Statistics of zeros of zeta functions of curves over a finite field

Zeev Rudnick, Tel Aviv University
Joint projects with D. Faifman (TAU) & P. Kurlberg (KTH)

Number theory over $F_q[x]$

Motivation – closer analogy between arithmetic and RMT

$\mathbf{F}_q$=finite field with q elements, e.g. $\mathbf{F}_3=\mathbf{Z}/3\mathbf{Z}$.

$\mathbf{F}_q[x]$ = polynomials $f(x) = a_0+a_1x+a_2x^2+\ldots+a_dx^d$, with coefficients a_i in $\mathbf{F}_q$

$\deg(f)=d$ if $a_d \neq 0$; set $\|f\|:=q^{\deg(f)}$

A polynomial f(x) in $F_q[x]$ is **reducible** if we can write $f(x)=a(x)b(x)$

with $\deg(a)>0$ and $\deg(b)>0$, **irreducible** otherwise

e.g. $P(x)= x^2+1$ is irreducible over $\mathbf{F}_3=\mathbf{Z}/3\mathbf{Z}$.

Extension fields: Can get an extension field $\mathbf{E}$ of $\mathbf{F}_q$, of degree $n=\dim[\mathbf{E}:\mathbf{F}_q]$ by adjoining to $\mathbf{F}_q$ a root of an irreducible polynomial of degree n.

e.g. adjoin a root of x^2+1 to $\mathbf{F}_3=\mathbf{Z}/3\mathbf{Z}$ gives field $\mathbf{F}_9=\mathbf{F}_3[\sqrt{-1}]$ of $3^2=9$ elements.

Analogy: $\mathbf{F}_q[x]$ vs. integers

integers $\mathbf{Z}$ ↔ polynomials $\mathbf{F}_q[x]$

primes p ↔ irreducible polynomial P(x) ("prime")

positive p>0 ↔ monic polynomial $P(x)=x^d+\ldots$

Counting prime polynomials :

$$\pi_q(n) := \#\{P \quad \text{monic irreducible}, \quad \deg(P) = n\}$$

Prime Polynomial Theorem :

$$\pi_q(n) = \frac{q^n}{n} + O(q^{n/2})$$

The zeta function for $F_q[x]$

$$\zeta_q(s) := \prod_{\substack{P \text{ monic} \\ \text{irreducible}}} \left(1 - \| P \|^{-s}\right)^{-1}, \quad \mathrm{Re}(s) > 1$$

Norm $\|f\| := q^{\deg(f)}$

Riemann ζ- function

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

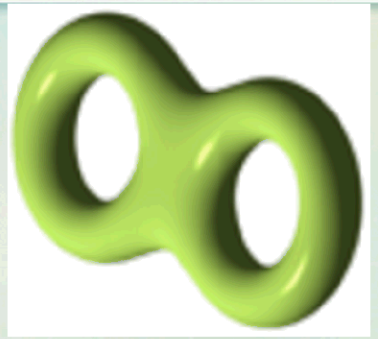
It is convenient to introduce new variable $u := q^{-s}$:

$$Z_q(u) := \prod_{\substack{P \\ \text{monic irred}}} \left(1 - u^{\deg P}\right)^{-1} = \sum_{f \text{ monic}} u^{\deg f}, \quad |u| < \frac{1}{q}$$

Here zeta is very simple:

$$\zeta_q(s) = Z_q(u) = \frac{1}{1 - qu}, \quad u := q^{-s}$$

Zeta functions of curves



C=curve/$\mathbf{F}_q$ (projective, smooth, geometrically connected)

. C= {h(x,y)=0}, h in $\mathbf{F}_q[x,y]$,

e.g. $h(x,y)=y^2-x^5-x$

Basic problem: Find number $N_1(C;\mathbf{F}_q)$ of points, i.e. # solutions of h(x,y)=0, x,y in $\mathbf{F}_q$

Counting points in field extensions:

N_n:=# pts of C in extension field of degree n

Zeta function of C

$$Z_C(u) := \exp \sum_{n=1}^{\infty} \frac{N_n u^n}{n}, \quad |u| < \frac{1}{q}$$

Example: projective line C=$\mathbf{P}^1$

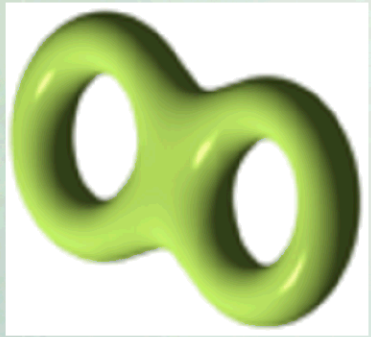
$$Z_{P^1}(u) = \frac{Z_q(u)}{1-u} = \frac{1}{(1-qu)(1-u)}$$

Example: Hyperelliptic curves

We take curves given in affine form as

$$C_Q : y^2 = Q(x), \quad Q(x) = x^{2g+1} + a_{2g}x^{2g} + \ldots + a_1 x + a_0$$

with Q(x) **square-free** of degree 2g+1 $\Longleftrightarrow$ C_Q **smooth**, of genus g



Zeta function: $Z_C(u) = Z_{P^1}(u)L^*(u,\chi_Q),$

$$L(u,\chi_Q) = \prod_P \left(1 - \chi_Q(P)u^{\deg P}\right)^{-1}$$

$L(u,\chi_Q)$ = Dirichlet L-function associated to the quadratic character

$$\chi_Q : \left(F_q[x]/(Q)\right)^* \to \{\pm 1\}$$

Properties of $Z_C(u)$, C=general curve

Rationality

$$Z_C(u) = \frac{P_C(u)}{(1-u)(1-qu)}$$

$$P_C(u) = 1 + A_1 u + \ldots + A_{2g} u^{2g} \in \mathbf{Z}[u]$$ g= genus of C

Functional equation $u \leftrightarrow 1/qu$

i.e. $s \leftrightarrow 1-s$ $(u=q^{-s})$

$$Z_C(u) = (qu^2)^{g-1} Z_C\left(\frac{1}{qu}\right)$$

Riemann Hypothesis (Weil): all zeros lie on $|u|=q^{1/2} \leftrightarrow \mathrm{Re}(s)=1/2$ $(u=q^{-s})$

Spectral interpretation

$$P_C(u) = \det(I - uq^{1/2}\Theta_C), \quad \Theta_C \in \mathrm{USp}(2g)$$

Θ_C =unitarized Frobenius conjugacy class

zeros of $Z_C(u)$ = eigenvalues of Θ_C

$$\Theta \approx \begin{pmatrix} e^{i\theta_1} & & & \\ & e^{i\theta_2} & & \\ & & \ddots & \\ & & & e^{i\theta_{2g}} \end{pmatrix}$$

The matrix Θ encodes arithmetic information about the curve C,

e.g. number of points in extension field of degree n is

$$N_n = q^n + 1 - q^{n/2}\,\mathrm{trace}\left(\Theta^{\mathrm{n}}\right)$$

Goal: compare how angles θ_j are distributed in USp(2g) as we **vary the curve** C

The hyperelliptic ensemble

To study statistics of zeros (i.e. of Θ_C), we consider a moduli space H(g,q) of hyperelliptic curves of genus g≥1 over $\mathbf{F}_q$ (q odd):

We take curves given in affine form as

$$C_Q : y^2 = Q(x), \quad Q(x) = x^{2g+1} + a_{2g}x^{2g} + \ldots + a_1 x + a_0$$

with Q(x) **square-free** of degree 2g+1

H(g,q) as a **probability space** – we pick Q **uniformly** among squarefrees;

- allows us to speak about probabilities and expected values :

$$\Pr_{H(g,q)}(\text{property}\,(C)) := \frac{\#\left\{ Q \in H(g,q) : \text{curve } C_Q \text{ has property} \right\}}{\#\, H(g,q)}$$

n.b. $\#\boldsymbol{H}(g,q) = (q-1)q^{2g}$

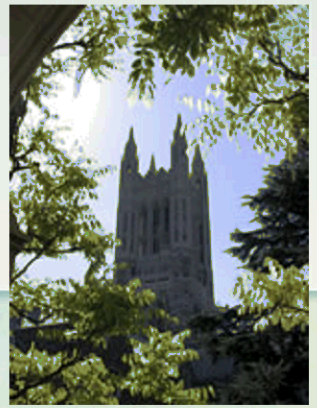
A tale of two limits

We will be interested in limiting values of various probabilities & expectations, with two main options:

genus fixed, $q \to \infty$

q fixed, genus $\to \infty$

Fixed genus, $q \to \infty$

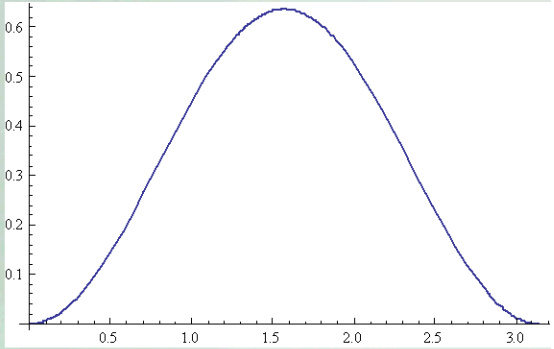


Deligne, Katz & Sarnak: for FIXED genus and large q, we have **equidistribution** of conjugacy class Θ in USp(2g).

Example: g=1 (elliptic curves) then get 2X2 matrix $\Theta = \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix}$

The angle θ is equidistributed w.r.t. Sato-Tate measure * $\sin^2 \theta \, d\theta$ on $[0,\pi]$

$$\lim_{q\to\infty} \Pr_{H(g,q)}\{a < \theta < b\} = \frac{2}{\pi}\int_a^b \sin^2(\theta)\, d\theta$$



For genus$\to\infty$ (q fixed), there is no "target space" - the ambient space USp(2g) **changes** with the genus g. So we cannot directly ask for distribution of conjugacy class Θ, instead ask for distribution of the angles $\theta_j \in \mathbf{R/Z}$

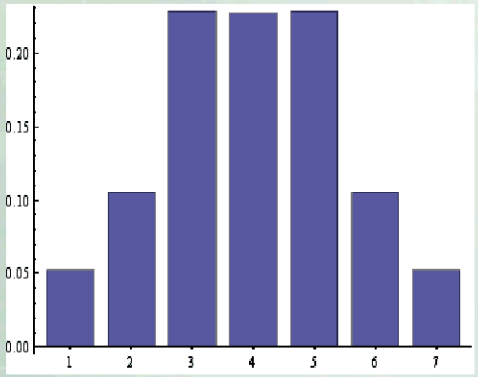
Fixed q, genus $\to\infty$

Fluctuations of trace(Θ) (Kurlberg & ZR):

For fixed q and genus $g\to\infty$, the distribution of trace(Θ) is that of a sum of q independent trinomials:

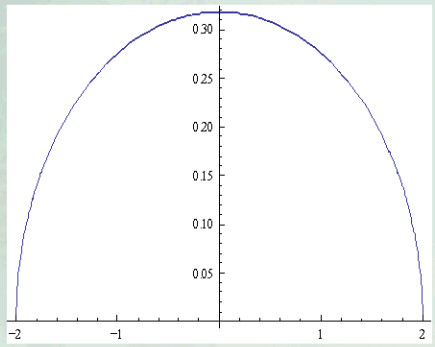
$$\lim_{g\to\infty}\Pr\left\{\mathrm{tr}(\Theta)=\frac{n}{\sqrt{q}}\right\}=\Pr\left\{X_1+\ldots+X_q=n\right\}$$

$$\Pr\left(X_i=\begin{cases}+1\\-1\\0\end{cases}\right)=\begin{cases}1/2(1+q^{-1})\\1/2(1+q^{-1})\\1/(q+1)\end{cases}$$



Compare: for fixed genus, say g=1,and **large** q, get **different** limit

$$\lim_{q\to\infty}\Pr\{\mathrm{tr}(\Theta)<x\}=\int_{2\cos\theta<x}\frac{2}{\pi}\sin^2\theta\,d\theta=\frac{2}{\pi}\int_{-1}^{x/2}\sqrt{1-w^2}\,dw$$



Reason for sum of trinomials:

NOT by method of moments or via characteristic functions !!

The trace of Θ is a (quadratic) character sum:

$$tr(\Theta) = -\frac{1}{\sqrt{q}} \sum_{x \in \mathrm{F}_q} \chi_2(Q(x))$$

we show that the value distribution of the random variable $Q \to Q(0)$ at a given point, say 0, in $\mathbf{F}_q$ is :

$$\lim_{g \to \infty} \Pr\{Q(0) = a\} = \begin{cases} q^{-1}/(1-q^{-2}), & a \neq 0 \\ 1/(q+1), & a = 0 \end{cases}$$

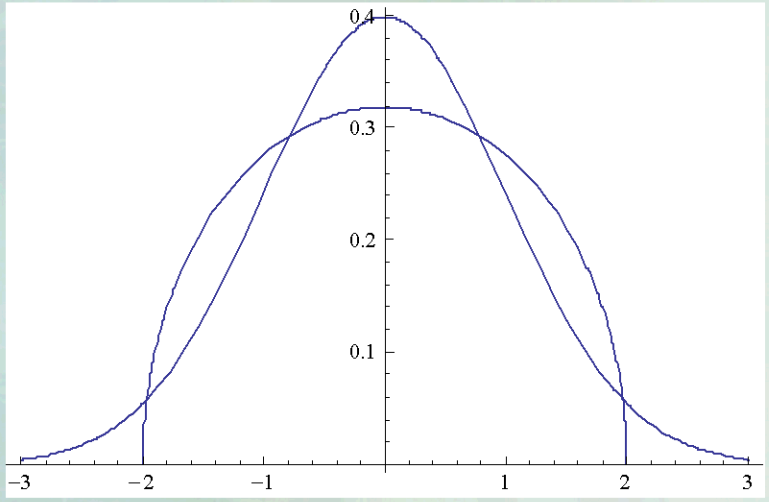
- here Q varies over all **square-free** monic polynomials of degree 2g+1

the value distribution of the random variable $Q \to \chi_2(Q(0))$ is that of a trinomial

Hybrid result –both g AND q large

THM (Kurlberg-Z.R.): If BOTH genus g and q tend to infinity then trace(Θ) has a **Gaussian** distribution.

$$\lim_{\substack{q \to \infty \\ g \to \infty}} \Pr\{\operatorname{tr}(\Theta) < x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-w^2/2)\,dw$$



Counting functions

For an interval I in **R/Z**, define a counting function

$$N_I(C) := \#\left\{ j : \theta_{j,C} \in I \right\}$$

$$\Theta \approx \begin{pmatrix} e^{i\theta_1} & & & \\ & e^{i\theta_2} & & \\ & & \ddots & \\ & & & e^{i\theta_{2g}} \end{pmatrix}$$

Fact: For for **fixed** q and genus$\to\infty$, the angles for an **individual** curve become **uniformly distributed**:

i.e., if we pick for each genus g a curve C/$\mathbf{F}_q$ of genus g, then

$$N_I(C) \sim 2g \bullet \frac{\text{length}(I)}{2\pi}, \quad g \to \infty$$

Goal: **Statistics** of $N_I(C)$ as C varies in family **H**(q,g) of hyperelliptic curves over $\mathbf{F}_q$ of given genus g

Fixed genus, q→∞

Katz & Sarnak: For fixed genus and q→∞, statistics of N_I coincide with corresponding quantity in USp(2g), e.g.

$$\lim_{q\to\infty} \Pr_{H(q,g)}(N_I = k) = \Pr_{USp(2g)}(N_I = k)$$

Reason: The conjugacy classes Θ_C become equidistributed in USp(2g).

Consequence: If we further take large genus limit, then will get Gaussian statistics

$$\lim_{genus\to\infty}\left\{\lim_{q\to\infty}\Pr_{H(q,g)}\left(a < \frac{N_I - 2g\,|I|}{\sqrt{\frac{2}{\pi^2}\log(2g\,|I|)}} < b\right)\right\} = \lim_{g\to\infty}\Pr_{USp(2g)}\left(a < \frac{N_I - 2g\,|I|}{\sqrt{\frac{2}{\pi^2}\log(2g\,|I|)}} < b\right)$$

$$= \int_a^b e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}$$

RMT: Diaconis-Shashahani, Keating-Snaith, Diaconis-Evans

Fixed q, large genus

Larsen, D. Faifman & ZR:

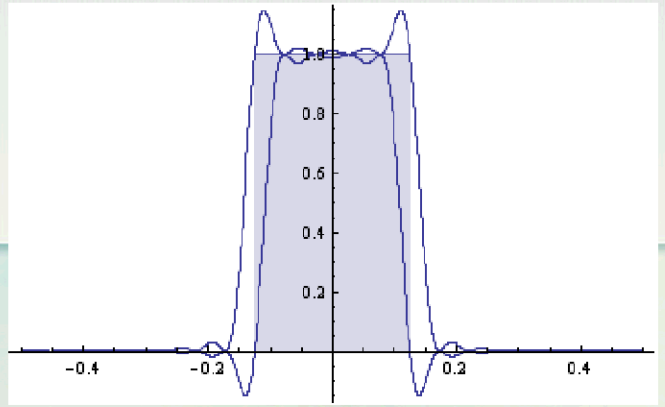
N_I has mean $2g|I|$, variance $\log(2g|I|)/2\pi^2$ & Central Limit Theorem:

$$\lim_{genus\to\infty} \Pr_{H(q,g)}\left(a < \frac{N_I - 2g\,|I|}{\sqrt{\frac{2}{\pi^2}\log(2g\,|I|)}} < b \right) = \int_a^b e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}$$

This is valid also for shrinking intervals: $|I|\to 0$ as long as $2g|I|\to\infty$ (mesoscopic regime)

Analogue: Selberg's theorem for S(t)

Method



Following Selberg, Hughes-Ng-Soundararajan

Step 1: Replace the window function for I by a smooth window, getting a new counting function ,

Then use an "explicit formula" to approximate (in L^2) the counting function N_I by a sum over primes.

$$S_K^{\pm}(Q) := -2 \sum_{\deg P \le K} \frac{\deg(P)}{\sqrt{\| P \|}} \chi_Q(P), \qquad K \approx \frac{g}{\log\log(g)}$$

Step 2: compute all moments of the approximation S_K, when

$K \approx g/\log\log(g)$

to get Gaussian distribution of S_K, hence of N_I .

The local regime length(I)≈1/g

The local regime: intervals of length $|I| \approx 1/g$, where we expect ≈ ONE zero (on average).

Study: linear statistics (=one-level density) - counting in "soft" intervals

$$Z_f(\Theta) := \sum_{j=1}^{2g} f^{\mathrm{per}}(2g \cdot \theta_j), \qquad f \in C_c^\infty(\mathbf{R})$$

The one-level density Z can be expressed via traces of Θ^n :

$$Z_f(\Theta) = \frac{1}{2g} \sum_{n=-\infty}^{\infty} \hat{f}\left(\frac{n}{2g}\right) \mathrm{tr}(\Theta^n)$$

Conj (Katz-Sarnak): for fixed q and genus→∞, local statistics will coincide with RMT statistics = limit of USp(2g).

Traces of large powers

In all regimes, distribution of N_I is dictated by expected values of products of traces of powers $\mathrm{tr}(\Theta^n)$.

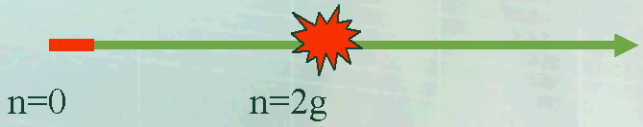
For local regime length(I)≈1/g , we need n≈g (as g→∞)

First problem : Asymptotics of $E(\mathrm{tr}(\Theta^n))$

$$\text{RMT}: \int_{USp(2g)} \mathrm{tr}\left(U^n\right) dU = \begin{cases} 0, & 0 < n < 2g \text{ odd} \\ -1, & 0 < n \le 2g \text{ even} \\ 0, & n > 2g \end{cases}$$

Thm: Assume log(g)< n < 4g - log(g). Then as g →∞

$$E_{H(g)}\left\{ \mathrm{tr}\left(\Theta^n\right) \right\} \sim \int_{USp(2g)} \mathrm{tr}\left(U^n\right) dU + \begin{cases} -\dfrac{1}{q-1}, & n = 2g \\ 0, & \text{otherwise} \end{cases}$$



RMT

deviations

Also get deviations for small n

Implications for one-level density

Corollary: If Support(f^) is in (-2,2) then

$$E_{H(g)}(Z_f) = \int_{USp(2g)} Z_f(U)dU + \frac{\mathrm{dev}(f)}{2g} + o\left(\frac{1}{g}\right)$$

Cf Ozluk-Snyder, Iwaniec, Luo & Sarnak, ……

Deviation from RMT in lower order terms (cf Miller, Ricotta, Royer, Young, Conrey-Snaith,….)

$$\text{If} \quad \int f(x)dx = 0 \quad \text{then} \quad \mathrm{dev}(f) = -\frac{f^{\wedge}(1)}{q-1}$$

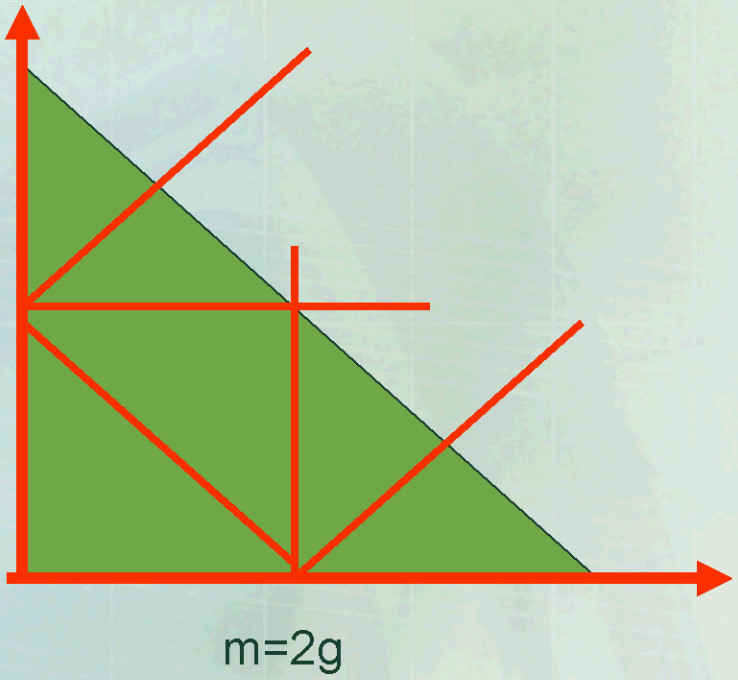
Products of two traces

Similar deviations are observed for products of two traces

$$\mathrm{E}_{H(g)}\left\{\mathrm{tr}\left(\Theta^{m}\right)\mathrm{tr}\left(\Theta^{n}\right)\right\} \qquad m+n<4g$$

agreement with RMT

deviations



Further directions

Moments of central value $\det(I-\Theta)$

- Keating-Snaith, Conrey et al.

For doing experiments, we have a numerical advantage – can compute Frobenius Θ in **polynomial time** in genus g (Kedlaya, Lauder, Wan,…)