More points on curves over finite field extensions than expected

Bradley W. Brock

Natural Science Division
Pepperdine University
Malibu, CA 90263-4321
USA
E-mail: bbrock@pepperdine.edu

and

Andrew Granville

Department of Mathematics
University of Georgia
Athens, GA 30602-7403
USA
E-mail: andrew@math.uga.edu

On average, there are $q^r + o(q^{r/2})$ $\mathbf{F}_{q^r}$-rational points on curves of genus g defined over $\mathbf{F}_{q^r}$. This is also true if we restrict our average to genus g curves defined over $\mathbf{F}_q$, provided r is odd or $r > 2g$. However, if $r = 2, 4, 6, \ldots$ or $2g$ then the average is $q^r + q^{r/2} + o(q^{r/2})$. We give a number of proofs of the existence of these $q^{r/2}$ "extra" points, and in some cases give a precise formula, but we are unable to provide a satisfactory explanation for this phenomenom.

1. INTRODUCTION

Let C be a nonsingular, projective, geometrically irreducible curve of genus g over the finite field $\mathbf{F}_q$ with q elements. Weil showed that there are g algebraic integers $\alpha_1, \alpha_2, \ldots, \alpha_g$ associated to C such that $|\alpha_j| = \sqrt{q}$ and such that for every r the number $N_r(C)$ of $\mathbf{F}_{q^r}$-rational points on C is

$$\# \ C(\mathbf{F}_{q^r}) = N_r(C) = q^r + 1 - a_r(C) \text{ where} \tag{1}$$

$$a_r(C) \ = \ (\alpha_1^r + \overline{\alpha}_1^r) + (\alpha_2^r + \overline{\alpha}_2^r) + \ldots + (\alpha_g^r + \overline{\alpha}_g^r). \tag{2}$$

In particular, all the $N_r(C)$ are determined by $N_1(C), \ldots, N_g(C)$. More precisely, for every g and r there is a polynomial over $\mathbf{Q}$ in the indeterminates $q, a_1, \cdots, a_g$ that evaluates to the power sum $a_r := a_r(C)$. Because of complex conjugation and permutation there are generically $2^g g!$ equivalent choices for the g-tuple $(\alpha_1, \dots, \alpha_g)$, and for our purposes we shall assume that one of these has been chosen at random.

Define

$$a_r(g,q) := \sum_{\substack{C/\mathbf{F}_q \\ \text{genus}(C)=g}} \frac{a_r(C)}{|\mathrm{Aut}(C/\mathbf{F}_q)|} \Bigg/ \sum_{\substack{C/\mathbf{F}_q \\ \text{genus}(C)=g}} \frac{1}{|\mathrm{Aut}(C/\mathbf{F}_q)|},$$

that is, the average of $a_r(C)$ as C varies over a set of representatives of each of the $\mathbf{F}_q$-isomorphism classes of curves of genus g defined over $\mathbf{F}_q$, weighted by $1/|\mathrm{Aut}(C/\mathbf{F}_q)|$ (this will be discussed further in section 7). An immediate consequence of Weil's result is that $|a_r(C)| \le 2gq^{r/2}$ and therefore $|a_r(g,q)| \le 2gq^{r/2}$. One might naïvely suppose that the real part of the rth moment of the $\alpha_j/\sqrt{q}$, as we vary of all curves of fixed genus g defined over $\mathbf{F}_q$ for large q are distributed roughly symmetrically about the origin, so that $a_r(g,q) = o(q^{r/2})$. This is true for odd r, and for $r > 2g$; moreover, when $r > 2g$, the rth moments of the $\alpha_j/\sqrt{q}$ are roughly uniformly distributed around the unit circle. On the other hand our naive supposition is untrue for even $r \le 2g$, in which case the mean of the rth moments of the $\alpha_j/\sqrt{q}$ is $-1/2g + o(1)$ rather than just $o(1)$: Consequently $a_r(g,q) = -\{1 + o(1)\}q^{r/2}$ for such r. We also show that the analogues of these results hold for the hyperelliptic curves of genus g over $\mathbf{F}_q$.

In section 2 we show that these asymptotic results, for genus one, are an immediate consequence of a result of Birch [1]: One can define the Sato-Tate distribution, X_{ST}, to be the unique distribution on the unit circle such that

$$\mathbf{E}(X_{ST}^r) = \begin{cases} -\frac{1}{2} & \text{if } r = -2 \text{ or } 2 \\ \ \ 1 & \text{if } r = 0 \\ \ \ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $\mathbf{E}(X)$ is the expected value of the random variable X. Birch [1] proved that as X varies over all curves of genus 1 defined over $\mathbf{F}_q$ the distribution of the values $\{\alpha_1(E)/\sqrt{q} : E/\mathbf{F}_q\}$ on the unit circle tends to the Sato-Tate distribution as $q \to \infty$. Katz and Sarnak [11] proved the genus g generalization of Birch's result, and in section 4 we use this to compute the moments of this generalized Sato-Tate distribution, which gives our asymptotic results for every genus g.

The Selberg trace formula is tailor-made for giving a precise formula for $a_r(1,q)$ in terms of the trace of Hecke operators on certain vector spaces of cusp forms (see section 3). For example $a_2(1,q) = -q - 1/q$, so that the average of $N_2(E)$ is $q^2 + q + 1 + 1/q$ as E varies over all curves of genus 1 defined over $\mathbf{F}_q$. In sections 6 and 7 we use elementary counting arguments to determine exact averages for various families of hyperelliptic curves. In particular we show that $N_2(C)$ is $q^2+q+1-(-1)^g q^{1-2g}$, on average, as C ranges over all hyperelliptic curves of genus g over $\mathbf{F}_q$.

In section 5 we guess at the average of $N_r(C)$ for hyperelliptic curves of genera 2 and 3 and small r, based on computer experimentation. Interestingly the averages that arise in the genus one case (section 3) are elementary functions of q, for $r < 10$. Then, for $r = 10$ and 12 the formula is in terms of the Ramanujan tau function, which seems to also occur for genera 2 and 3, according to our experiments.

So why is $a_2(1,q) = -q + o(q)$, whereas $a_r(1,q) = o(q)$ for all $r > 2$? That is, where do the extra q points come from in the quadratic extension? And, why are there no such extra points over extensions of higher degree? Moreover what explains the analogous phenomenom for higher genus curves? We will restrict our discussion here to hyperelliptic curves; that is, curves of the form

$$y^2 = f(x) \text{ where } f(x) \in \mathbf{F}_q[x] \text{ has no repeated roots.} \tag{4}$$

Let q be a power of prime $p > 3$. Suppose that χ is the character of order two in $\mathbf{F}_q$; that is, $\chi(n) = 1$ if n is a nonzero square in $\mathbf{F}_q$, and $\chi(n) = -1$ if n is not a square in $\mathbf{F}_q$. The number of (affine) $\mathbf{F}_q$-points on (4) is $q + \sum_{m\in\mathbf{F}_q} \chi(f(m))$. Select $a \in \mathbf{F}_q$ so that $\chi(a) = -1$. Then the number of (affine) points on $y^2 = f(x)$ and the number on $y^2 = af(x)$ is precisely $2q + \sum_{m\in\mathbf{F}_q}(\chi(f(m)) + \chi(af(m))) = 2q$. There are total of two $\mathbf{F}_q$ points at infinity on these two curves after resolving singularities (see Lemma 7.1). Thus on average they each have $q+1$ $\mathbf{F}_q$ points, so that $N_1(C)$ is $q+1$ on average. If r is odd then a is not a square in $\mathbf{F}_{q^r}$ either, and so the same argument in that field shows that $N_r(C)$ is $q^r + 1$ on average.

There are $(q^r-1)/2$ nonzero elements in $\mathbf{F}_{q^r}$ which are squares of other elements in the field, and thus each have two square roots; and a similar number which are not squares. Thus, on average, an element of $\mathbf{F}_{q^r}$ has one square root. The squares seem to be more-or-less randomly distributed in the field. If one grants that the values taken by a polynomial $f(x)$ are also randomly distributed in the field then one might expect, on average, that the number of solutions to (4) is roughly $q^r + O(1)$. However if f is defined over $\mathbf{F}_q$ then $f(m) \in \mathbf{F}_q$ for every $m \in \mathbf{F}_q$ and so is a square in $\mathbf{F}_{q^2}$. Thus for each of these m there are two $\mathbf{F}_{q^2}$-solutions to $y^2 = f(m)$ if $f(m) \neq 0$. Thus we get an "extra" $q + O(1)$ points on (5) over what we had previously expected — this does seem to explain why $a_2(g,q) = -q + o(q)$. However the same argument suggests that $a_4(g,q) = -q^2 + o(q^2)$, but this is false for $g = 1$ as we have seen, so our heuristic is misleading! We would like to see a better elementary explanation (though see [10]).

Instead of fixing the finite field and varying over curves of a given genus, one can ask similar questions fixing a curve over $\mathbf{Q}$ and reducing mod$p$ to finite fields. For a given elliptic curve define $\alpha_p = \alpha_1(E \bmod p)$ when E reduces to an elliptic curve in $\mathbf{F}_p$ (as happens for all but finitely many p):

If E does not have complex multiplication then the Sato-Tate conjecture (which remains open) states that $\alpha_p/\sqrt{p}$ is distributed according to the Sato-Tate distribution, as above. Hence, by (3) the average value of $\#E(\mathbf{F}_{p^r})$ is $p^r + o(p^{r/2})$ if $r \neq 2$; and $p^2 + p + o(p)$ for $r = 2$.

If E does have complex multiplication (which is relatively rare) then we understand the distribution of the α_p: If p does not split in the endomorphism ring of E, then $\alpha = \pm i\sqrt{p}$ so that $\#E(\mathbf{F}_{p^r}) = p^r + 1 - (i^r + (-i)^r)\sqrt{p}^r$. If p does split in the endomorphism ring of E then, by Hecke's equidistribution theorem for L-functions with unitary grossencharacter of infinite order, the $\alpha_p/\sqrt{p}$ are equidistributed around the unit circle. Therefore the mean value of $(\alpha_p/\sqrt{p})^r$ is $o(1)$ for all $r \geq 1$, and so the average value of $\#E(\mathbf{F}_{p^r})$, for such p, is $p^r + o(p^{r/2})$. Since p splits half the time, the average value of $\#E(\mathbf{F}_{p^r})$ is $p^r + 1 - (i^r + (-i)^r + o(1))\sqrt{p}^r/2$.

Therefore, whether or not E has complex multiplication the average value of $\#E(\mathbf{F}_p)$ is $p + o(\sqrt{p})$, and the average value of $\#E(\mathbf{F}_{p^2})$ is $p^2 + p + o(p)$.

We can deduce this directly from results in the literature on L-functions though with a different definition of "average"; that is, a different probability measure on the primes. Before we were implicitly defining the expected value of a function f on the primes by

$$\mathbf{E}(f(p)) = \lim_{x\to\infty} \frac{\sum_{p\leq x} f(p)}{\sum_{p\leq x} 1};$$

but now we define the expected value to be

$$\mathbf{E}(f(p)) = \lim_{x\to\infty} \frac{\sum_{p\leq x} f(p)/p}{\sum_{p\leq x} 1/p}.$$

We will prove that $\mathbf{E}(a_p/\sqrt{p}) = 0$ and $\mathbf{E}(a_{p^2}/p) = -1$, where $a_p = \alpha_p + \overline{\alpha}_p$ and $a_{p^2} = \alpha_p^2 + \overline{\alpha}_p^2 = a_p^2 - 2p$; that is, the a_1 and a_2 of E mod p.

Wiles et al. [2, 19] have shown that

$$\tilde{L}(E,s) = \prod_{p\nmid N} \left(1 - \frac{\alpha_p}{p^s}\right)^{-1} \left(1 - \frac{\overline{\alpha}_p}{p^s}\right)^{-1} \quad \text{for } \mathrm{Re}(s) > 3/2,$$

can be analytically continued to the whole complex plane where N is the conductor. In particular the product converges for $s = 3/2$, and from the usual contour integration one can then deduce that

$$\sum_{p\leq x} \frac{a_p/\sqrt{p}}{p}$$

converges, so that $\mathbf{E}(a_p/\sqrt{p}) = 0$. Similarly Wiles et al. and Shimura [17] have shown that the symmetric square L-function

$$\tilde{L}(\mathrm{Sym}^2 E, s) = \zeta(s-1) \prod_{p\not| N} \left(1 - \frac{\alpha_p^2}{p^s}\right)^{-1} \left(1 - \frac{\overline{\alpha}_p^2}{p^s}\right)^{-1} \quad \text{for } \mathrm{Re}(s) > 2,$$

can be analytically continued to the whole complex plane and, in particular, the product converges at $s = 2$. (Note that by convention $\mathrm{Sym}^2 E$ is not the usual symmetric product of varieties but rather a variety $A/\mathbf{Q}$ such that the action of $\mathrm{Frob}_p(A)$ on $H^2(A)$ is isomorphic to the action of $\mathrm{Sym}^2\mathrm{Frob}_p(E)$ on $\mathrm{Sym}^2 H^1(E)$ for every p where $H^{\cdot}$ is a Weil cohomology.) Thus we can deduce

$$\sum_{p\leq x} \frac{1 + a_{p^2}/p}{p}$$

converges, so that $\mathbf{E}(a_{p^2}/p) = -1$.

If one can prove analytic continuation for the L-functions of arbitrary symmetric powers of E, at least up to and including the edge of the critical strip, then one can determine that the average value of $\#E(\mathbf{F}_{p^r})$ is as predicted above, as noted by Serre [15]. Recently Bump et al. [3] have given such a result for the third symmetric power but the fourth symmetric power seems beyond reach for now.

2. THE SATO-TATE DISTRIBUTION

The Sato-Tate distribution for the random variable $X = e^{i\theta}$, $-\pi < \theta \leq \pi$ is given by

$$\mathrm{Prob}(a \leq \theta \leq b) = \frac{1}{2\pi} \int_a^b 2\sin^2 t \; dt = \frac{1}{2\pi} \int_a^b (1 - \cos 2t) \; dt$$

Thus

$$\begin{aligned} \mathbf{E}(X^r + \overline{X}^r) &= \mathbf{E}(2\cos r\theta) = \frac{2}{\pi} \int_a^b \cos rt \;\; \sin^2 t \; dt \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} (\cos rt)(1 - \cos 2t) \; dt \end{aligned}$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi}\left(\cos rt - \frac{1}{2}\cos(r+2)t - \frac{1}{2}\cos(r-2)t\right)\,dt$$
$$= \begin{cases} -1 & \text{if } r=-2 \text{ or } 2 \\ 2 & \text{if } r=0 \\ 0 & \text{if } r\neq 0,-2 \text{ or } 2 \end{cases}$$

which implies (3) since the distribution is symmetric about the real axis.

We note that for the Sato-Tate distribution we have

$$\text{Prob } (a \le \{n\theta \bmod 2\pi\} \le b) = \begin{cases} \frac{1}{2\pi}\int_a^b (1-\cos t)\ dt & \text{for } \ n=2, \\ (b-a)/2\pi & \text{for } \ n\ge 3 \end{cases}$$

because

$$\frac{1}{n}\sum_{j=0}^{n-1}\mu\left(\frac{t+2\pi j}{n}\right) = \begin{cases} 1-\cos 2t & \text{if } n=1 \\ 1-\cos t & \text{if } n=2 \\ 1 & \text{if } n\ge 3 \end{cases}$$

where $\mu(t) = 1-\cos 2t$. What other probability distributions on the circle have this property?

Proposition 2.1. *Let*

$$\mu(t) = a_0 + \sum_{n\ge 1}(a_n\cos(nt) + b_n\sin(nt))$$

be a real function with period 2π and $a_0 = \frac{1}{2\pi}\int_0^{2\pi}\mu(t)dt = 1$. Then

$$\frac{1}{N}\sum_{j=0}^{N-1}\mu\left(\frac{\theta+2\pi j}{N}\right) = 1$$

for all θ and all integers $N > m$ if and only $a_n = b_n = 0$ for $n > m$.

Proof.

$$(1/N)\sum_{j=0}^{N-1}\exp(in(\theta+2\pi j)/N) = \begin{cases} \exp(in\theta/N) & \text{if } N|n \\ 0 & \text{otherwise.} \end{cases}$$

since we get a sum over N/nth roots of unity. Thus the hypothesis is equivalent to $\sum_{n>0,\ N|n}(a_n\cos(nt) + b_n\sin(nt)) = 0$ for every $N > m$ and every t. So by the uniqueness of the Fourier expansion this is equivalent to $a_n = b_n = 0$ for every n divisible by some $N > m$. In particular

$n = N$ implies our result. Conversely, if $a_n = b_n = 0$ for every $n > m$ then certainly $a_n = b_n = 0$ for every n divisible by some $N > m$. ▌

Note that for such a distribution function $\mathbf{E}(\cos(r\theta)) = a_{|r|}/2$ if $r \neq 0$, which again gives us (3).

3. SELBERG'S TRACE FORMULA

Selberg's trace formula [13] implies that for all even integers $k \geq 0$

$$\sigma_k(T_p) + 1 = -\frac{1}{2} \sum_{E/\mathbf{F}_p} \frac{\alpha_E^{k-1} - \overline{\alpha}_E^{k-1}}{\alpha_E - \overline{\alpha}_E} \tag{5}$$

where the sum is over representatives of the $\mathbf{F}_p$-isomorphism classes of elliptic curves weighted by $2/|\mathrm{Aut}(E/\mathbf{F}_p)|$, where $\alpha_E = \alpha_1(E)$, and $\sigma_k(T_p)$ is the trace of the Hecke operator T_p acting on the cusp forms of weight k in $\mathrm{SL}(2, \mathbf{Z})$ for $k \geq 4$, with $\sigma_0(T_p) = 0$ and $\sigma_2(T_p) = -p-1$ (for $k = 0$ and 2 this follows from $\sum_{E/\mathbf{F}_p} 1 = 2p$.) Now since

$$\frac{\alpha_E^{k+1} - \overline{\alpha}_E^{k+1}}{\alpha_E - \overline{\alpha}_E} - p\frac{\alpha_E^{k-1} - \overline{\alpha}_E^{k-1}}{\alpha_E - \overline{\alpha}_E} = \alpha_E^k + \overline{\alpha}_E^k$$

we deduce that

$$\mathrm{mean}\left(\frac{a_k(E)}{p^{k/2}}\right) = \frac{1}{2p} \sum_{E/\mathbf{F}_p} \frac{a_k(E)}{p^{k/2}} = \frac{\sigma_k(T_p) + 1}{p^{k/2}} - \frac{\sigma_{k+2}(T_p) + 1}{p^{k/2+1}} \tag{6}$$

for even $k \geq 2$. Now $\sigma_k(T_p) = 0$ for $k = 4, 6, 8, 10, 14$ and $\sigma_{12}(T_p) = \tau(p)$, Ramanujan's τ-function, so that

$$\begin{aligned}
&\mathrm{mean}\ \ (\#E(\mathbf{F}_{p^2})) = p^2 + p + 1 + 1/p \\
&\mathrm{mean}\ \ (\#E(\mathbf{F}_{p^k})) = p^k + 1/p \ \text{ for } k = 4, 6, 8 \\
&\mathrm{mean}\ \ (\#E(\mathbf{F}_{p^{10}})) = p^{10} + \frac{\tau(p) + 1}{p} \\
&\mathrm{mean}\ \ (\#E(\mathbf{F}_{p^{12}})) = p^{12} - \tau(p) + 1/p.
\end{aligned}$$

By Deligne's proof [4, 5] of the Ramanujan-Petersson conjecture $\sigma_k(T_p) = O(p^{(k-1)/2+\epsilon})$, and so we have

$$\mathrm{mean}(\#E(\mathbf{F}_{p^k})) = p^k + O(p^{(k-1)/2+\epsilon}),\ \text{for even } k \geq 4,$$

or in other words $\mathrm{mean}(a_k(E)/p^{k/2}) = O(p^{-1/2+\epsilon}) = o(1)$.

More generally for q a power of a prime p, the average number of points over $\mathbf{F}_{q^k}$ on an elliptic curve over $\mathbf{F}_q$ is

$$q^k + \frac{1+\sigma_{k+2}(T_q) - p^{k+1}\sigma_{k+2}(T_{q/p^2})}{q} - \sigma_k(T_q) + p^{k-1}\sigma_k(T_{q/p^2})$$

where we set $\sigma_2(T_q) = -(pq-1)/(p-1)$ and $\sigma_k(T_1)$ equal to the dimension of the space of cusp forms of weight k in $\mathrm{SL}(2,\mathbf{Z})$. In particular if $k=10$ the average is $q^{10}+1/q+\tau'(q)/q$, and if $k=12$ the average is $q^{12}+1/q-\tau'(q)$, where $\tau'(q)=\tau(q)$ if $q=p$ and is equal to $\tau(q)-p^{11}\tau(q/p^2)$ if q is a higher power of p.

4. THE GENERALIZED SATO-TATE DISTRIBUTION

We are now going to study the distribution of $(\alpha_1, \alpha_2, \ldots, \alpha_g)$, as in (1.1), as we vary all curves of genus g, defined over $\mathbf{F}_q$. Note that these come in conjugate pairs and we do not distinguish α and $\overline{\alpha}$. Since each $|\alpha_j| = \sqrt{q}$ we renormalize and ask for the distribution of $(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_g})$ where $\alpha_j = \sqrt{q}e^{i\theta_j}$. In section 2 we saw the distribution function for $g=1$. In the remarkable book [11] Katz and Sarnak show that such distribution functions for families of varieties satisfying certain monodromy conditions, are intimately related with the distribution functions for the eigenvalues of the compact classical groups in their standard representations. In the case of the curves of genus g, that classical group in question is $\mathrm{USp}(2g)$. Weyl [18] gave the distribution law for these eigenvalues:

$$Pr(a_1 \le \theta_1 \le b_1, \ldots,\ a_g \le \theta_g \le b_g) = \frac{1}{(2\pi)^g}\int_{a_1}^{b_1} \cdots \int_{a_g}^{b_g} \rho(\theta_1,\ldots,\theta_g)d\theta_g \ldots d\theta_1,$$

where

$$\rho(\theta_1,\ldots,\theta_g) = \frac{1}{g!}\prod_{1\le i<j\le g}(2\cos\theta_i - 2\cos\theta_j)^2 \prod_{1\le i\le g}(2\sin^2\theta_i).$$

Note that ρ is symmetric and even in all the variables. Let $t_j = e^{i\theta_j}$. Note that ρ has degree $2g$ in both t_j and $1/t_j$, so we may write

$$\begin{aligned}\rho(\theta_1,\ldots,\theta_g) &= \sum_{|k_1|\le 2g} \cdots \sum_{|k_g|\le 2g} c(k_1,\ldots,k_g)t_1^{k_1}\ldots t_g^{k_g} \qquad (7)\\ &= \frac{1}{(-2)^g g!}\prod_{1\le i<j\le g}(t_i-t_j)^2(t_it_j-1)^2 \prod_{1\le i\le g}\frac{(t_i^2-1)^2}{t_i^{2g}}.\end{aligned}$$

for some coefficients $c(\mathbf{k})$. These coefficients are just the product moments because $\mathbf{E}(t_1^{a_1}\dots t_g^{a_g})$

$$\begin{aligned}
&= \frac{1}{(2\pi i)^g}\oint\cdots\oint\sum_{k_1}\cdots\sum_{k_g} c(k_1,\dots,k_g)t_1^{a_1+k_1-1}\cdots t_g^{a_g+k_g-1}\,dt_g\dots dt_1\\
&= \sum_{k_1}\cdots\sum_{k_g} c(k_1,\dots,k_g)\delta_{a_1+k_1}\cdots\delta_{a_g+k_g}\\
&= c(-a_1,\dots,-a_g)=c(a_1,\dots,a_g).
\end{aligned}$$

The generating function of the product moments (7) implies that $\mathbf{E}(t_1^{k_1}\dots t_g^{k_g})=0$ if $\sum k_i$ is odd or if $|k_{\sigma(1)}+\dots+k_{\sigma(t)}|>(2g-t+1)t$ for some permutation σ of the indices and some $t\le g$.

We note that the density function for $m\theta$ is

$$\frac{1}{m^g}\sum_{0\le j_1\le m-1}\cdots\sum_{0\le j_g\le m-1}\rho\left(\frac{\theta_1+2\pi j_1}{m},\dots,\frac{\theta_g+2\pi j_g}{m}\right)=c(0,0,\dots,0),$$

for $m\ge 2g+1$. In other words the vectors $(e^{im\theta_1},e^{im\theta_2},\dots,e^{im\theta_g})$ are equidistributed on the g-dimensional torus, once $m\ge 2g+1$. (Note that we observed this for $g=1$ in section 2).

In order to determine the mean of $(\alpha_1^r+\overline{\alpha}_1^r)+(\alpha_2^r+\overline{\alpha}_2^r)+\ldots+(\alpha_g^r+\overline{\alpha}_g^r)$ we can use our formula above for ρ: Because of its symmetry this equals $2g\mathbf{E}(\alpha_1^r)=2gq^{r/2}c(r,0,0,\dots,0)=2gq^{r/2}\mathbf{E}(\cos(r\theta_1))$, and so we might as well determine the distribution function for θ_1:

$$\begin{aligned}
Pr(a_1\le\theta_1\le b_1) &= \frac{1}{2\pi}\int_{a_1}^{b_1}\left\{\frac{1}{(2\pi)^{g-1}}\int_0^{2\pi}\cdots\int_0^{2\pi}\rho(\theta_1,\dots,\theta_g)d\theta_g\dots d\theta_2\right\}d\theta_1\\
&= \frac{1}{2\pi}\int_{a_1}^{b_1}\sum_{|k_1|\le 2g}c(k_1,0,0,\dots,0)t_1^{k_1}d\theta_1.
\end{aligned}$$

Now if $m\ge 2g+1$ then

$$\frac{1}{m^{g-1}}\sum_{0\le j_2\le m-1}\cdots\sum_{0\le j_g\le m-1}\rho\left(\theta_1,\frac{2\pi j_2}{m},\dots,\frac{2\pi j_g}{m}\right)=\sum_{|k_1|\le 2g}c(k_1,0,\dots,0)t_1^{k_1}.$$

Note that $\rho(.)=0$ if any $j_i=0$ or if any $j_i=\pm j_k \bmod m$. So, selecting $m=2g+1$ we find that the only non-zero values of ρ occur when $\{j_2,j_3,\dots,j_g\}=\{\pm1,\pm2,\dots,\pm g\}\setminus\{\pm i\}$, for some i. At such a point

$\rho = c_1 h_i(\theta_1)/g! h_i(2\pi i/m)$ where

$$h_i(\theta) = 4\sin^2\theta \prod_{1\leq j\leq g, j\neq i} (2\cos\theta - 2\cos(2\pi j/m))^2$$

and

$$c_1 = \prod_{1\leq j<k\leq g} (2\cos(2\pi j/m) - 2\cos(2\pi k/m))^2 \prod_{1\leq j\leq g} (2\sin^2(2\pi j/m))$$

From $2(\cos x - \cos y) = e^{-ix}(e^{ix} - e^{iy})(e^{ix} - e^{-iy})$ we deduce that (writing $\theta = \theta_1$ and $t = t_1$)

$$h_i(\theta) = -t^{-(m-1)}((t+1)(t^m-1)/(t-\zeta^i)(t-\zeta^{-i}))^2,$$

and thus $h_i(2\pi i/m) = m^2/(1-\zeta^i)(1-\zeta^{-i})$ where $\zeta = e^{2i\pi/m}$. Similarly $c_1 = (m/2)^g$. Since each possibility occurs $2^{g-1}(g-1)!$ times, we thus have

$$\begin{aligned}\sum_{|k|\leq 2g} c(k,0,0,\dots,0)t^k &= \frac{2^{g-1}c_1}{m^{g-1}g} \sum_{1\leq i\leq g} \frac{h_i(\theta)}{h_i(2\pi i/m)} \\ &= -\frac{t^{-(m-1)}}{2gm} \sum_{1\leq i\leq g} (1-\zeta^i)(1-\zeta^{-i}) \left(\frac{(t+1)(t^m-1)}{(t-\zeta^i)(t-\zeta^{-i})}\right)^2.\end{aligned}$$

Now $(\zeta^{1/2} - \zeta^{-1/2})(t+1)(t^m-1)/(t-\zeta)(t-\zeta^{-1}) = \sum_{j=0}^{m-1}(\zeta^{j+1/2} - \zeta^{-j-1/2})t^j$, and so the above becomes

$$\begin{aligned}&= \frac{t^{-(m-1)}}{4gm} \sum_{\xi:\ \xi^m=1} \sum_{j=0}^{m-1} (\xi^{j+1/2} - \xi^{-j-1/2})t^j \sum_{k=0}^{m-1} (\xi^{k+1/2} - \xi^{-k-1/2})t^k \\ &= \frac{t^{-(m-1)}}{4g} \sum_{0\leq j,k\leq m-1} t^{j+k} \frac{1}{m} \sum_{\xi:\ \xi^m=1} (\xi^{j+k+1} + \xi^{-j-k-1} - \xi^{j-k} - \xi^{k-j}) \\ &= \frac{t^{-(m-1)}}{2g}(mt^{m-1} - \sum_{0\leq j\leq m-1} t^{2j}) = \frac{1}{2g}\left(m - \left(\frac{t^m - t^{-m}}{t - t^{-1}}\right)\right) \\ &= \frac{1}{2g}\left(2g+1 - \frac{\sin((2g+1)\theta)}{\sin\theta}\right) = 1 - \frac{1}{g}\sum_{j=1}^{g} \cos(2j\theta).\end{aligned}$$

So we have proved

$$Pr(a \leq \theta_1 \leq b) = \frac{1}{2\pi}\int_a^b \frac{1}{2g}\left(2g+1 - \frac{\sin((2g+1)\theta)}{\sin\theta}\right) d\theta, \qquad (8)$$

generalizing the Sato-Tate measure. (Note that the case $g = 1$ is as in section 2 since $(1/2)(3 - \sin(3\theta)/\sin\theta) = 2\sin^2\theta$). Therefore,

$$2g\mathbf{E}(\cos(r\theta_1)) = \frac{1}{2\pi}\int_0^{2\pi} 2\cos(r\theta)\left(g - \sum_{j=1}^{g}\cos(2j\theta)\right)d\theta = -1$$

if $r = \pm 2, \pm 4, \ldots, \pm 2g$, $= 2g$ if $r = 0$, and $= 0$ otherwise. Thus, the result of Katz and Sarnak implies the following result:

Theorem 4.1. *Curves of genus g defined over $\mathbf{F}_q$ have, on average, $q^r + o(q^{r/2})$ points in $\mathbf{F}_{q^r}$, except if $r = 2j$ for some j in the range $1 \le j \le g$, in which case the average is $q^r + q^{r/2} + o(q^{r/2})$.*

Remark 1. In light of Deligne's proof of the Ramanujan-Petersson conjecture it is natural to conjecture that as C varies over all curves of genus g over $\mathbf{F}_q$ the average number of $\mathbf{F}_{q^r}$-rational points on C is $q^r + q^{r/2} + O(q^{(r-1)/2+\epsilon})$ if r is even and $r \le 2g$ and $q^r + O(q^{(r-1)/2+\epsilon})$ otherwise. Indeed this is essentially the Symplectic Higher Degree Excess Theorem of Katz [10].

The distribution functions of the eigenvalues of all of the compact classical groups lend themselves to analogous observations: $\mathrm{O}_-(2g+2)$ has the same distribution function as $\mathrm{USp}(2g)$. For $\mathrm{SO}(2g)$ we have the density function

$$\frac{1}{2^{g-1}g!}\prod_{1\le i<j\le g}(2\cos\theta_i - 2\cos\theta_j)^2,$$

in which case $m\theta$ are uniformly distributed on the g-dimensional torus once $m \ge 2g - 1$. The probability distribution function for each eigenvalue is

$$\frac{1}{2g}\left(2g - 1 + \frac{\sin((2g-1)\theta)}{\sin\theta}\right) = 1 + \frac{1}{g}\sum_{j=1}^{g-1}\cos(2j\theta)$$

so that $\mathbf{E}(\cos(r\theta_1)) = 1/2g$ if $r = 2, 4, \ldots, 2(g-1)$ and $= 0$ otherwise, for $r > 0$.

For $\mathrm{SO}(2g+1)$ we have the density function

$$\frac{1}{g!}\prod_{1\le i<j\le g}(2\cos\theta_i - 2\cos\theta_j)^2\prod_{1\le i\le g}(2\sin^2(\theta_i/2));$$

here the $m\theta$ are uniformly distributed on the g-dimensional torus once $m \geq 2g$. The probability distribution function for each eigenvalue is

$$1 - \frac{1}{2g}\frac{\sin(2g\theta)}{\sin\theta} = 1 - \frac{1}{g}\sum_{j=1}^{g}\cos((2j-1)\theta)$$

so that $\mathbf{E}(\cos(r\theta_1)) = -1/2g$ if $r = 1, 3, \ldots, 2g-1)$ and $= 0$ otherwise, for $r > 0$.

For $U(g)$ we have the density function

$$\frac{1}{g!}\prod_{1\leq j<k\leq g} |e^{i\theta_j} - e^{i\theta_k}|^2;$$

here the $m\theta$ are uniformly distributed on the g-dimensional torus once $m \geq g$. The probability distribution function for each eigenvalue is 1, that is, they are uniformly distributed, though this is not true of pairs of eigenvalues.

5. EXPERIMENTAL RESULTS

It is feasible that we will be able to deduce further trace formulae for higher genus curves. For example, we can try to use the Deligne Equidistribution Theorem though the problem will certainly be to understand the restrictions of monodromy in such families. For now we have some experimental results. It seems that, on average, $\#C(\mathbf{F}_{q^r})$, as C varies over curves of genus 2 defined over $\mathbf{F}_q$, is $q^r + 1$ for r odd, and

$$\begin{aligned} q^2 + q + 1 - 1/q^3 &\qquad \text{for } r = 2\\ q^4 + q^2 + 1 + 1/q + 1/q^2 - 1/q^3 &\qquad \text{for } r = 4\\ q^6 + 1/q &\qquad \text{for } r = 6\\ q^8 + 1/q + 1/q^2 - 1/q^3 &\qquad \text{for } r = 8\\ q^{10} + (1/q^2 + 1/q^3)\tau'(q) + 1 + 1/q - 1/q^3 &\qquad \text{for } r = 10\\ q^{12} - \tau'(q)/q^2 - 1 + 1/q + 1/q^2 &\qquad \text{for } r = 12 \end{aligned}$$

where $\tau'(q)$ is the modified Ramanujan tau function defined in section 3. As C varies over hyperelliptic curves of genus 3 we get $q^r + 1$ for r for odd, and

$$q^2 + q + 1 + 1/q^5 \qquad \text{for } r = 2$$

$$\begin{aligned} q^4+q^2+1/q+1/q^2-1/q^3+1/q^5 &\qquad \text{for } r=4\\ q^6+q^3+q+1/q+1/q^3+1/q^5 &\qquad \text{for } r=6\\ q^8-1+1/q+1/q^2-1/q^3+1/q^5 &\qquad \text{for } r=8\\ q^{10}+\tau'(q)/q^4+1-1/q^2+1/q^5 &\qquad \text{for } r=10 \end{aligned}$$

Here we are weighting by $1/|\mathrm{Aut}(C/\mathbf{F}_q)|$ as explained in the section 7.

6. EXTRA POINTS ON HYPERELLIPTIC CURVES IN $\mathbf{F}_{q^r}$

Hyperelliptic curves are a special case of cyclic covers of the projective line, for which we prove the following theorem.

THEOREM 6.1. *Let $k \geq 1$ and $f(x) \in \mathbf{F}_q[x]$. The average number of $\mathbf{F}_{q^r}$-rational affine points on the q^r curves $y^k = f(x)+g(x)$, where $g(x) \in \mathbf{F}_q[x]$ runs through all polynomials of degree less than r, is*

$$q^r + \sum_{m|r}((k, \frac{q^r-1}{q^m-1}) - 1)I_m(q)(1-q^{-m}).$$

Here $(\cdot,\cdot)$ is the greatest common divisor and $I_m(q) = \sum_{d|m}\mu(d)q^{m/d}$ where $\mu(d)$ is the Möbius μ function. By inclusion-exclusion or Möbius inversion $I_m(q)$ is the number of elements of $\mathbf{F}_{q^m}$ that are in no proper subfield containing $\mathbf{F}_q$, and therefore $I_m(q)/m$ is the number of monic irreducible polynomials of degree m over $\mathbf{F}_q$. Note that the average taken in Theorem 4.1 depends on q, r, and k but not $f(x)$, so we get the same result if we average over all polynomials of a given degree at least r.

Proof. The number of affine $\mathbf{F}_{q^r}$-rational points on the curve $y^k = f(x)+g(x)$ is

$$q^r + \sum_{x\in\mathbf{F}_{q^r}} \sum_{\chi:\ \chi^k=\chi_0,\ \chi\neq\chi_0} \chi(f(x)+g(x)).$$

Here χ is a multiplicative character of $\mathbf{F}_{q^r}$, and χ_0 is the trivial character. Let $\mathbf{F}'_{q^m}$ denote those elements of $\mathbf{F}_{q^m}$ that are in no subfield containing $\mathbf{F}_q$, and therefore $I_m(q) = |\mathbf{F}'_{q^m}|$. For a fixed χ the average of $\sum_{x\in\mathbf{F}_{q^r}}\chi(f(x)+g(x))$ over all g's is

$$\tfrac{1}{q^r} \sum_{x\in\mathbf{F}_{q^r}} \sum_{a_0,a_1,\ldots,a_{r-1}\in\mathbf{F}_q} \chi(f(x)+a_{r-1}x^{r-1}+\ldots+a_1x+a_0) \tag{9}$$

$$
\begin{aligned}
&= \tfrac{1}{q^r} \sum_{m|r} \sum_{x\in \mathbf{F}'_{q^m}} \sum_{a_0,a_1,\ldots,a_{r-1}\in \mathbf{F}_q} \chi(f(x)+a_{r-1}x^{r-1}+\ldots+a_1x+a_0)\\
&= \tfrac{1}{q^r} \sum_{m|r} \sum_{x\in \mathbf{F}'_{q^m}} \sum_{z\in \mathbf{F}_{q^m}} q^{r-m}\chi(z)\\
&= \sum_{m|r} I_m(q)q^{-m} \sum_{z\in \mathbf{F}_{q^m}} \chi(z)
\end{aligned}
$$

since $\mathbf{F}_{q^m} = \{a_{m-1}x^{m-1}+\ldots+a_0 : a_0,\ldots a_{m-1} \in \mathbf{F}_q\}$ and $f(x) + a_{r-1}x^{r-1}+\ldots a_m x^m \in \mathbf{F}_{q^m}$. Suppose χ has order exactly $j > 1$ which divides k. Every element of $\mathbf{F}_{q^m}$ is a jth power of an element of $\mathbf{F}_{q^r}$ if and only if $(q^m-1)|(q^r-1)/j$. Hence,

$$
\sum_{z\in \mathbf{F}_{q^m}} \chi(z) = \begin{cases} 0 & \text{if } j(q^m-1) \nmid q^r-1 \\ q^m-1 & \text{if } j(q^m-1) \mid q^r-1. \end{cases}
$$

The number of characters with order exactly j is $\phi(j)$ if $j|(q^r-1)$ and 0 otherwise where ϕ is the Euler ϕ-function. Thus, summing (4.1) over all $\chi \neq \chi_0$ we obtain

$$
\sum_{1<j|(k,q^r-1)} \phi(j) \sum_{\substack{m|r\\ j(q^m-1)|(q^r-1)}} I_m(q)(1-q^{-m}).
$$

Switching the order of summation and noting that $\sum_{j|n}\phi(j) = n$ finishes the proof. ▮

Corollary 6.1. *Let $f(x) \in \mathbf{F}_q[x]$. The average number of $\mathbf{F}_{q^r}$-rational affine points on the q^r curves $y^2 = f(x) + g(x)$, where $g(x) \in \mathbf{F}_q[x]$ runs through all polynomials of degree less than r, is q^r if q is even or r is odd. If q is odd and r is even then the average is*

$$
\begin{aligned}
q^r + q^{r/2} - \sum_{m|r/2} \frac{I_m(q)}{q^m} &= q^r + q^{r/2} - \tau(r/2) - \sum_{\substack{dt|(r/2)\\ t>1}} \frac{\mu(t)}{q^{d(t-1)}}\\
&= q^r + q^{r/2} - \tau(r/2) + O(1/q),
\end{aligned}
$$

where, here, $\tau(d)$ denotes the number of divisors of d.

Proof. In Theorem 6.1 with $k = 2$ if q is even q^r-1 is odd so the summation is 0. If q and r are odd $\frac{q^r-1}{q^m-1}$ is odd so again the summation is 0. If q is odd and r is even then $2(q^m-1)|(q^r-1)$ if and only if $m|r/2$. To conclude note that $\sum_{m|r/2} I_m(q)$ is the number of elements in $\mathbf{F}_{q^{r/2}}$, which is $q^{r/2}$, and the lead term of $I_m(q)$ is q^m. ▌

Corollary 6.2 is a corollary to both the Corollary 6.1 and the following proposition and will be needed in section 7 to prove Theorem 7.1. Corollary 6.1 and Proposition 6.1 for $r = 2$ support the asymptotic result in Theorem 4.1.

Proposition 6.1. *Let $f(x, y)$ be any function $\mathbf{F}_{q^2} \times \mathbf{F}_{q^2} \to \mathbf{F}_{q^2}$ such that image of the restriction $\mathbf{F}_q \times \mathbf{F}_q$ is in $\mathbf{F}_q$. (For example, $f(x, y) \in \mathbf{F}_q[x, y]$.) The average number of $\mathbf{F}_{q^2}$-rational affine points on the q^3 curves $f(x, y) + ax + by + c$, $a, b, c \in \mathbf{F}_q$, is $q^2 + q - 1$. The average number of $\mathbf{F}_q$-rational affine points is q.*

Proof. For $x, y \in \mathbf{F}_{q^2}$ and $a, b \in \mathbf{F}_q$ we have $c \in \mathbf{F}_q$ satisfying $f(x, y) + ax + by + c = 0$ if and only if $f(x, y) + ax + by \in \mathbf{F}_q$. Note that if $t \in \mathbf{F}'_{q^2}$ then $\mathbf{F}_{q^2} = \{mt + n : \ m, n \in \mathbf{F}_q\}$, so that if $g \in \mathbf{F}_{q^2}$ then there exists $m \in \mathbf{F}_q$ such that $g - mt \in \mathbf{F}_q$.

Thus, for each $x \in \mathbf{F}'_{q^2}, y \in \mathbf{F}_{q^2}, b \in \mathbf{F}_q$ there is a unique such a. Also for each $y \in \mathbf{F}'_{q^2}, a, x \in \mathbf{F}_q$ there is a unique such b. Finally if $x, y \in \mathbf{F}_q$ then any $a, b \in \mathbf{F}_q$ will do. Thus there is a total of $(q^2 - q)q^3 + (q^2 - q)q^2 + q^4 = q^3(q^2 + q - 1)$ on these q^3 curves. ▌

Corollary 6.2. *Assume q is odd and $d \geq 2$. The average number of $\mathbf{F}_{q^2}$-rational affine points on the q^d curves $y^2 = f(x)$, where $f(x) \in \mathbf{F}_q[x]$ runs through all polynomials of degree less than d, is $q^2 + q - 1$.*

7. AN EXACT FORMULA FOR THE AVERAGE NUMBER OF QUADRATIC POINTS

In this section we give an exact formula for the average number of $\mathbf{F}_{q^2}$-rational points on an hyperelliptic curve of genus g over $\mathbf{F}_q$ when q is odd.

Theorem 7.1. *Fix integer $g \geq 1$ and an odd prime power q. The weighted average number of $\mathbf{F}_{q^2}$-rational points on an hyperelliptic curve of genus g over $\mathbf{F}_q$ is exactly $q^2 + q + 1 - (-1)^g q^{1-2g}$.*

We begin by determining the number of rational points at ∞ on a given hyperelliptic curve, though leave the proof as an exercise.

Lemma 7.1. *There is exactly one $\mathbf{F}_q$-rational point on (4) at ∞ if the degree of f is odd. If the degree of f is even then let a be the leading* *coefficient of f: If a is a square in $\mathbf{F}_q$, then there are two $\mathbf{F}_q$-rational points on (4) at ∞, once we have resolved the singularity there. If a is not a square in $\mathbf{F}_q$, then there are no $\mathbf{F}_q$-rational points on (4) at ∞.*

In the introduction we defined our "weighted average", which we now motivate. The first step is the following lemma, which has appeared in [7, 5.1] and [11, 10.7.5]. For elliptic curves it appeared in [8, 2.2] but was known to Serre as early as 1984. Denote the algebraic closure of a field k by $\overline{k}$.

Lemma 7.2. *Let C be a curve (or a pointed curve) over $\overline{\mathbf{F}}_q$ with finite (geometric) automorphism group whose isomorphism class is* $\mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q)$-*invariant. Then*

$$\sum_X \frac{1}{|\mathrm{Aut}(X/\mathbf{F}_q)|} = 1$$

where the sum is over representatives of the $\mathbf{F}_q$-isomorphism classes of curves $X/\mathbf{F}_q$ that are $\overline{\mathbf{F}}_q$-isomorphic to C.

Let $N_n(X)$ denote the number of $\mathbf{F}_{q^n}$-rational points on X, and define

$$N_n = N_n(C) = \sum_X \frac{N_n(X)}{|\mathrm{Aut}(X/\mathbf{F}_q)|}.$$

(Note that we have abused the notation in that the meaning of N_n depends on whether the argument is a curve over $\mathbf{F}_q$ or $\overline{\mathbf{F}}_q$.) Ofer Gabber has sketched a proof that N_n is always an integer and has given a geometric interpretation of N_n.

Example 7.1. The curve

$$C:\ x^4+y^4+z^4+x^2y^2+x^2z^2+y^2z^2+x^2yz+xy^2z+xyz^2=0$$

is the Klein curve over $\mathbf{F}_2$, for which $\mathrm{Aut}(C/\overline{\mathbf{F}}_2)=\mathrm{Aut}(C/\mathbf{F}_2)=\mathrm{GL}_3(\mathbf{F}_2)$, the simple group of order 168. There are five other $\mathbf{F}_2$-isomorphism classes that are $\overline{\mathbf{F}}_2$-isomorphic to C, namely

$$\begin{aligned}
x^4+xy^3+xz^3+y^2z^2+x^2yz &= 0,\\
x^4+y^4+z^4+x^3y+y^3z+z^3x+x^2yz &= 0,\\
x^3y+y^3z+z^3x+x^2y^2+y^2z^2+x^3z &= 0,\\
x^4+xy^3+xz^3+y^2z^2+x^2yz+x^2y^2+x^3z &= 0,\\
x^3y+y^3z+z^3x &= 0,
\end{aligned}$$

with $\mathbf{F}_2$-automorphism groups of orders 8, 7, 7, 4, and 3, respectively; note that $1 = \frac{1}{168} + \frac{1}{8} + \frac{1}{7} + \frac{1}{7} + \frac{1}{4} + \frac{1}{3}$. (The fourth curve above is isomorphic to the curve of genus 3 with 7 points mentioned in [16].) These six curves have respective Weil polynomials

$$\begin{aligned}
z_1(t) &= 1 - 3t + 9t^2 - 13t^3 + 18t^4 - 12t^5 + 8t^6, \\
z_2(t) &= 1 + t + 5t^2 + 3t^3 + 10t^4 + 4t^5 + 8t^6, \\
z_3(t) &= 1 - 3t + 2t^2 + t^3 + 4t^4 - 12t^5 + 8t^6, \\
z_4(t) &= 1 + 4t + 9t^2 + 15t^3 + 18t^4 + 16t^5 + 8t^6, \\
z_5(t) &= 1 - t - t^2 + 3t^3 - 2t^4 - 4t^5 + 8t^6, \\
z_6(t) &= 1 + 5t^3 + 8t^6.
\end{aligned}$$

The generating function

$$\sum_{n=1}^{\infty} (2^n + 1 - N_n) t^n = -t \sum_{i=1}^{6} \frac{z_i'(t)}{z_i(t) |\mathrm{Aut}(C_i/\mathbf{F}_2)|} = \frac{g(t)}{\prod_{i=1}^{6} z_i(t)},$$

where $g(t) \in \mathbf{Z}[t]$ (and has degree 36), thus proving that in this case all N_n's are integers.

Katz and Sarnak [11, 10.7.4] define the intrinsic cardinality of the set $\mathcal{M}_g(\mathbf{F}_q)$ of all $\mathbf{F}_q$-isomorphism classes of curves X of genus g to be

$$\mathit{IntrinCard}(\mathcal{M}_g(\mathbf{F}_q)) = \sum_X \frac{1}{|\mathrm{Aut}(X/\mathbf{F}_q)|}.$$

(For $g = 1$ we should perhaps write "pointed curves of genus 1," throughout; that is, elliptic curves.) In light of Lemma 7.2 we could define this intrinsic cardinality to be the number of $\overline{\mathbf{F}}_q$-isomorphism classes of curves C of genus g that are $\mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q)$-invariant. Likewise we could define the intrinsic cardinality of any $\mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q)$-invariant subset, such as $\mathcal{H}_g(\mathbf{F}_q)$ of all $\mathbf{F}_q$-isomorphism classes of hyperelliptic curves of genus g. By Lemma 7.2 intrinsic cardinalities are indeed integers. By "the average number of $\mathbf{F}_{q^n}$-rational points on an hyperelliptic curve of genus g over $\mathbf{F}_q$" we shall mean

$$\frac{\sum_C N_n(C)}{\mathit{IntrinCard}(\mathcal{H}_g(\mathbf{F}_q))}$$

where the sum is over all $\overline{\mathbf{F}}_q$-isomorphism classes of genus g hyperelliptic curves C that are $\mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q)$-invariant. In other words, it is the average number of points on such C, where we weight each curve X by $1/|\mathrm{Aut}(X/\mathbf{F}_q)|$.

Our next task is to compute the denominator, *IntrinCard*$(\mathcal{H}_g(\mathbf{F}_q))$. Lemma 7.3 appears in [6, Proposition 13], is alluded to in [11, 10.5.12], and was known to Michael Larsen as early as 1984.

Proposition 7.1. *IntrinCard*$(\mathcal{H}_g(\mathbf{F}_q)) = q^{2g-1}$ *for any* $g \geq 1$*.*

To prove this we will need the following lemma.

Lemma 7.3. *There are exactly* $q^d - q^{d-1}$ *monic squarefree polynomials of degree* $d > 1$ *in* $\mathbf{F}_q[x]$*.*

Proof. Let n_d be the number of monic squarefree polynomials of degree d. Clearly the number of monic polynomials of degree d is q^d. Any monic $f(x) \in \mathbf{F}_q[x]$ can be written uniquely as $f(x) = g(x)k(x)^2$ where $g(x)$ is monic and squarefree and $k(x)$ is monic. Counting the number of possible polynomials on both sides of the equation we see that

$$q^d = n_d + n_{d-2}q + n_{d-4}q^2 + \cdots .$$

Subtracting q times this from the equation for $d+2$ gives $n_{d+2} = q^{d+2} - q^{d+1}$, as required. ▮

Proof of Proposition 7.1. We shall only prove the result when q is odd, since we only apply it then (and since our proof is too long when q is even).

Every hyperelliptic curve over $\mathbf{F}_q$ of genus g is isomorphic to a curve of the form $y^2 = f(x)$ where $f(x)$ is squarefree of degree $2g+2$ or $2g+1$. By Lemma 7.3 there are $(q^{2g+2} - q^{2g+1}) + (q^{2g+1} - q^{2g}) = q^{2g+2} - q^{2g}$ such monic f; and so there are $(q^2-1)(q^2-q)q^{2g-1}$ such f. To determine the number of curves isomorphic to a given curve, note that all isomorphisms between such curves are given by a linear fractional map on x, and a scalar multiple of y. There are $(q^2-1)(q^2-q)/(q-1)$ elements of $\mathrm{PGL}_2(\mathbf{F}_q)$ and $(q-1)$ scalar multiples, giving a total of $(q^2-1)(q^2-q)$ such isomorphic curves. ▮

Thus *IntrinCard*$(\mathcal{M}_2(\mathbf{F}_q)) = q^3$, and we conjecture that

$$\mathit{IntrinCard}(\mathcal{M}_3(\mathbf{F}_q)) = q^6 + q^5 + 1.$$

Proof of Theorem 7.1 Let P_d be the set of monic squarefree polynomials in $\mathbf{F}_q[x]$, and $n_d = |P_d|$ which equals $q^d - q^{d-1}$ for $d \geq 2$ by Lemma 7.3. Let r_d be the proportion of $f \in P_d$ for which $f(0) = 0$. Note that $f \in P_d$ with $f(0) = 0$ if and only if we can write $f(x) = xg(x)$ where $g \in P_{d-1}$ with $g(0) \neq 0$, so that

$$n_d r_d + n_{d-1} r_{d-1} = n_{d-1}.$$

Let $a_{d,e}$ be the average number of affine $\mathbf{F}_{q^2}$ points on the curves $y^2 = f(x)g(x)^2$, where $f \in P_d$ and $g(x) \in \mathbf{F}_q[x]$ is monic of degree e; and write $a_d = a_{d,0}$. Note that, for a given a, we have the same number of solutions to $y^2 = f(a)$ as to $y^2 = f(a)g(a)^2$ if $g(a) \neq 0$. Now, the proportion of the degree e monics for which $g(a) = 0$ is $1/q$ if $e \geq 1$ and 0 if $e = 0$ when $a \in \mathbf{F}_q$; whereas it is $1/q^2$ if $e \geq 2$ and 0 if $e = 0$ or 1 when $a \in \mathbf{F}'_{q^2}$. Thus we deduce that

$$a_{d,e} = a_{d,e-1} \text{ when } e \geq 3.$$

Also $a_{d,1} - a_{d,0}$ is the mean value of $1 - \#\{y \in \mathbf{F}_{q^2} : \ y^2 = f(a)\}$, over $f \in P_d$ and $a \in \mathbf{F}_q$. As $a \in \mathbf{F}_q$ thus $f(a) \in \mathbf{F}_q$, and so there are two square roots of $f(a)$ in $\mathbf{F}_q$ unless $f(a) = 0$. Letting $x \to x - a$ we deduce that

$$a_{d,1} - a_{d,0} = r_d - 1.$$

Finally $a_{d,2} - a_{d,1}$ is $I_2(q)/q^2$ times the mean value of $1 - \#\{y \in \mathbf{F}_{q^2} : y^2 = f(a)\}$, varying over $f \in P_d$ and $a \in \mathbf{F}'_{q^2}$. This equals the mean value of $1 - \#\{y \in \mathbf{F}_{q^2} : \ y^2 = f(a)\}$, varying over $f \in P_d$ and $a \in \mathbf{F}_{q^2}$, minus $(1/q)$ times the mean value varying only over $a \in \mathbf{F}_q$. Combining this with the previous paragraph gives

$$a_{d,2} - a_{d,1} = 1 - a_d/q^2 - (r_d - 1)/q.$$

By Corollary 6.2 we have, for $d \geq 2$,

$$(q^2 + q - 1)q^d = \sum_{e=0}^{\lfloor d/2 \rfloor} a_{d-2e,e} n_{d-2e} q^e. \tag{10}$$

For $d \geq 4$ we subtract q times $(10)_{d-2}$ away from $(10)_d$ to get the right hand side

$$a_d n_d + q n_{d-2}(a_{d-2,1} - a_{d-2,0}) + q^2 n_{d-4}(a_{d-4,2} - a_{d-4,1}).$$

Using the equations above we deduce that for $d \geq 6$,

$$a_d - a_{d-4}/q^4 = q^2 + q - 1 + 1/q - 2/q^2. \tag{11}$$

The average we are looking for is $\mu_g = ((a_{2g+1} + 1) + q(a_{2g+2} + 2))/(q + 1)$, which takes account of the points at ∞, as explained in Lemma 7.1. Adding $(11)_{2g+1}$ to q times $(11)_{2g+2}$ we deduce

$$(\mu_g - (q^2 + q + 1)) = \frac{1}{q^4}(\mu_{g-2} - (q^2 + q + 1))$$

for all $g \geq 3$. Using the above equations to evaluate μ_g for $g = 0, 1, 2$ we then can prove by induction that $\mu_g - (q^2 + q + 1) = -(-1)^g/q^{2g-1}$. ▮

ACKNOWLEDGMENTS

The first author would like to thank Michael Zieve for observing that curves with many points over a finite field often could be written over the prime field, which points toward some of these results, and for connecting him with the second author. He also thanks Nick Katz, Ofer Gabber, Noam Elkies, and Peter Sarnak for generously answering numerous questions by email, and Kevin Iga for a fruitful conversation. The first author used Magma to perform the numerical experiments in section 6 while he was at the Center for Communications Research in Princeton. The second author would like to thank William Stein for a discussion that led to this paper, and Tom Tucker for useful conversations.

REFERENCES

1. B.J. Birch, "How the number of points of an elliptic curve over a fixed prime field varies", *J. London Math. Soc.* **43** (1968), 57-60.
2. C. Breuil, B. Conrad, F. Diamond, R. Taylor, "On the modularity of elliptic curves over **Q**", *http://math.harvard.edu/∼rtaylor/st.dvi*, preprint.
3. D. Bump, D. Ginzburg, J. Hoffstein, "The symmetric cube", *Invent. Math.*, **125** (1996), 413-449.
4. P. Deligne, "Formes modulaires et representations l-adiques", *Seminaire Bourbaki, Vol. 1968/69, Exp. 355, Lecture Notes in Math.* **179**, Springer-Verlag, New York, 1971.
5. P. Deligne, "La conjecture de Weil I", *Inst. Hautes Études Sci. Publ. Math.* **43**(1974), 273-307.
6. P. Fleischmann, I. Janiszczak, R. Knôrr, "The number of regular semisimple classes of special linear and unitary groups", *Linear Algebra Appl.*, **274**(1998)17-26.
7. G. van der Geer, M. van der Vlugt, "Supersingular curves of genus 2 over finite fields of characteristic 2", *Math. Nachr.* **159**(1992), 73-81.
8. E.W. Howe, "On the group orders of elliptic curves over finite fields", *Compositio Math.* **85**(1993), 229-247.
9. P.I. Katsylo, "Rationality of the moduli spaces of hyperelliptic curves", *Izv. Akad. Nauk SSSR Ser. Mat.*, **48**(1984), 705-710.
10. N.M. Katz, "Frobenius-Schur indicator and the ubiquity of Brock-Granville quadratic excess", (to appear).
11. N.M. Katz, P. Sarnak, " Random Matrices, Frobenius Eigenvalues, and Monodromy", *AMS Colloquium Publications* vol. 45, Providence, Amer. Math. Soc., 1999.
12. R. Schoof, "Nonsingular plane cubic curves over finite fields" *J. Combin. Theory Ser. A*, **46** (1987), 183-211.
13. A. Selberg, "Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series", *J. Indian Math. Soc.* **20** (1956), 47-88.
14. A. Selberg, "On the estimation of Fourier coefficients of modular forms", *Proc. Amer. Math. Soc., Symp. Pure Math. VIII: Theory of Numbers*, Pasadena, 1963, 1-15.
15. J.-P. Serre, "Abelian l-adic Representations and Elliptic Curves", New York, Benjamin, 1968, 1-15.

16. J.-P. Serre, "Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini", *C. R. Acad. Sci. Paris Sér. I Math.*, **296** (1983), 397-402; *Oeuvres. Collected Papers*, Vol. III 1972-1984, Springer-Verlag, New York, 1986, pp. 658-663.

17. G. Shimura, "On the holomorphy of certain Dirichlet series", *Proc. London Math. Soc.* (3) **31**(1975) 79-98.

18. H. Weyl, "Classical Groups", Princeton U. Press, Princeton, 1946.

19. A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141** (1995), 443-551.

20. D. Zagier, "The Eichler-Selberg trace formula on $SL_2(\mathbf{Z})$", Appendix to S. Lang's Introduction to Modular Forms, *Grundlehren der Mathematischen Wissenschaften* vol. 222, Springer-Verlag, New York, 1976, 44-54.