Hecke operators with odd determinant and binary frameproof codes beyond the probabilistic bound?

Hugues Randriam
Telecom ParisTech – LTCI CNRS UMR 5141
46, rue Barrault – 75634 Paris Cedex 13 – France

***Abstract*—We give a slight improvement on Xing's lower bound for frameproof codes constructed from algebraic curves. Combined with some additional number-theoretic assumptions (still conjectural) and a concatenation process, this should lead to the existence of a family of binary 2-frameproof codes of asymptotic rate going beyond the up to now best known (non-constructive) lower bound.**

I. Introduction

In [10], Körner considers the following problem:

How many different points can we find in the n-dimensional Hamming space so that no three of them are on a line?

(We say three points in a metric space are on a line if they satisfy the triangle inequality with equality. Recall also the n-dimensional Hamming space is the set of length n binary sequences, the Hamming distance between two sequences being the number of coordinates in which they differ.)

Denote by $M(n)$ the maximal size of such a configuration, and define its asymptotic exponent

$$\rho = \limsup_{n\to\infty} \frac{\log_2 M(n)}{n}.$$

It is shown in [10] that ρ satisfies the inequalities

$$1 - \frac{1}{2}\log_2(3) \leq \rho \leq \frac{1}{2}$$

where the derivation of the lower bound

$$1 - \frac{1}{2}\log_2(3) \approx 0.207518\ldots$$

is a typical example of the power of the so called *probabilistic method* in extremal combinatorics.

This said, the gap between these two bounds is very big, and the problem of narrowing it is still open. One of the main results of this paper is that the lower bound could be (slightly!) improved, under some additional number-theoretic assumption whose validity is unfortunately still conjectural:

Conjecture A: Let $p > 2$ be an odd prime. Does there exist infinitely many prime numbers N such that the Hecke operator T_p acting on the space of weight 2 cusp forms $S_2(\Gamma_0(N))$ has odd determinant?

Theorem 1: If Conjecture A holds true for $p = 11$, then the asymptotic exponent ρ satisfies the lower bound

$$\rho \geq \frac{3}{50}\log_2(11) \approx 0.207565\ldots$$

The reader might wonder how our initial problem come to be related to such a question from the theory of modular forms. In fact, as noted in [10], estimating $M(n)$ is an old problem, and different equivalent versions had been considered by people working in various areas. Its first occurence can be traced to the paper [6], dealing with state assignment for asynchronous automata. In this context, our configurations of points (no three on a line) were known as $(2,1)$-separating systems — see the recent survey [15] for more on this theory and its history. Later, $(s,1)$-separating codes were re-introduced under the name of "s-frameproof codes" in a cryptographic context, namely for digital fingerprinting purposes, and more generally in relation with the theory of traitor-tracing schemes (see [1] and [17]).

Frameproof codes were then extensively studied, and many constructions were given. In [19] (reproduced in [14]), Xing gives a criterion for an algebraic-geometry (AG) code to be frameproof. Using this criterion, and a point-counting argument on the Jacobian of the curve on which the code is defined, he derives an asymptotic lower bound for the achievable rate of such codes. In the next section, we look carefully and rephrase some steps of this counting argument. In some cases, this allows us to improve slightly on Xing's bound.

We then formulate some conjectures which lead to a much bigger improvement. Loosely speaking, what we require is the existence of curves having many points but whose class groups have few s-torsion elements. For example, the modular curves $X_0(N)$ are asymptotically maximal for the number of points ([9][18]), and have class number given by some determinant of Hecke operators, so Conjecture A fits in this framework. We will see some alternatives, weaker or stronger. The strongest one, is that some family of maximal curves (e.g. modular curves, or the tower in [8]) should satisfy the Cohen-Lenstra heuristics ([4][7]). Although much stronger than what we need, this is a natural assumption that looks very plausible.

From these conjectures, one deduces a new lower bound for frameproof codes over "big" fields, and, using a concatenation process, this leads to the lower bound for *binary* codes of Theorem 1. As the reader noticed, this improvement over the probabilistic bound is extremely small, so it really seems to rely on the validity of our conjectures. A bare use of Xing's original result would not be sufficient.

In the final section of this paper, arguments supporting these conjectures and possible strategies of proof are then discussed, as well as constructiveness issues.

II. On Xing's lower bound

A. Frameproof codes

If n, M and q are three integers, we define an $(n, M)_q$-code as any subset $C \subset Q^n$ of size $|C| \geq M$, where Q is a finite set (the "alphabet") of size $|Q| = q$. Elements of C are called "codewords". The rate of C is defined as the real number $\frac{\log_q |C|}{n} \geq \frac{\log_q M}{n}$.

If q is a prime power, we denote by $\mathbb{F}_q$ the field with q elements, and if n, k are two integers, we define an $[n, k]_q$-(linear) code as any vector subspace $C \subset (\mathbb{F}_q)^n$ of dimension $\dim_{\mathbb{F}_q} C \geq k$. Clearly an $[n, k]_q$-code is then also an $(n, q^k)_q$-code, with rate at least $\frac{k}{n}$.

If $C \subset Q^n$ is a code and $S \subset C$ is any set of codewords, we define the set $\mathrm{desc}(S)$ of descendants of S, as the set of all $w \in Q^n$ such that for all $1 \leq i \leq n$, there is at least one $c \in S$ having same i-th coordinate as w. For any integer $s \geq 2$, one says that C is s-frameproof if for any $S \subset C$ of size $|S| \leq s$, the intersection $\mathrm{desc}(S) \cap (C \setminus S)$ is empty.

An $(n, M)_q$-code (resp. an $[n, k]_q$-code) that is also s-frameproof will be denoted an s-FPC$(n, M)_q$ (resp. an s-FPC$[n, k]_q$). We define $M_q(n, s)$ (resp. $k_q(n, s)$) as the maximum M (resp. the maximum k) for which there exists an s-FPC$(n, M)_q$ (resp. an s-FPC$[n, k]_q$). We also define the asymptotic rates

$$R_q(s) = \limsup_{n\to\infty} \frac{\log_q M_q(n, s)}{n}$$

$$R_q^{\mathrm{lin}}(s) = \limsup_{n\to\infty} \frac{k_q(n, s)}{n}.$$

B. Curves and their Jacobians

We fix q a non-trivial power of some prime number p. If X is an algebraic curve (smooth, projective, absolutely irreducible 1-dimensional scheme) over $\mathbb{F}_q$, we denote by $g = g(X)$ its genus and $J = J(X)$ its Jacobian. We suppose given a point $P_0 \in X(\mathbb{F}_q)$, we let $j = j_{P_0}$ be the corresponding Abel-Jacobi embedding of X in J, $j(P) = [P - P_0]$. We let $W_1 = j(X)$ be its image and, for any r, W_r the r-fold sum of W_1 with itself (*i.e.* the image of X^r under the r-fold Abel-Jacobi map $j^{(r)}$). These W_r form an increasing sequence of closed subvarieties of J. In particular, W_g is the whole of J, while $W_{g-1} = \Theta$ is its theta divisor. By convention we set $W_0 = \{0_J\}$ and $W_r = \emptyset$ for $r < 0$.

We say that a divisor on X is a *simple effective divisor* (s.e.d.) if it is a (formal) sum of pairwise distinct $\mathbb{F}_q$-points of X, each with multiplicity 1. From now on we choose such a s.e.d. G, let $n = \deg G \leq |X(\mathbb{F}_q)|$ its degree, and choose an ordering $P_1, \ldots, P_n$ of the points in its support, so

$$G = P_1 + \cdots + P_n$$

(these points are pairwise distinct, although one of them maybe equal to P_0). We also denote by κ_G its image in J,

$$\kappa_G = j^{(n)}(G) = [G - nP_0].$$

Remark that on X there is a biggest s.e.d., namely the sum of all points in $X(\mathbb{F}_q)$. We will denote it by G^*.

For any integer $s \geq 2$, we denote by $s\cdot(J(\mathbb{F}_q))$ the image of the multiplication-by-s map acting on the group of $\mathbb{F}_q$-points of J, and by $J(\mathbb{F}_q)[s]$ its kernel, *i.e.* the s-torsion subgroup of the class group of X.

Definition 2: For $s \geq 2$, let the *s-frameproof Xing number* of (X, G) be the biggest integer $x_s = x_s(X, G)$ such that the translate $s \cdot (J(\mathbb{F}_q)) - \kappa_G$ is not contained in W_{x_s}.

Remark $x_s < g$, since $W_g = J$.

Later on we will need normalized versions of the quantities just introduced. So we define:

$$\alpha(X) = \frac{|X(\mathbb{F}_q)|}{g} \qquad \delta_s(X) = \frac{\log_s |J(\mathbb{F}_q)[s]|}{g}$$

$$\nu(X, G) = \frac{n}{g} = \frac{\deg G}{g} \qquad \xi_s(X, G) = \frac{x_s(X, G)}{g}.$$

Proposition 3: These quantities satisfy:

(i) $\nu(X, G) \leq \alpha(X)$

(ii) $\xi_s(X, G) < 1$

(iii) if $s = s_1 s_2$ with s_1, s_2 relatively prime, then

$$\delta_s(X) = \delta_{s_1}(X) \log_s s_1 + \delta_{s_2}(X) \log_s s_2$$

(iv) $0 \leq \delta_s(X) \leq 2$ for any s

(v) $0 \leq \delta_s(X) \leq 1$ if s a power of p.

Proof: For (i), remark $n \leq |X(\mathbb{F}_q)|$ by definition. We know $x_s < g$, hence (ii). Then (iii) is the Chinese remainder theorem, while (iv)–(v) follow from [13] p. 39 (or p. 64). ■

C. AG codes

Given a ($\mathbb{F}_q$-rational) divisor D on X, we denote by $\mathcal{O}(D)$ the associated invertible sheaf, $\mathcal{L}(D)$ its space of global sections, and $l(D) = \dim_{\mathbb{F}_q} \mathcal{L}(D)$. The "Riemann" part of the Riemann-Roch theorem asserts that $l(D) \geq \deg D + 1 - g$, with equality when $\deg D \geq 2g - 1$. Also, $l(D) > 0$ if and only if $[D - (\deg D)P_0]$ lies in $W_{\deg D}$.

For each i, choose a local parameter t_i at P_i. From this choice one deduces a trivialization of the 1-dimensional vector space $\mathcal{O}(D)|_{P_i} \simeq \mathbb{F}_q$. We then define a linear code $C(G, D) \subset \mathbb{F}_q^n$ by evaluation at the P_i (relatively to these trivializations) of the functions in $\mathcal{L}(D)$. As noted by Xing, this construction generalizes Goppa's evaluation codes, while allowing the supports of G and D to overlap (in fact Xing's original definition also asked D to be positive, but this condition is clearly unnecessary). He then proves:

Theorem 4 (Xing's criterion, [19] Th. 3.5): Suppose $\deg D < n$ and

$$l(sD - G) = 0.$$

Then $C(G, D)$ is an s-FPC$[n, l(D)]_q$.

Corollary 5: Suppose $n \geq g$. Then there exists an

$$s\text{-FPC}[n, \lfloor \tfrac{n+x_s}{s} \rfloor + 1 - g]_q.$$

Proof: Set $d = \lfloor \frac{n+x_s}{s} \rfloor$. Then, from $n \geq g > x_s$ and $s \geq 2$, one deduces $d < n$. On the other hand, by construction $sd - n \leq x_s$, which means there is a degree 0 divisor D_0 such that $s\cdot[D_0] - \kappa_G$ is not in W_{sd-n}. This implies that the degree d divisor $D = D_0 + dP_0$ satisfies the hypotheses in Xing's

criterion. Thus $C(G,D)$ is frameproof, and its dimension is $l(D) \geq d+1-g$ by Riemann-Roch. ∎

This corollary motivates the search for lower bounds on x_s:

Proposition 6: Set $c(q) = 1 + \log_q \frac{(3\sqrt{q}-1)}{(q-1)(\sqrt{q}-1)}$. Then

$$x_s \geq \left\lfloor g - \log_q |J(\mathbb{F}_q)[s]| - \log_q g - c(q) \right\rfloor.$$

Proof (compare with [19], Lemmas 3.7–3.10): To ease notations, let $\mathcal{H} = J(\mathbb{F}_q)$ and $h = |\mathcal{H}|$. By the rank theorem, $s \cdot \mathcal{H}$ is a subgroup of index $|\mathcal{H}[s]|$ in $\mathcal{H}$. Thus

$$\log_q |s \cdot \mathcal{H} - \kappa_G| = \log_q |s \cdot \mathcal{H}| = \log_q h - \log_q |\mathcal{H}[s]|.$$

On the other hand, for $0 \leq r < g$, [19], Lemma 3.9 (proved in [14], Lemma 2.24) gives

$$\log_q |W_r(\mathbb{F}_q)| < \log_q h + r - g + \log_q g + c(q).$$

Now, if $s \cdot \mathcal{H} - \kappa_G$ is to be contained in W_r, it will in fact be contained in $W_r(\mathbb{F}_q)$. This implies $|s \cdot \mathcal{H} - \kappa_G| \leq |W_r(\mathbb{F}_q)|$ hence

$$r > g - \log_q |\mathcal{H}[s]| - \log_q g - c(q)$$

and the lower bound on x_s follows. ∎

Remark that this lower bound on x_s does not depend on G.

D. Asymptotics

Say that a sequence of curves X_k over $\mathbb{F}_q$ form an ∞-sequence if $g(X_k)$ tends to infinity as k tends to infinity.

Let $A(q)$ be the *largest* real number such that there exists an ∞-sequence of curves X_k over $\mathbb{F}_q$ with

$$\alpha(X_k) \underset{k\to\infty}{\longrightarrow} A(q).$$

Let $\delta_s^-(q)$ be the *smallest* real number such that there exists an ∞-sequence of curves X_k over $\mathbb{F}_q$ satisfying:

$$\begin{cases} \alpha(X_k) \longrightarrow A(q) \\ \delta_s(X_k) \longrightarrow \delta_s^-(q). \end{cases}$$

Let $\xi_s^+(q)$ be the *biggest* real number such that there exists an ∞-sequence of curves X_k over $\mathbb{F}_q$, and G_k s.e.d. on X_k, satisfying:

$$\begin{cases} \nu(X_k, G_k) \longrightarrow A(q) \\ \xi_s(X_k, G_k) \longrightarrow \xi_s^+(q). \end{cases}$$

With these notations:

Theorem 7: Let q be a prime power. Then:

(i) if $A(q) > 1$,

$$R_q^{\text{lin}}(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{\xi_s^+(q)}{sA(q)}$$

(ii) $\xi_s^+(q) \geq 1 - \delta_s^-(q) \log_q s$.

Proof: Inequality (i) is just the asymptotic version of Corollary 5: consider an ∞-sequence of curves X_k over $\mathbb{F}_q$, and G_k s.e.d. on X_k, such that $\nu(X_k, G_k) \longrightarrow A(q) > 1$ and $\xi_s(X_k, G_k) \longrightarrow \xi_s^+(q)$. Then for k big enough one can apply Corollary 5 to (X_k, G_k), which gives an s-frameproof code of rate lower bounded by

$$\frac{1}{n}\left(\left\lfloor \frac{n+x_s}{s} \right\rfloor + 1 - g\right) \geq \frac{1}{n}\left(\frac{n+x_s}{s} - g\right)$$

(where $n = \deg G_k$, $x_s = x_s(X_k, G_k)$, $g = g(X_k)$). Now when k tends to infinity, this last quantity tends to

$$\frac{1}{A(q)}\left(\frac{A(q) + \xi_s^+(q)}{s} - 1\right)$$

so (i) follows.

In the same way, inequality (ii) is the asymptotic version of Proposition 6, applied to any ∞-sequence of curves X_k satisfying $\alpha(X_k) \longrightarrow A(q)$ and $\delta_s(X_k) \longrightarrow \delta_s^-(q)$, by choosing $G_k = G_k^*$ the maximal s.e.d. on X_k (hence $\nu(X_k, G_k) = \alpha(X_k)$). ∎

Combining (i) and (ii) gives:

Corollary 8: Suppose $A(q) > 1$. Then:

$$R_q^{\text{lin}}(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{1 - \delta_s^-(q) \log_q s}{sA(q)}.$$

For this to be useful, we need upper bounds on $\delta_s^-(q)$. For example, Proposition 3 (iv) gives $\delta_s^-(q) \leq 2$, and this is essentially how in [19] Xing derives: $R_q^{\text{lin}}(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{1-2\log_q s}{sA(q)}$. We can do slightly better, as follows:

Proposition 9: Let v_p denote the p-adic valuation, so $q = p^{v_p(q)}$, and write $s = p^{v_p(s)} s'$ (so s' is prime to p). Then

$$\delta_s^-(q) \;\leq\; 2\log_s s' + \log_s p^{v_p(s)} \;=\; 2 - \frac{v_p(s)}{v_p(q)} \log_s q.$$

Proof: Follows from Proposition 3 (iii)–(v). ∎

Corollary 10: With the same notations, if $A(q) > 1$, then

$$R_q^{\text{lin}}(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{1 - 2\log_q s + \frac{v_p(s)}{v_p(q)}}{sA(q)}.$$

Remark that when s is prime to p, this gives again Xing's lower bound. When p divides s, it improves on it by $\frac{v_p(s)}{sA(q)v_p(q)}$.

III. Conjectures and application to $R_2(2)$

We would like to improve still further on the previous results, but for this we will rely on one of the following two conjectures, which should hold for any prime power q:

Conjecture B-1: $\delta_s^-(q) = 0$.

Conjecture B-2: $\xi_s^+(q) = 1$.

Proposition 11: Conjecture B-1 implies Conjecture B-2, which in turn implies

$$R_q^{\text{lin}}(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{1}{sA(q)}.$$

Proof: Use Proposition 3 (ii) and Theorem 7. ∎

Recall ([5]) that if C_1 is an $(n_1, M_1)_{q_1}$-code and C_2 an $(n_2, M_2)_{q_2}$-code with $q_1 \leq M_2$, then from any choice of mapping Q_1 into C_2 (where Q_1 is the alphabet of C_1) one deduces a "concatenated code" $C_1 \circ C_2$, which will be an $(n_1 n_2, M_1)_{q_2}$-code. The rate of this code is $\frac{r_1 \log_{q_2} q_1}{n_2} \leq r_1 r_2$ where r_i is the rate of C_i. One then easily checks:

Lemma 12: If C_1 and C_2 are s-FPC, then so is $C_1 \circ C_2$.

From this one deduces:

Lemma 13: If C_2 is an s-FPC$(n_2, M_2)_{q_2}$ then, for any integer $q_1 \leq M_2$,

$$R_{q_2}(s) \geq \frac{\log_{q_2} q_1}{n_2} R_{q_1}(s).$$

Proof: Let Q_1 be an alphabet of size q_1, take C_1 in a family of s-FPC over Q_1 of asymptotic rate $R_{q_1}(s)$, and concatenate with C_2. ∎

Theorem 14: If Conjecture B-1 (or B-2) is true for $p = 11$, $q = p^2$, $s = 2$, then

$$R_2(2) \geq \frac{3}{50}\log_2(11) \approx 0.207565\ldots$$

Proof: From [11] we know the one-shortened Kerdock code $K(2^m - 1, 2^{2m-1})$ for $m \geq 4$ is 2-frameproof. We let C_2 be this code for $m = 4$, hence C_2 is a 2-FPC$(15, 128)_2$. Applying the lemma with $q_1 = q = 121$ will give

$$R_2(2) \geq \frac{2\log_2 11}{15} R_{121}(2).$$

On the other hand, for $q = p^2$, we know $A(q) = \sqrt{q} - 1 = 10$ ([9][18]). Proposition 11 then gives

$$R_{121}(2) \geq R^{\text{lin}}_{121}(2) \geq \frac{1}{2} - \frac{1}{10} + \frac{1}{20} = \frac{9}{20} = 0.45$$

which gives the result. ∎

As noted before, it is known that $R_2(2)$ satisfies

$$\frac{1}{2} \geq R_2(2) \geq 1 - \frac{1}{2}\log_2(3) \approx 0.207518\ldots$$

so our (conjectural) Theorem gives a slight improvement on the lower bound. Remark that Xing's lower bound in [19] gives

$$R_{121}(2) \geq \frac{1}{2} - \frac{1}{10} + \frac{1 - 2\log_{121} 2}{20} \approx 0.435546\ldots$$

hence only

$$R_2(2) \geq \frac{3}{50}\log_2(11) - \frac{1}{150} \approx 0.200899\ldots$$

One should also remark that, although Proposition 11 deals with linear codes over $\mathbb{F}_q$, the binary codes we obtain after concatenation are nonlinear. By comparison, for a linear code, being 2-frameproof is equivalent to being an *intersecting* code, hence ([12][3]):

$$x^* \geq R^{\text{lin}}_2(2) \geq 1 - \frac{1}{2}\log_2(3)$$

where the upper bound $x^* \approx 0.283476$ is solution of $x = H(\frac{1}{2} - \sqrt{x(1-x)})$ (if the Varshamov-Gilbert bound were exact, one could even replace x^* with $x^{**} \approx 0.227092$ solution of $x = 1 - H(x)$). So, as expected, the upper bound in the linear case is smaller, and in fact much smaller, than in the non-linear case. On the other hand, it is quite remarkable that the (non-conjectural but also non-constructive) lower bounds for $R_2(2)$ and $R^{\text{lin}}_2(2)$ are the same.

IV. More about the conjectures

A. Conjecture B-1 and B-2 and constructiveness

First we deal with the weaker of these two conjectures, which is B-2 (see Proposition 11). What makes Conjecture B-2 plausible is that, on any algebraic curve X, a "generic" divisor B of degree $b = \deg B < g$ has $l(B) = 0$, or equivalently, the image of B in the Jacobian does not lie in W_b.

One would like to apply this idea with B of the form $B = sD - G$ with the notations of Theorem 4. Of course this does not work because such a B is not "generic". However, by letting the curve X and the divisor G vary, and taking b a little smaller than g, we see there are many degrees of freedom. Remark that one wants G to be almost of the size of $X(\mathbb{F}_q)$, so it is convenient to write $G = G^* - T$ where G^* consists of all points in $X(\mathbb{F}_q)$ (i.e. G^* is the maximal s.e.d. on X), and T is a "small" subset in G^*. This suggests the following constructive form of Conjecture B-2:

Construct an ∞-sequence of curves X_k over $\mathbb{F}_q$ and a sequence of integers t_k, with $\frac{t_k}{g(X_k)} \longrightarrow 0$, and T_k s.e.d. on X_k of degree $\deg T_k \leq t_k$, such that for each k one can find explicitly an element of the translate $s \cdot (J_k(\mathbb{F}_q)) - \kappa_{G^*_k} + \kappa_{T_k}$ that does not belong to W_{g-t_k}.

Or equivalently:

Construct an ∞-sequence of curves X_k over $\mathbb{F}_q$, and D_k divisor on X_k, and T_k s.e.d. on X_k, such that:

$$\frac{|X_k(\mathbb{F}_q)|}{g(X_k)} \longrightarrow A(q) \qquad \frac{\deg T_k}{g(X_k)} \longrightarrow 0$$

$$\frac{\deg(sD_k + T_k - G^*_k)}{g(X_k)} \longrightarrow 1 \qquad l(sD_k + T_k - G^*_k) = 0.$$

This would lead to a constructive proof of the lower bound $R^{\text{lin}}_q(s) \geq \frac{1}{s} - \frac{1}{A(q)} + \frac{1}{sA(q)}$ in Proposition 11, and hence also to a constructive proof of $R_2(2) \geq \frac{3}{50}\log_2(11)$.

Consider now Conjecture B-1. The key point in the proof that Conjecture B-1 implies Conjecture B-2 is Proposition 6. In this Proposition, one proves that $s \cdot (J(\mathbb{F}_q)) - \kappa_G$ is not included in $W_r(\mathbb{F}_q)$ by proving that the former has cardinality larger than the latter. This is a strong requirement, indicating that Conjecture B-2 should be quite easier to prove than Conjecture B-1, and at the same time it is a non-constructive argument. Thus even a constructive proof for Conjecture B-1 would not lead to an effective construction of good frameproof codes.

Anyway this should not prevent us to try to prove Conjecture B-1 constructively. We formulate it explicitely:

Construct an ∞-sequence of curves X_k over $\mathbb{F}_q$ such that $\alpha(X_k) \longrightarrow A(q)$ and $\delta_s(X_k) \longrightarrow 0$.

When $s = l$ is a prime number, $J_k(\mathbb{F}_q)[l]$ is a $\mathbb{Z}/l\mathbb{Z}$-vector space, and $\log_l |J_k(\mathbb{F}_q)[l]|$ is its dimension, so this can be rewritten as:

Construct an ∞-sequence of curves X_k over $\mathbb{F}_q$, of genus g_k, and with Jacobian J_k, such that:

$$\frac{|X_k(\mathbb{F}_q)|}{g_k} \longrightarrow A(q) \qquad \frac{\dim_{\mathbb{Z}/l\mathbb{Z}} J_k(\mathbb{F}_q)[l]}{g_k} \longrightarrow 0.$$

For $q = p^2$, several families of curves attaining $A(q) = p - 1$ have been constructed. For example in [8] such curves are given by an explicit tower of Artin-Schreier extensions. Then a way to prove Conjecture B-1 would be to control the l-primary part of the class number in such a tower.

B. Modular curves and Cohen-Lenstra heuristics

Another family of curves attaining $A(q)$, when $q = p^2$, is given by the modular curves $X_0(N)$. The class number of these curves is then given by the following, which is a consequence of the Eichler-Shimura relation:

Proposition 15: For N prime to p, and $q = p^2$, denote by $T_p(N)$ the Hecke operator T_p acting on the space of weight 2 cusp forms $S_2(\Gamma_0(N))$. Then

$$|J_0(N)(\mathbb{F}_q)| = \det((p+1)^2 - T_p(N)^2).$$

Corollary 16: With the same notations,

$$\dim_{\mathbb{Z}/l\mathbb{Z}} J_0(N)(\mathbb{F}_q)[l] \leq v_l(\det((p+1)^2 - T_p(N)^2)).$$

From this we see that in order to prove Conjecture B-1, it would suffice to prove:

$$\liminf_{\substack{N\longrightarrow\infty \\ (N,p)=1}} \frac{v_l(\det((p+1)^2 - T_p(N)^2))}{g_0(N)} = 0.$$

This could be put in a more general framework. The Cohen-Lenstra heuristics [4] predict the distribution of the l-primary part of class groups in families of number fields. They were later extended to algebraic curves in [7]. If one knew that these heuristics hold for the modular curves $X_0(N)$, one would know the precise distribution of the $J_0(N)(\mathbb{F}_q)[l]$, and all the previous conjectures would follow.

A possible approach for results of this type would follow the lines of [16]. There, Serre first estimates the trace of powers of $T_p(N)$, given explicitely by the Eichler-Selberg trace formula. From this he then derives an equidistribution result for the eigenvalues of $T_p(N)$, relative to the Archimedean topology. What one would like to do is the analogue for the l-adic topology.

As a conclusion, we consider the case $l = 2$. There it is easy to relate Conjectures A and B-1: if p is an odd prime, then $\det T_p(N)$ is odd if and only if $\det((p+1)^2 - T_p(N)^2)$ is, so Conjecture A would give an infinite number of primes N such that $v_2(\det((p+1)^2 - T_p(N)^2) = 0$. This is already more than what we need, but we could ask even more: do these primes form a set of positive Dirichlet density?

So let $p > 2$ be prime, and let $\Sigma(p)$ be the set of prime numbers N such that $T_p(N)$ has odd determinant.

Observe that a matrix M, with coefficients in $\mathbb{Z}$, has odd determinant, if and only if the matrix $\overline{M}$, obtained by reducing its coefficients modulo 2, is invertible. Now for any integer g define P_g as the probability that a uniformly distributed random matrix $\overline{M} \in \mathbf{M}_g(\mathbb{Z}/2\mathbb{Z})$ is invertible. It is easily shown that $P_g = (1-1/2)(1-1/4)\cdots(1-1/2^g)$, which tends to

$$P^* = \textstyle\prod_{k\geq 1}\left(1-\frac{1}{2^k}\right) \approx 0.288788...$$

as g tends to infinity. From this discussion we see that P^* would be a natural candidate for the density of $\Sigma(p)$. In fact, as noted in [7], this is precisely what the Cohen-Lenstra heuristics would predict. As a first step in this direction, one can prove the following:

Proposition 17: Let $p > 2$ be prime. Then, except perhaps for a finite number of values, all $N \in \Sigma(p)$ satisfy:

1) $-N$ is quadratic residue modulo p
2) $N \neq 1 \mod 8$.

Remark that condition 1) eliminates half of the primes, while condition 2) eliminates one fourth, independently. Hence, $\Sigma(p)$ is included in a subset of Dirichlet density

$$\left(1-\tfrac{1}{2}\right)\left(1-\tfrac{1}{4}\right) = 0.375$$

in the set of all prime numbers. It is very tempting to see this as the first two factors in the infinite product defining P^*.

ACKNOWLEDGMENT

The author got interested in frameproof codes through discussions with G. Cohen, J. Körner, and G. Zémor. Also, an apparently innocuous comment by F. Morain during a talk of the author had a profound impact on this work.

All the material here was known to the author around 2003, but never appeared in print (except for an announcement in [2]) because of its largely conjectural nature. However, considering only limited progress was made since then, time has probably come to share it with a broader audience.

REFERENCES

[1] D. Boneh and J. Shaw, *Collusion-secure fingerprinting for digital data*, IEEE Trans. Inform. Theory **44** (1998) 1897–1905.
[2] G. Cohen, "Separation and witnesses", in C. Xing et al. (Eds), *IWCC 2009*, LNCS **5557**, 12–21, Springer, 2009.
[3] G. Cohen and A. Lempel, *Linear intersecting codes*, Discr. Math. **56** (1985) 35–43.
[4] H. Cohen and H.W. Lenstra, Jr., "Heuristics on class groups of number fields", in: *Number theory, Noordwijkerhout 1983*, LNM **1068**, Springer.
[5] G.D. Forney, *Concatenated codes*, MIT Press, 1966.
[6] A. D. Friedman, R. L. Graham, and J. D. Ullman, *Universal single transition time asynchronous state assignments*, IEEE Trans. Comput. **18** (1969) 541–547.
[7] E. Friedman and L.C. Washington, "On the distribution of divisor class groups of curves over a finite field", in: *Théorie des nombres (Quebec, 1987)*, de Gruyter, 227–239.
[8] A. Garcia and H. Stichtenoth, *A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound*, Invent. Math. **121** (1995) 211–222.
[9] Y. Ihara, *Some remarks on the number of rational points of algebraic curves over finite fields*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **28** (1981) 721–724.
[10] J. Körner, *On the extremal combinatorics of the Hamming space*, J. Combin. Theory Ser. A **71** (1995) 112–126.
[11] A. Krasnopeev and Yu. L. Sagalovich, "The Kerdock codes and separating systems", in: *ACCT-8, Tsarskoe Selo (St Petersbourg) 2002*, 165–167.
[12] D. Miklós, *Linear binary codes with intersection properties*, Discr. Appl. Math. **9** (1984) 187–196.
[13] D. Mumford, *Abelian varieties*, Oxford Univ. Press, 1974.
[14] H. Niederreiter, H. Wang, and C. Xing, "Function fields over finite fields and their applications to cryptography", in: A. Garcia and H. Stichtenoth, *Topics in geometry, coding theory and cryptography*, Springer, 2007.
[15] Yu. L. Sagalovich and A. G. Chilingarjan, *Separating systems and new scopes of its application*, Information Processes **9** (2009) 225–248.
[16] J.-P. Serre, *Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p*, J. Amer. Math. Soc. **10** (1997) 75–102.
[17] D. R. Stinson and R. Wei, *Combinatorial properties and constructions of traceability schemes and frameproof codes*, SIAM J. Discr. Math. **11** (1998) 41–53.
[18] M.A. Tsfasman, S.G. Vladut, and T.Zink, *Modular curves, Shimura curves, and Goppa codes better than Varshamov-Gilbert bound*, Math. Nachr. **109** (1982) 21–28.
[19] C. Xing, *Asymptotic bounds on frameproof codes*, IEEE Trans. Inform. Theory **48** (2002) 2991–2995.