambda(p) = \tau(p)/p^{11/2}$ and examine your sums in Figure 1.6. Let $H(\theta)$ be the Heavy side function

$$H(\theta) = \begin{cases} 1 & \text{if} \quad \theta \in [0, \pi/2) \\ -1 & \text{if} \quad \theta \in (\frac{\pi}{2}, \pi] \end{cases} \tag{30}$$

Then

$$\sum_{p \le x} H(\theta_p) = \sum_{\substack{p \le x \\ \tau(p) > 0}} 1 - \sum_{\substack{p \le x \\ \tau(p) < 0}} 1. \tag{31}$$

For $n > 0$,

$$\begin{aligned} A_n(H) &= \langle H, U_n \rangle \\ &= \frac{1}{\pi} \left[\int_0^{\pi/2} U_n(\theta) \sin^{+2} \theta \, d\theta - \int_{\pi/2}^{\pi} U_n(\theta) \sin^{-2} \theta \, d\theta \right] \\ &= \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{(-1)^{(n-1)/2} 2}{\pi} \left[\frac{1}{n} + \frac{1}{n+2} \right] & \text{if } n \text{ is odd}. \end{cases} \end{aligned} \tag{32}$$

As far as $M_n\left(\frac{1}{2}\right)$, we know that $M_1\left(\frac{1}{2}\right) = 0$ and we extend our working hypothesis for this example of $\pi = \triangle$, to;

$$M_n\left(\frac{1}{2}\right) = \frac{1 - \epsilon\left(\frac{1}{2}, \pi, \mathsf{sym}^n\right)}{2} \tag{33}$$

where $\epsilon\left(\frac{1}{2}, \pi, \mathsf{sym}^n\right)$ is the global root number of $L(S, \pi, \mathsf{sym}^n)$. For this case of $\pi = \triangle$ which is unramified at all finite places and is a discrete series of weight $k = 12$ at infinity, this root number can be computed easily (Shahidi) and equals

$$\epsilon\left(\frac{1}{2}, \pi, \mathsf{sym}^n\right) = \begin{cases} 1 & \text{if } n \equiv 0, 1, 2, 4, 6, 7 \mod 8 \\ -1 & \text{if } n \equiv 3, 5 \mod 8. \end{cases} \tag{34}$$

So for n odd

$$M_n\left(\frac{1}{2}\right) = \begin{cases} 0 & \text{if } n \equiv 1, -1 \mod 8 \\ 1 & \text{if } n \equiv 3, 5 \mod 8 \end{cases} \tag{35}$$

For $N \ge 1$ let

$$V_N(\theta) = \sum_{n=1}^{N} A_n(H) \, U_n(\theta). \tag{36}$$

So V_N is a smoothed out version of $H(\theta)$ and $V_N(\theta) \to H(\theta)$ as $N \to \infty$.

$$S_N(x) = \frac{\log x}{\sqrt{x}} \sum_{p \le x} V_N(\theta_p). \tag{37}$$

is a smooth version of what we want, viz

$$S(x) := \frac{\log x}{\sqrt{x}} \sum_{p \le x} H(\theta_p) \,. \tag{38}$$

Applying (18), (19), (20), (21) to V_N yields that $S_N(x)$ has a limiting distribution μ_N with mean

$$E(S_N) = \frac{2}{\pi} \sum_{n\,\mathsf{odd}}^{N} (-1)^{\frac{n-1}{2}} \left[\frac{1}{n} + \frac{1}{n+2}\right] (1 - 2\, I_{3,5}(n)) \tag{39}$$

and variance

$$V(S_N) = \sum_{n\,\mathsf{odd}}^{N} \frac{4}{\pi^2} \left(\frac{1}{n} + \frac{1}{n+2}\right)^2 \sum_{\gamma_{j,n} > 0} \frac{1}{\frac{1}{4} + \gamma_{j,n}^2} \,. \tag{40}$$

Here

$$I_{3,5}(n) = \begin{cases} 1 & \mathsf{if}\ n \equiv 3, 5(8) \\ 0 & \mathsf{if}\ n \equiv 1, 7(8) \end{cases}$$

Let $N \to \infty$ in (39) and summing the infinite series one finds that the bias tends to

$$E(S_N) \to E(S) = \frac{2}{\pi} \left(1 + 2 \int_0^1 \frac{x^2}{1 + x^4}\, dx\right) . \tag{41}$$

This then should be the mean in your graph in Figure 1.6 (after scaling by $\log x/\sqrt{x}$). It agrees quite well.

However, as $N \to \infty$ there is a new feature with the variance which has to do with $L(s, \pi, \mathsf{sym}^n)$ being the L-function of an automorphic form on GL_{n+1}! An analysis of the zeros of height say less than 100 with n large using the method in [3] (after computing the archimedian factors $L(S, \pi_\infty, \mathsf{sym}^n)$) shows that

$$\#\{0 \le \gamma_{j,n} \le 100\} \ge c_1 n \tag{42}$$

(here $c_1 > 0$ is fixed). This will suffice for our purposes but in fact one can give a lower bound of $cn \log n$ in (42)$^{(*)}$. It follows that there is $c_2 > 0$ s.t.

$$V(S_N) \ge c_2 \sum_{\substack{n\,\mathsf{odd} \\ n \le N}} \frac{n}{n^2} \gg \log N \,. \tag{43}$$

This indicates that for S in (38)

$$V(S) = \infty \,.$$

That is for your fluctuations in (1.6), while the mean is positive the fluctuations are big enough to dissolve the bias. Put another way $\delta(\{x : S(x) > 0\}) = \frac{1}{2}$. The same applies to any elliptic curve $E/\mathbb{Q}$ and to your graphs 2.2 to 2.5, as long as the curve is not CM. In these cases the variance will be infinite and wipe out the bias from the mean which can be computed in terms of the rank of E and the Epsilon factors in $L(s, \pi_E, \mathsf{sym}^n)$. For a CM elliptic curve the number of zeros of the symmetric power L-functions up to a fixed height grows only like $\log n$. Hence the variance will remain bounded and a definite bias will remain (and can be computed) for $S(x)$.

Best regards,

Peter Sarnak

1. B. Mazur, "Finding meaning in error terms," *BAMS* **45**, 185-228, (2009).
2. M. Rubinstein & P. Sarnak, "Experimental Math.", Vol. 3, 173-197, (1994).
3. S. D. Miller, *Duke Math. Journal*, Vol. 112, 83-116, (2000).

August 12, 2008: PS/gpp.

Added - January, 2008

Here is a more detailed statement about the density of zeroes of $L(s, \triangle, \mathsf{sym}^n)$ in $(42)^*$ above. The exact archimedean factor of $L(s, \triangle, \mathsf{sym}^n)$ is (see [G] for example) given by

$$L(s, \triangle_\infty, \mathsf{sym}^n) \;=\; \prod_{0 \le j < \frac{n-1}{2}} \Gamma_{\mathbb{C}} \left(s + \frac{(n-2j)+1}{2} \right) \quad \text{if } n \text{ is odd} \tag{i}$$

and

$$= \;\Gamma_{\mathbb{R}}(s + 12) \prod_{0 \le j \le \frac{n}{2}} \Gamma_{\mathbb{C}}\,(s + 11j) \quad \text{if } n \text{ is even.}$$

Hence the "analytic conductor" $c(t, \triangle, \mathsf{sym}^n)$ in [I-S] see also [I-K], satisfies

$$\log c(t, \triangle, \mathsf{sym}^n) \;\sim\; n\, \log(n(2+|t|)) \tag{ii}$$

for $t \in \mathbb{R}$, $n \ge 1$.

The analytic conductors controls the density of zeroes of $L(s, \triangle, \mathsf{sym}^n)$ as follows:

Proposition *For $t \in \mathbb{R}$, $n \ge 1$, $\frac{100}{\log n} \le \epsilon \le 1$ we have*

$$\epsilon \log c(t, \triangle, \mathsf{sym}^n) \;\ll\; \sum_{|\gamma_{j,n} - t| \le \epsilon} 1 \;\ll\; \epsilon \log(c(t, \triangle, \mathsf{sym}^n))\,,$$

the implied constant is universal.

This follows using positivity with a carefully chosen test function in the explicit formula (Theorem 5.12, page 109 of [I-K]). $(42)^*$ follows immediately from the Proposition as does the fact that the zeroes of the individual $L(s, \triangle, \mathsf{sym}^n)$ becoming dense in $\Re(s) = \frac{1}{2}$, as $n \to \infty$. Note that the latter is not true for any sequence π_n of automorphic forms on GL_n as $n \to \infty$, as the example of $(\zeta(s))^n$ shows. However, if the π_n's are cuspidal then I would bet it is true for the standard L-functions $L(s, \pi_n)$.

In connection with these oscillatory sums, (31) has the largest fluctuations while (26) is smoothed out but still has large fluctuations. One can consider the further smoothed sums (for say an elliptic curve)

$$G(x) \;=\; \sum_{p \le x} \frac{a_E(p)}{p} \tag{iii}$$

and

$$F(x) \;=\; \sum_{p \le x} \frac{a_E(p)\, \log p}{p} \tag{iv}$$

As we show below these are of historical interest.

Now

$$F(x) \;=\; \frac{1}{2\pi i} \int\limits_{\Re(s)=1} \frac{L'}{L}\left(s + \frac{1}{2}, E\right) x^s \,\frac{ds}{s} \;+\; c_1(E) \;+\; o(1) \tag{v}$$

as $x \to \infty$. Here $c_1(E)$ is a constant. Shifting the contour to the left of $\Re(s) = 0$ yields

$$F(x) \;=\; r \log x \;+\; c_2(E) \;+\; \sum_{\rho \neq \frac{1}{2}} \frac{x^{i\gamma}}{i\gamma} \;+\; o(1) \tag{vi}$$

where $\rho = \frac{1}{2} + i\gamma$ runs of the zeroes of $L(s, E)$ and r is the order of vanishing of $L(s, E)$ at $s = 1/2$.

Thus, for $F(x)$ the secondary term still oscillates unboundly. However, for $G(x)$ the behavior is probably different.

$$\begin{aligned} G(x) \;&=\; \int_2^x \frac{dF(t)}{\log t} \\ &=\; \int_2^x \frac{1}{\log t}\,\frac{r\,dt}{t} \;+\; \int_2^x \frac{dS(t)}{\log t} \end{aligned}$$

where

$$S(x) \;=\; \sum_{\rho \neq \frac{1}{2}} \frac{x^{i\gamma}}{i\gamma}\,. \tag{vii}$$

Hence

$$\begin{aligned} G(x) \;&=\; r\,\log\log x \;+\; c_3(E) \;+\; \frac{S(x)}{\log x} \\ &\quad - \int_2^x S(t)\,\frac{d}{dt}\left(\frac{1}{\log t}\right) dt \qquad \text{(viii)} \\ &=\; r \log\log x \;+\; c_3(E) \;+\; \frac{S(x)}{\log x} \\ &\quad - \sum_{\rho \neq \frac{1}{2}} \frac{1}{i\gamma} \int_2^\infty \frac{e^{i\gamma y}}{y^2}\,dy \;+\; o(1) \qquad \text{(ix)} \\ &=\; r \log\log x \;+\; c_4(E) \;+\; \frac{S(x)}{\log x} \;+\; o(1) \qquad \text{(x)} \end{aligned}$$

Thus

$$G(x) \;=\; r \log\log x \;+\; c_4(E) \;+\; o(1)$$

iff

$$S(x) \,=\, o(\log x) \tag{xi}$$

It is hard to imagine that the second condition in (xi) is not true since this would require a remarkable uniform lining up of the phases $\gamma \log x$. Hence for $G(x)$ I (and I think most others) would expect that the oscillations tend to zero with x.

The historal significance of this is its connection with the original two Conjectures of Birch and Swinnerton-Dyer [B-SD]. In this first paper they make two conjectures

Conjecture 1: *If $N_p(E)$ is the number of points in E over $\mathbb{F}_p$ then as $x \to \infty$*

$$\prod_{p \le x} \frac{N_p}{p} \,\sim\, c(E)(\log x)^r$$

$r \,=\, \mathsf{rank}\,(E/\mathbb{Q}$.

Conjecture 2: *The usual* B-SD *conjecture in terms of the order of vanishing of $L(s,E)$ at $s = 1/2$ being the rank.*

Conjecture 1 was the starting point for their numerical investigation as it is the analogue for an elliptic curve of Siegel's mass formula. Today, Conjecture 2 is the important one in terms of what is proven. Are these Conjectures equivalent (assuming GRH as we are)? The answer is yes iff the second part of (xi) holds. Indeed Conjecture 1 is, on taking logs equivalent to

$$\sum_{p \le x} \frac{a_p}{p} \,=\, r \log\log x \,+\, c_5(E) \,+\, o(1)\,, \quad \text{as} \;\; x \to \infty\,.$$

The equivalence of Conjectures 1 and 2 under a hypothesis equivalent to $S(x) = o(\log x)$, was recently established in [Co] and [K-M].

References

[G] A. Guloglu, *IMRN*, 2005, **9**, 517-550.

[I-S] H. Iwainec & P. Sarnak, *GAFA*, 2000, 705-741.

[I-K] H. Iwainec & E. Kowalski, "Analytic Number Theory," *Coll. AMS*, Vol. 53, 2004.

[B-SD] B. Birch & Swinnerton-Dyer *Journal Reine Angew Math.*, **218**, 1965, 79-108.

[Co] K. Conrad, *Canadian Journal*, Vol. 57, 2005, 267-297.

[K-M] W. Kuo and R. Murty, *Canadian Journal Math.*, **57**, 2005, 328-337.

August 12, 2008: PS/gpp.