Hyperelliptic curves, L-polynomials and random matrices

Andrew V. Sutherland

Massachusetts Institute of Technology

February 15, 2011

joint work with Kiran Kedlaya

http://arxiv.org/abs/0803.4462

Distributions of Frobenius traces

Let $E/\mathbb{Q}$ be a non-singular elliptic curve.
Let $t_p = \#E(\mathbb{F}_p) - p + 1$ denote the trace of Frobenius.

Consider the distribution of

$$x_p = -t_p/\sqrt{p} \in [-2, 2]$$

as $p \leqslant N$ varies over primes of good reduction.

What happens as $N \to \infty$?

`http://math.mit.edu/~drew`

Trace distributions in genus 1

1. Typical case (no CM)

All elliptic curves without CM have the Sato-Tate distribution.

[Clozel, Harris, Shepherd-Barron, Taylor, Barnet-Lamb, and Geraghty].

2. Exceptional cases (CM)

All elliptic curves with CM have the same exceptional distribution.

[classical]

Zeta functions and L-polynomials

For a smooth projective curve $C/\mathbb{Q}$ and a good prime p define

$$Z(C/\mathbb{F}_p;T) = \exp\left(\sum_{k=1}^{\infty} N_k T^k/k\right),$$

where $N_k = \#C/\mathbb{F}_{p^k}$. This is a rational function of the form

$$Z(C/\mathbb{F}_p;T) = \frac{L_p(T)}{(1-T)(1-pT)},$$

where $L_p(T)$ is an integer polynomial of degree $2g$. For $g = 2$:

$$L_p(T) = p^2T^4 + c_1pT^3 + c_2pT^2 + c_1T + 1.$$

Unitarized L-polynomials

The polynomial

$$\bar{L}_p(T) = L_p(T/\sqrt{p}) = \sum_{i=0}^{2g} a_i T^i$$

has coefficients that satisfy $a_i = a_{2g-i}$ and $|a_i| \leqslant \binom{2g}{i}$.

Given a curve C, we may consider the distribution of $a_1, a_2, \ldots, a_g$, taken over primes $p \leqslant N$ of good reduction, as $N \to \infty$.

In this talk we will focus on genus $g = 2$.

http://math.mit.edu/~drew

The random matrix model

$\bar{L}_p(T)$ is a real symmetric polynomial whose roots lie on the unit circle.

The random matrix model

$\bar{L}_p$(T) is a real symmetric polynomial whose roots lie on the unit circle.

Every such polynomial arises as the characteristic polynomial $\chi(T)$ of a unitary symplectic matrix in $\mathbb{C}^{2g\times 2g}$.

The random matrix model

$\bar{L}_p$(T) is a real symmetric polynomial whose roots lie on the unit circle.

Every such polynomial arises as the characteristic polynomial $\chi(T)$ of a unitary symplectic matrix in $\mathbb{C}^{2g\times 2g}$.

Conjecture (Katz-Sarnak)

For a typical curve of genus g, the distribution of $\bar{L}_p$ converges to the distribution of χ in $USp(2g)$.

This conjecture has been proven "on average" for universal families of hyperelliptic curves, including all genus 2 curves, by Katz and Sarnak.

The Haar measure on $USp(2g)$

Let $e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_g}$ denote the eigenvalues of a random conjugacy class in $USp(2g)$. The Weyl integration formula yields the measure

$$\mu = \frac{1}{g!}\Big(\prod_{j<k}(2\cos\theta_j - 2\cos\theta_k)\Big)^2 \prod_j \left(\frac{2}{\pi}\sin^2\theta_j d\theta_j\right).$$

In genus 1 we have $USp(2) = SU(2)$ and $\mu = \frac{2}{\pi}\sin^2\theta d\theta$, which is the Sato-Tate distribution.

Note that $-a_1 = \sum 2\cos\theta_j$ is the trace.

$\bar{L}_p$-distributions in genus 2

Our goal was to understand the $\bar{L}_p$-distributions that arise in genus 2, including not only the generic case, but all the exceptional cases.

This presented three challenges:

- Collecting data.
- Identifying and distinguishing distributions.
- Classifying the exceptional cases.

Collecting data

There are four ways to compute $\bar{L}_p$ in genus 2:

1. point counting: $\tilde{O}(p^2)$.
2. group computation: $\tilde{O}(p^{3/4})$.
3. p-adic methods: $\tilde{O}(p^{1/2})$.
4. ℓ-adic methods: $\tilde{O}(1)$.

Collecting data

There are four ways to compute $\bar{L}_p$ in genus 2:

1. point counting: $\tilde{O}(p^2)$.
2. group computation: $\tilde{O}(p^{3/4})$.
3. p-adic methods: $\tilde{O}(p^{1/2})$.
4. ℓ-adic methods: $\tilde{O}(1)$.

For the feasible range of $p \leqslant N$, we found (2) to be the best.
We can accelerate the computation with partial use of (1) and (4).

Computing L-series of hyperelliptic curves, ANTS VIII, 2008, KS.

Performance comparison

$p \approx 2^k$	points+group	group	p-adic
2^{14}	**0.22**	0.55	4
2^{15}	**0.34**	0.88	6
2^{16}	**0.56**	1.33	8
2^{17}	**0.98**	2.21	11
2^{18}	**1.82**	3.42	17
2^{19}	**3.44**	5.87	27
2^{20}	**7.98**	10.1	40
2^{21}	18.9	**17.9**	66
2^{22}	52	**35**	104
2^{23}		**54**	176
2^{24}		**104**	288
2^{25}		**173**	494
2^{26}		**306**	871
2^{27}		**505**	1532

Time to compute $L_p(T)$ in CPU milliseconds on a 2.5 GHz AMD Athlon

Time to compute $\bar{L}_p$ for all $p \leqslant N$

N	2 cores	16 cores
2^{16}	1	< 1
2^{17}	4	2
2^{18}	12	3
2^{19}	40	7
2^{20}	2:32	24
2^{21}	10:46	1:38
2^{22}	40:20	5:38
2^{23}	2:23:56	19:04
2^{24}	8:00:09	1:16:47
2^{25}	26:51:27	3:24:40
2^{26}		11:07:28
2^{27}		36:48:52

Characterizing distributions

The *moment sequence* of a random variable X is

$$M[X] = (\,\mathrm{E}[X^0], \mathrm{E}[X^1], \mathrm{E}[X^2], \ldots).$$

Provided X is suitably bounded, $M[X]$ exists and uniquely determines the distribution of X.

Given sample values $x_1, \ldots, x_N$ for X, the nth *moment statistic* is the mean of x_i^n. It converges to $\mathrm{E}[X^n]$ as $N \to \infty$.

Characterizing distributions

The *moment sequence* of a random variable X is

$$M[X] = (\mathrm{E}[X^0], \mathrm{E}[X^1], \mathrm{E}[X^2], \ldots).$$

Provided X is suitably bounded, $M[X]$ exists and uniquely determines the distribution of X.

Given sample values $x_1, \ldots, x_N$ for X, the nth *moment statistic* is the mean of x_i^n. It converges to $\mathrm{E}[X^n]$ as $N \to \infty$.

If X is a symmetric integer polynomial of the eigenvalues of a random matrix in $USp(2g)$, then $M[X]$ is an *integer* sequence.

This applies to all the coefficients of $\chi(T)$.

The typical trace moment sequence in genus 1

Using the measure μ in genus 1, for $t = -a_1$ we have

$$E[t^n] = \frac{2}{\pi}\int_0^\pi (2\cos\theta)^n \sin^2\theta d\theta.$$

The typical trace moment sequence in genus 1

Using the measure μ in genus 1, for $t = -a_1$ we have

$$E[t^n] = \frac{2}{\pi}\int_0^{\pi} (2\cos\theta)^n \sin^2\theta d\theta.$$

This is zero when n is odd, and for $n = 2m$ we obtain

$$E[t^{2m}] = \frac{1}{2m+1}\binom{2m}{m}.$$

and therefore

$$M[t] = (1, 0, 1, 0, 2, 0, 5, 0, 14, 0, 42, 0, 132, \ldots).$$

This is sequence A126120 in the OEIS.

The typical trace moment sequence in genus $g > 1$

A similar computation in genus 2 yields

$$M[t] = (1, 0, 1, 0, 3, 0, 14, 0, 84, 0, 594, \ldots),$$

which is sequence A138349, and in genus 3 we have

$$M[t] = (1, 0, 1, 0, 3, 0, 15, 0, 104, 0, 909, \ldots),$$

which is sequence A138540.

In genus g, the nth moment of the trace is the number of returning walks of length n on $\mathbb{Z}^g$ with $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_g \geqslant 0$ [Grabiner-Magyar].

The exceptional trace moment sequence in genus 1

For an elliptic curve with CM we find that

$$E[t^{2m}] = \frac{1}{2}\binom{2m}{m}, \qquad \text{for } m > 0$$

yielding the moment sequence

$$M[t] = (1, 0, 1, 0, 3, 0, 10, 0, 35, 0, 126, 0, \ldots),$$

whose even entries are A008828.

An exceptional trace moment sequence in Genus 2

For a hyperelliptic curve whose Jacobian is isogenous to the direct product of two elliptic curves, we compute $M[t] = M[t_1 + t_2]$ via

$$\mathrm{E}[(t_1 + t_2)^n] = \sum \binom{n}{i} \mathrm{E}[t_1^i]\mathrm{E}[t_2^{n-i}].$$

For example, using

$$M[t_1] = (1, 0, 1, 0, 2, 0, 5, 0, 14, 0, 42, 0, 132, \ldots),$$
$$M[t_2] = (1, 0, 1, 0, 3, 0, 10, 0, 35, 0, 126, 0, 462, \ldots),$$

we obtain $A138551$,

$$M[t] = (1, 0, 2, 0, 11, 0, 90, 0, 889, 0, 9723, \ldots).$$

The second moment already differs from the standard sequence, and the fourth moment differs greatly (11 versus 3).

Sieving for exceptional curves

We surveyed the $\bar{L}_p$-distributions of genus 2 curves

$$y^2 = x^5 + c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0,$$

$$y^2 = b_6x^6 + b_5x^5 + b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0,$$

with integer coefficients $|c_i| \leqslant 64$ and $|b_i| \leqslant 16$, over 10^{10} curves.

We initially set $N \approx 2^{12}$, discarded about 99% of the curves (those whose moment statistics were "unexceptional"), then repeated this process with $N \approx 2^{16}$ and $N \approx 2^{20}$.

We eventually found 30,000 non-isomorphic curves with apparently exceptional distributions, many of which coincided.

Representative examples were computed to high precision $N \approx 2^{26}$.

Survey highlights

- The moment statistics always appear to converge to integers.
- 20 distinct trace distributions (eventually found 23 of 24 predicted). This exceeds the possibilities for $\mathrm{End}(\mathrm{Jac}(C))$, $\mathrm{Aut}(C)$, or $\mathrm{MT}(C)$.
- The same $\bar{L}_p$-distribution can arise for split and simple Jacobians.
- The density of zero traces can be any of

$$\{0, 1/6, 1/4, 1/2, 7/12, 5/8, 3/4, 13/16, 7/8\}.$$

Density 0 occurs in several exceptional cases.

Survey highlights (new results)

- The moment statistics always appear to converge to integers.
- 26 distinct $\bar{L}_p$-distributions (out of 26 predicted). This exceeds the possibilities for $\mathrm{End}(\mathrm{Jac}(C))$, $\mathrm{Aut}(C)$, or $\mathrm{MT}(C)$.
- The same $\bar{L}_p$-distribution can arise for split and simple Jacobians.
- The density of zero traces can be any of

$$\{0, 1/6, 1/4, 1/2, 7/12, 5/8, 3/4, 11/16, 19/24, 13/16, 7/8\}.$$

Density 0 occurs in several exceptional cases.

Survey highlights (new results)

- The moment statistics always appear to converge to integers.
- 26 distinct $\bar{L}_p$-distributions (out of 26 predicted).
  This exceeds the possibilities for $\mathrm{End}(\mathrm{Jac}(C))$, $\mathrm{Aut}(C)$, or $\mathrm{MT}(C)$.
- The same $\bar{L}_p$-distribution can arise for split and simple Jacobians.
- The density of zero traces can be any of

  $$\{0, 1/6, 1/4, 1/2, 7/12, 5/8, 3/4, 11/16, 19/24, 13/16, 7/8\}.$$

  Density 0 occurs in several exceptional cases.
- Distinct $\bar{L}_p$-distributions may have identical trace distributions.
  As of 2/15/2011, we have identified 30 distinct $\bar{L}_p$-distributions.

Random matrix subgroup model

Conjecture

For a genus g curve C, the distribution of $\bar{L}_p$ converges to the distribution of χ in some infinite compact subgroup $H \subseteq USp(2g)$.

*Equality holds if and only if C has large Galois image.**

*image of $\rho_\ell : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(T_\ell(C))$ Zariski dense in $GSp(2g, \mathbb{Z}_\ell)$.

Representations of genus 1 distributions

The Sato-Tate distribution has $H = USp(2g)$, the typical case.

For CM curves, consider the subgroup of $USp(2) = SU(2)$:

$$H = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} i\cos\theta & i\sin\theta \\ i\sin\theta & -i\cos\theta \end{pmatrix} : \theta \in [0, 2\pi] \right\}.$$

This is a compact group (the normalizer of $SO(2)$ in $SU(2)$).

Its Haar measure yields the desired moment sequence.

Candidate subgroups in genus 2

Let $G_1 = SU(2)$ and $G_2 = N(SO(2)) \subset SU(2)$.

- $USp(4)$ — generic genus 2 curve.
- Index 2 subgroup K of $N(SO(2) \times SO(2))$ — genus 2 CM curve.
- $G_1 \times G_1$, $G_1 \times G_2$, $G_2 \times G_2$ — products of 2 elliptic curves.
- $J(G_1 \times G_1)$ (but not $J(G_2 \times G_2)$ [Serre]).
- $G_i \otimes G_0$ for some finite subgroup G_0 of $SU(2)$ —
  "twisted" product of an elliptic curve with itself (22 cases!).

We require elements of G_0 to have traces whose squares lie in $\mathbb{Z}$.
We may assume $-I \in G_0$.

A very recent example

The genus 2 curve

$$y^2 = 297x^6 - 324x^5 - 2970/37x^4 + 720/37x^3 + 1980/1369x^2 - 144/1369x - 88/50653$$

found by Fité and Lario in December 2010 has $\bar{L}_p$-distribution matching $G_2 \otimes G_0$, where G_0 is a binary dihedral group of order 24.

This distribution was predicted by our model that did not show up in our survey. It also occurs for the curve

$$y^2 = x^6 - 9x^5 - 15x^4 + 30x^3 + 15x^2 - 9x - 1$$

whose coefficients lie just beyond the range of our search.

A very recent example

The genus 2 curve

$$y^2 = 297x^6 - 324x^5 - 2970/37x^4 + 720/37x^3 + 1980/1369x^2 - 144/1369x - 88/50653$$

found by Fité and Lario in December 2010 has $\bar{L}_p$-distribution matching $G_2 \otimes G_0$, where G_0 is a binary dihedral group of order 24.

This distribution was predicted by our model that did not show up in our survey. It also occurs for the curve

$$y^2 = x^6 - 9x^5 - 15x^4 + 30x^3 + 15x^2 - 9x - 1$$

whose coefficients lie just beyond the range of our search.

The paremetrizations they used (due to Cardona) also yielded two new distributions that were not predicted by our model!

Finite subgroups of $SU(2)$

A finite subgroup of $SU(2)$ is isomorphic to one of the following:

- Cyclic C_n group of order n.
- Binary dihedral group BD_n of order $4n$.
- Binary tetrahedral group BT (order 24).
- Binary octahedral group BO (order 48).
- Binary icosahedral group BI (order 120).

There are 12 groups on this list that are candidates for G_0.

All of these give rise to distributions that match an exceptional $\bar{L}_p$-polynomial distribution in genus 2.

Finite subgroups of $SU(2)$

A finite subgroup of $SU(2)$ is isomorphic to one of the following:

- Cyclic C_n group of order n.
- Binary dihedral group BD_n of order $4n$.
- Binary tetrahedral group BT (order 24).
- Binary octahedral group BO (order 48).
- Binary icosahedral group BI (order 120).

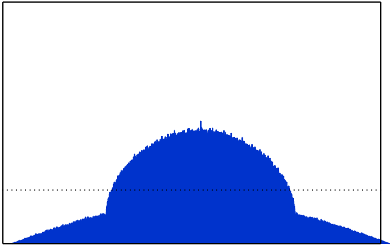
There are 12 groups on this list that are candidates for G_0.

All of these give rise to distributions that match an exceptional $\bar{L}_p$-polynomial distribution in genus 2.

This includes the two new distributions, arising from BT and BO, which only seem applicable to G_2.

H	#	d	$c(H)$	$z(H)$	M_2	M_4	M_6	M_8	M_{10}
$USp(4)$	1	10	1	0	1	3	14	84	594
K	19	2	4	3/4	1	9	100	1225	15876
$G_1 \times G_1$	2	6	1	0	2	10	70	588	5544
$G_1 \times G_2$	3	4	2	0	2	11	90	889	9723
$G_2 \times G_2$	8	2	4	1/4	2	12	110	1260	16002
$J(G_1 \times G_1)$	9	6	2	1/2	1	5	35	294	2772
$G_1 \otimes C_2$	5	3	1	0	4	32	320	3584	43008
$G_1 \otimes C_4$	11b	3	2	1/2	2	16	160	1792	21504
$G_1 \otimes C_6$	4	3	3	0	2	12	110	1204	14364
$G_1 \otimes C_8$	7	3	4	1/4	2	12	100	1008	11424
$G_1 \otimes C_{12}$	6	3	6	1/6	2	12	100	980	10584
$G_1 \otimes BD_1$	11	3	2	1/2	2	16	160	1792	21504
$G_1 \otimes BD_2$	18	3	4	3/4	1	8	80	896	10752
$G_1 \otimes BD_3$	10	3	6	1/2	1	6	55	602	7182
$G_1 \otimes BD_4$	16	3	8	5/8	1	6	50	504	5712
$G_1 \otimes BD_6$	14	3	12	7/12	1	6	50	490	5292
$G_2 \otimes C_2$	13	1	2	1/2	4	48	640	8960	129024
$G_2 \otimes C_4$	21b	1	4	3/4	2	24	320	4480	64512
$G_2 \otimes C_6$	12	1	6	1/2	2	18	220	3010	43092
$G_2 \otimes C_8$	17	1	8	5/8	2	18	200	2520	34272
$G_2 \otimes C_{12}$	15	1	12	7/12	2	18	200	2450	31752
$G_2 \otimes BD_1$	21	1	4	3/4	2	24	320	4480	64512
$G_2 \otimes BD_2$	23	1	8	7/8	1	12	160	2240	32256
$G_2 \otimes BD_3$	20	1	12	3/4	1	9	110	1505	21546
$G_2 \otimes BD_4$	22	1	16	13/16	1	9	100	1260	17136
$G_2 \otimes BD_6$	24	1	24	19/24	1	9	100	1225	15876
$G_2 \otimes BT$	25	1	24	5/8	1	6	60	770	10836
$G_2 \otimes BO$	26	1	48	11/16	1	6	50	525	6426

`smalljac` now available in purple Sage.



drew@math.mit.edu

Hyperelliptic curves, L-polynomials and random matrices

Andrew V. Sutherland

Massachusetts Institute of Technology

February 15, 2011

joint work with Kiran Kedlaya

http://arxiv.org/abs/0803.4462