Curves over Finite Fields and Codes

Gerard van der Geer

Abstract. This paper gives a review of recent developments in this field and discusses some questions.

1. Introduction

The history of counting points on curves over finite fields goes back at least to C.F. Gauss who counted the number of points on several types of curves defined over the prime field $\mathbf{Z}/p\mathbf{Z}$. For example, in §358 of his Disquisitiones of 1801 he counts the number of points on the Fermat curve

$$x^3 + y^3 + z^3 \equiv 0 \pmod{p} \tag{1}$$

with p an odd prime greater than 3 and gives a beautiful answer: if $p \not\equiv 1 \,(\mathrm{mod} 3)$ then the number $\#C(\mathbf{F}_p)$ of $\mathbf{F}_p$-rational points on the projective curve C defined by (1) is $p+1$, while if $p \equiv 1 \,(\mathrm{mod} 3)$ there is a unique way of writing $4p = a^2 + 27b^2$ with a and b integers and $a \equiv 1 \,(\mathrm{mod} 3)$ and then $\#C(\mathbf{F}_p) = p + 1 + a$. Note that one has $|a| < 2\sqrt{p}$. Also Jacobi worked on the number of solutions of such congruences because he wanted to obtain estimates for Gauss sums. After them the problem sunk into oblivion for a long time.

In his 1924 thesis E. Artin introduced a zeta function $\zeta_F(s)$ for hyperelliptic function fields $F = \mathbf{F}_q(x, y)$ over a finite field $\mathbf{F}_q$, with q odd, where y satisfies $y^2 = f(x)$, in analogy with the Dedekind zeta function $\zeta_K(s) = \sum_{\mathfrak{a}} N(\mathfrak{a})^{-s}$ for a number field K. Artin noticed that by substituting $t = q^{-s}$ this zeta function became a rational function $Z_F(t)$ of t and that it satisfied a functional equation relating $\zeta_F(s)$ and $\zeta_F(1-s)$. He also advanced a conjecture as analogue of the Riemann hypothesis saying that the zeros of $Z_F(t)$ satisfy $|t| = q^{-1/2}$. Artin formulated everything in terms of ideals and ideal classes, but shortly afterwards F.K. Schmidt introduced a more geometric point of view and wrote the zeta function for a smooth (absolutely) irreducible projective curve X defined over $\mathbf{F}_q$ in the form

$$Z_X(t) = \exp\Big(\sum_{r=1}^{\infty} c(r)\frac{t^r}{r}\Big)$$

with $c(r) = \#X(\mathbf{F}_{q^r})$ and observed that the theorem of Riemann-Roch implied that for a curve of genus g the function $Z_X(t)$ is of the form

$$Z_X(t) = \frac{P_X(t)}{(1-t)(1-qt)}, \tag{2}$$

with $P(t)$ a polynomial of degree $2g$, and that it satisfies a functional equation

$$Z_X(1/qt) = q^{1-g}t^{2-2g}Z_X(t).$$

Around 1932 Hasse noticed that the conjecture by Artin implied

$$|\#X(\mathbf{F}_q) - (q+1)| \leq 2g\sqrt{q} \tag{3}$$

and proved it for $g = 1$ using correspondences. Deuring observed then that in order to extend this proof to higher genus one needed a theory of correspondences. This theory was developed by A. Weil and within 16 years after Artin put forward his conjecture it was proved by Weil. He showed that the polynomial $P_X(t)$ in (2) is a polynomial with integral coefficients of the form $P_X(t) = \prod_{i=1}^{2g}(1-\alpha_i t)$, where the α_i are algebraic integers with $|\alpha_i| = \sqrt{q}$ and this implies the famous Hasse-Weil bound (3). One also finds the formula

$$\#X(\mathbf{F}_{q^r}) = q^r + 1 - \sum_{i=1}^{2g} \alpha_i^r. \tag{4}$$

Note that there exist curves which attain this bound over $\mathbf{F}_{q^2}$; for example the so-called hermitian curves defined by

$$x^{q+1} + y^{q+1} + z^{q+1} = 0$$

have genus $g = q(q-1)/2$ and satisfy $\#X(\mathbf{F}_{q^2}) = q^3 + 1 = q^2 + 1 + 2gq$.

After Weil obtained his result interest waned again, but it was brought back by coding theory.

2. Error correcting codes

The theory of error-correcting codes was born out of frustration about the frequent stops made by early computers every time these observed an inaccuracy when a parity check failed. Early pioneers like Hamming, Golay and others invented schemes to add redundancy with which inaccuracies due to noise could be repaired and tried to do this efficiently. They found some beautiful mathematical structures while doing this. For a review of the history and the results we refer to [M-S].

Fix a finite field $\mathbf{F}_q$ of cardinality q and consider the $\mathbf{F}_q$-vector space $\mathbf{F}_q^n$. The set $\mathbf{F}_q$ is called the alphabet, the elements of $\mathbf{F}_q^n$ are called *words* and the integer n is called the word length. The Hamming distance on $\mathbf{F}_q^n$ is the distance function with $d(x,y) = \#\{i : 1 \leq i \leq n, x_i \neq y_i\}$. The weight $w(x)$ of a word x is its distance to the origin.

By a linear code we shall mean a linear subspace C of $\mathbf{F}_q^n$. We say that $C \subseteq \mathbf{F}_q^n$ is a (n,k,d)-code if its dimension is k and the minimum weight of a non-zero word in C is d. A rough gauge of the quality of a code is provided by two invariants: the transmission rate $R = k/n$ and the relative distance $\delta = d/n$. In essence coding theory is a game where one tries to find codes that optimize these invariants. We shall restrict ourselves to linear codes.

If $C \subseteq \mathbf{F}_q^n$ is a linear code, then

$$C^{\perp} = \{x \in \mathbf{F}_q^n : \langle x, y \rangle = 0, \quad \text{for all} \quad y \in C\}$$

with $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ is called the dual code.

An important problem of coding theory is the determination of the weight distribution of a code, that is, of the polynomial

$$\sum_{c \in C} X^{w(c)} \quad \in \mathbf{Z}[X].$$

The weight distribution and that of the dual code can be expressed in each other by the MacWilliams identities.

The notion of weight of a word admits an extension to the weight of a subspace of C. If $D \subseteq C$ is a linear subspace of dimension r we define the weight of D as the number of coordinate places for which C contains a codeword with a non-vanishing coordinate at that place, or equivalently as

$$w(D) = \frac{1}{q^r - q^{r-1}} \sum_{c \in D} w(c).$$

The weight hierarchy of C is the set $\{d_r(C) : 1 \leq r \leq n\}$ of generalized Hamming weights $d_r(C)$ defined by

$$d_r(C) = \min\{w(D) : D \subseteq C, \dim(D) = r\}.$$

The numbers $d_r(C)$ are a measure for the reliability of a transmission channel where part of the data fall prey to hostile eavesdroppers, cf. [Wei].

An important class of codes is given by the generalized Reed-Muller codes. Consider the set of polynomials

$$P_s = \{f \in \mathbf{F}_q[X_1, \ldots, X_r] : \deg(f) \leq s\}.$$

Here the degree is the total degree. We can evaluate elements of P_s at the points of affine r-space over $\mathbf{F}_q$:

$$\beta : P_s \to \mathbf{F}_q^n, \qquad f \mapsto (f(P)_{P \in \mathbf{F}_q^r}).$$

where $n = q^r$. The image $\beta(P_s)$ of this map is the generalized q-ary Reed-Muller code $R_q(s,r)$. These codes were studied intensively from the middle of the 20th century. The weight of a code word $\beta(f)$ is $n - \#H_f(\mathbf{F}_q)$, with $H_f(\mathbf{F}_q)$ the set of $\mathbf{F}_q$-rational points on the hypersurface H_f given by $f = 0$. Determining the weight distribution of this code is equivalent to determining the distribution of the number of points in the family of all hypersurfaces of degree $\leq s$ in $\mathbf{F}_q^r$. This makes the difficulty of the problem apparent and it will be no surprise that the weight distribution of $R_q(s,m)$ in general is not known for $s \geq 3$.

3. Asymptotically good sequences of codes

There are classical bounds on the parameters of codes, like the simple Singleton bound $k+d \leq n+1$ for codes with dimension k and minimum distance d, which show that the parameters of codes underly certain restrictions.

A central problem of coding theory from its early beginnings has been to find long codes which can correct a fixed percentage of errors per codeword and which have a strictly positive transmission rate. To a linear code one can associate the pair $(\delta = d/n, R = k/n) \in [0,1]^2$ of relative parameters. Let $U_q^{\rm lin}$ be the set of limit points of the set of all such pairs coming from linear codes. The region $U_q^{\rm lin}$ is called the domain of codes. It is bounded in the unit square by the sides of the unit squares on the δ- and R-axis and by the graph of a function $\alpha_q^{\rm lin} : [0,1] \to [0,1]$ defined by

$$\alpha_q^{\rm lin}(\delta) = \sup\{R : (\delta, R) \in U_q^{\rm lin}\}.$$

A sequence of codes (C_i) with parameters (n_i, k_i, d_i) such that the ratios d_i/n_i and k_i/n_i converge to δ and R with $\delta R > 0$ is called an *asymptotically good sequence*.

Define the entropy function $H_q(\delta)$ by

$$H_q(\delta) = \begin{cases} 0 & \delta = 0, \\ \delta \log_q(q-1) - \delta \log_q(\delta) - (1-\delta)\log_q(1-\delta) & 0 < \delta \leq (q-1)/q. \end{cases}$$

We denote by $b_q(n,d)$ the number of points in a ball of radius d in $\mathbf{F}_q^n$. Gilbert made the simple observation that, if $q^n > q^{k-1} b_q(n, d-1)$, there exists a (n,k,d)-code over $\mathbf{F}_q$. This together with the fact that $\lim_{n\to\infty}(\log_q b_q(n, [\delta n]))/n$ equals the entropy function for $0 \leq \delta \leq (q-1)/q$ implies the following bound.

Proposition 3.1. (Gilbert-Varshamov bound) *For $0 \leq \delta \leq (q-1)/q$ the function $\alpha_q^{\rm lin}$ satisfies $\alpha_q^{\rm lin}(\delta) \geq 1 - H_q(\delta)$.*

Another elementary bound from coding theory, the Plotkin bound, implies

$$\alpha_q^{\rm lin}(\delta) \leq 1 - (\frac{q}{q-1})\delta \qquad \text{for} \quad 0 \leq \delta \leq (q-1)/q,$$

and

$$\alpha_q^{\rm lin}(\delta) = 0 \qquad \text{for} \quad (q-1)/q \leq \delta \leq 1.$$

Thus on the interval $[0, (q-1)/q]$ the function $\alpha_q^{\rm lin}$ is bounded from below and above by the graphs of $1 - H_q(\delta)$ and $1 - (q/(q-1))\delta$. Manin showed moreover that $\alpha_q^{\rm lin}$ is a continuous function function of δ which is decreasing on $[0, (q-1)/q]$.

For a long time coding theorists were unable to construct explicit sequences of codes with limit points on or above the Gilbert-Varshamov bound and they were thus led to suspect that $\alpha_q^{\rm lin}(\delta) = 1 - H_q(\delta)$ for $0 \leq \delta \leq (q-1)/q$.

4. Goppa codes

In 1973 V.D. Goppa succeeded in constructing sequences of codes attaining the Gilbert-Varshamov bound. The codes were obtained by taking as entries of their parity check matrix the values of rational functions. When he tried to improve upon this he had the idea of taking values of rational functions on algebraic curves and thus discovered around 1980 an unexpected relationship between algebraic curves and codes, see [Go1], [Go2].

If X a a smooth irreducible projective curve of genus g defined over a field $\mathbf{F}_q$ with a given set $P = \{P_1, \ldots, P_n\}$ of n distinct $\mathbf{F}_q$-rational points and if $L \subset \mathbf{F}_q(X)$ is a $\mathbf{F}_q$-linear subspace of the function field $\mathbf{F}_q(X)$ such that no $f \in L$ has a pole in the points P_i we can define an evaluation map

$$\alpha : L \to \mathbf{F}_q^n, \quad f \mapsto (f(P_i)_{i=1}^n).$$

The image of L under α is then a code. In case

$$L = L(D) = \{f \in k(X)^* : (f) + D \geq 0\} \cup \{0\}$$

with D a $\mathbf{F}_q$-divisor we find the code $C(D, P)$. Invariants of these codes, like the dimension k and minimum distance d, can be expressed in terms of properties of the curve; for example, one has the elementary result:

Theorem 4.1. *If D is a divisor defined over $\mathbf{F}_q$ of degree $g \leq \deg D \leq n$ and with $\mathrm{supp}(D) \cap P = \emptyset$, then we have $k \geq \deg(D) + 1 - g$ with equality if $\deg(D) \geq 2g - 1$. Moreover, $d \geq n - \deg(D)$.*

If we fix the ratio $\deg(D)/n$ then the transmission rate $R = k/n$ increases with the ratio n/g. Therefore to obtain good codes one has to construct curves with as many points as possible.

If X_ℓ is a sequence of curves defined over $\mathbf{F}_q$ such that their genera g_ℓ tend to ∞ and such that $\lim_{\ell \to \infty} \#X_\ell(\mathbf{F}_q)/g_\ell = \gamma > 0$ then the part of the line $\delta + R = 1 - 1/\gamma$ in the positive quadrant is contained in the domain U_q^{lin}. This follows by taking divisors D_ℓ of degree $[\#X_\ell(\mathbf{F}_q)(1-\delta)]$ and evaluating the functions in $L(D_\ell)$ in the rational points $X_\ell(\mathbf{F}_q)$. Then Theorem 4.1 tells us that for the geometric Goppa codes $C_\ell = C((D_\ell, X_\ell(\mathbf{F}_q))$ we have

$$R_\ell + \delta_\ell \geq 1 + (1 - g_\ell)/\#X_\ell(\mathbf{F}_q)$$

which tends to $1 - 1/\gamma$. Hence this sequence of codes has a limit point on or above the line $R + \delta = 1 - 1/\gamma$.

The fact that for q a square there exists a sequence of curves X_ℓ defined over $\mathbf{F}_q$ of genus g_ℓ with the ratio $\#X_\ell(\mathbf{F}_q)/g_\ell$ tending to $\sqrt{q} - 1$ was observed by Ihara (see [I1]) and independently by Tsfasman, Vladuts and Zink who applied it to coding theory, see [T-V-Z]. They used modular curves $X_0(N)$ and the rational points on them provided by the supersingular points.

Theorem 4.2. (Tsfasman, Vladuts, Zink) *Let q be a square. There exists a sequence of geometric Goppa codes over $\mathbf{F}_q$ with limit point on or above the line $R+\delta = 1-1/(\sqrt{q}-1)$.*

For $q \geq 49$ this line comes above the Gilbert-Varshamov bound and one thus gets asymptotically good sequences of codes above the Gilbert-Varshamov bound, and this came at that time as quite a surprise for coding theorists.

Goppa's discovery led to renewed interest in curves over finite fields and drew attention to new aspects of these curves.

5. Upper bounds for the number of points on a curve

The new interest for curves over finite fields soon paid its dividends. It was noted by Ihara (see [I2]) that the Hasse-Weil bound (3) could be improved. He did this by comparing $\#X(\mathbf{F}_q)$ and $\#X(\mathbf{F}_{q^2})$ using (4) and the Cauchy-Schwartz inequality. His bound

$$\#X(\mathbf{F}_{q^2}) \leq q+1+[(\sqrt{(8q+1)g^2+4(q^2-q))g}-g)/2]$$

is better than the Hasse-Weil bound for $g > (q-\sqrt{q})/2$. Drinfeld and Vladuts generalized this idea by using all extensions $\mathbf{F}_{q^r}/\mathbf{F}_q$ instead of just the quadratic one (see [V-D]). Let us define

$$N_q(g) := \text{maximum value of} \, \#X(\mathbf{F}_q), \tag{5}$$

where X runs through all curves of genus g defined over $\mathbf{F}_q$. Moreover, we define

$$A(q) := \limsup_{g\to\infty} N_q(g)/g.$$

The result of Drinfeld and Vladuts says that $A(q) \leq \sqrt{q}-1$.

Serre started the study of the actual value of $N_q(g)$, see [Se2]. He improved the Hasse-Weil bound slightly by applying some arithmetic to the α_i (cf. (4)):

$$N_q(g) \leq q+1+g[2\sqrt{q}]. \tag{6}$$

Serre also transplanted the 'formules explicites' from number theory. He takes a trigoniometric polynomial

$$f = 1+2\sum_{n\geq 1} u_n \cos n\theta$$

with real coefficients $u_n \geq 0$ such that $f(\theta) \geq 0$ for all $\theta \in \mathbb{R}$ and sets $\psi = \sum_{n\geq 1} u_n t^n$. Then he gets the estimate

$$N_q(g) \leq a_f g + b_f$$

with

$$a_f = \frac{1}{\psi(1/\sqrt{q})} \quad \text{and} \quad b_f = 1+\frac{\psi(\sqrt{q})}{\psi(1/\sqrt{q})}.$$

Finding the optimal choices for such functions f is a linear programming problem which was solved by Oesterlé, see [Se4], or [Sch] for an exposition. The resulting bounds are called Oesterlé bounds.

6. Asymptotics

Combining the result of Drinfeld and Vladuts $A(q) \leq \sqrt{q} - 1$ with the result of Ihara shows that

$$A(q) = \sqrt{q} - 1 \quad \text{for } q \text{ a square.}$$

In 1996 Garcia and Stichtenoth constructed an explicit tower of curves X_ℓ defined over $\mathbf{F}_{q^2}$. This tower starts with the rational line X_1 with coordinate x_1 and consists of successive Artin-Schreier extensions defined by

$$y_{\ell+1}^q + y_\ell = x_\ell^{q+1},$$

where x_ℓ is defined recursively by

$$x_\ell = y_\ell / x_{\ell-1}$$

and for which the ratio $\#X_\ell(\mathbf{F}_{q^2})/g(X_\ell)$ tends to $q-1$. Later Elkies showed that these towers are modular towers, cf. [E1].

For q not a square the situation is yet unclear. Zink used in [Z] degenerations of Shimura surfaces to show that if $q = p^3$ then there exists a sequence of curves over $\mathbf{F}_q$ with $\gamma = \lim X_\ell/g_\ell \geq 2(p^2-1)/(p+2)$. In certain cases one can construct explicit towers of curves reaching this bound γ, cf. [G-V5].

In any case we know since the 1980's by work of Serre who used class field towers that $A(q) > c \log q$ for an absolute constant $c > 0$. We refer to the paper by Hajir and Maire in this volume for a survey of asymptotically good towers.

These results make use of towers where the genera that occur may be rather sparse. If one insists on using all sufficiently large genera the problem becomes more difficult.

In [K-W-Z] it is shown that $N_q(g)$ goes to ∞ with g:

Theorem 6.1. ([K-W-Z]) *For fixed q we have* $\lim_{g\to\infty} N_q(g) = \infty$.

The following question suggests itself:

Question 6.2. *What is* $\liminf_{g\to\infty} N_q(g)/g$*?*

It follows from the work [E-Z] that $\liminf_{g\to\infty} N_q(g)/g \geq (\sqrt{q}-1)/3$ if q is a square.

The following results gives restrictions for the number $N_q(g)$ for all g:

Theorem 6.3. ([K-W-Z]) *For fixed q there are constants e_q and f_q depending on q with $0 < e_q < f_q$ such that for every $g > 0$ one has $e_q g < N_q(g) < f_q g$.*

7. Maximal Curves

We shall call a curve X defined over $\mathbf{F}_q$ *maximal* if $\#X(\mathbf{F}_q) = q + 1 + 2g\sqrt{q}$, i.e. if the curve attains the Hasse-Weil bound. Then q is a square and since the Ihara bound improves the Hasse-Weil bound if $g > (q - \sqrt{q})/2$, this can only happen if $g \leq (q - \sqrt{q})/2$. If X is a maximal curve and Y a curve dominated by X then Y is also a maximal curve, since the Jacobian of Y is an isogeny factor of the Jacobian of X.

For a maximal curve X the action of Frobenius F on the Jacobian is by $-\sqrt{q}$. Choose a $\mathbf{F}_{q^2}$-rational point P_0 on X and map X to the Jacobian $\mathrm{Jac}(X)$ by $P \mapsto [P - P_0]$. If $F(P)$ is the Frobenius image of P on X then one has

$$-\sqrt{q}[P - P_0] = [F(P) - P_0]$$

in the Jacobian, so for any point P the divisor $\sqrt{q}P + F(P)$ is linearly equivalent to $(\sqrt{q} + 1)P_0$ and this gives a canonically defined linear system on such curves. Fuhrmann and Torres and Garcia applied ideas of Stöhr and Voloch ([St-V]; see also Sect. 8 below) to this linear system to prove the following restrictions on the genera of maximal curves. Let $g_0 = g_0(q) = (q - \sqrt{q})/2$ and $g_1 = g_1(q) = (\sqrt{q} - 1)^2/4$.

Theorem 7.1. ([F-T], [F-G-T]) *If X is a maximal curve of genus g over $\mathbf{F}_q$ then either $g = g_0$ or $g \leq g_1$. Moreover, if q is odd and $(\sqrt{q} - 1)(\sqrt{q} - 2)/4 < g \leq g_1$ then $g = g_1$.*

A natural question arises:

Question 7.2. *Determine the genera for which there are maximal curves. Characterize these curves.*

Maximal curves with $g = g_0$ or $g = g_1$ have been characterized: for genus $g = g_0$ the curve is the hermitian curve, for $g = g_1$ it is given by $y^q + y = x^{\sqrt{q}+1)/2}$, which is dominated by a hermitian curve. For curves of genus $g_2 = (\sqrt{q} - 1)(\sqrt{q} - 3)/8$ there are two non-isomorphic types of maximal curves for $\sqrt{q} \equiv 3 (\mathrm{mod} 4)$, the Fermat curve of degree $(\sqrt{q} + 1)/2$ and the Artin-Schreier curve $y^{\sqrt{q}} + y = x^{(\sqrt{q}+1)/4}$. It is not known whether there are other types. We refer to the paper by Garcia in this volume for more details.

The maximal curves that we know all seem to come from the hermitian curve, although this has not been verified for all of them, which could motivate the following question, but the evidence is limited.

Question 7.3. (Stichtenoth) *Is every maximal curve the image under a dominant map of a hermitian curve?*

8. The function $N_q(g)$

For each pair (q,g) there is the constant of nature $N_q(g)$ defined by (5). Besides the asymptotic behavior the actual value of $N_q(g)$ for relatively small q and g is interesting, just as a test for our knowledge on curves over finite fields, but also with an eye to practical applications in coding theory or cryptography.

The Hasse-Weil bound and its improvements (Ihara, Serre, Oesterlé) give us an upper bound for $N_q(g)$. To see how good this upper bound is one has to construct curves with as many points as possible. In practice this results in an interval $[a,b]$ to which $N_q(g)$ is constrained. Here b the best upper bound we know and a is the largest number of points we know to occur for a curve of genus g over $\mathbf{F}_q$.

Serre determined the actual value of $N_q(g)$ for a number of small values of g and q. For example he determined $N_q(2)$ for all q and $N_q(g)$ for certain pairs (q,g) with $q=2$.

In some cases for small genera certain values of $N_q(g)$ can be eliminated by listing all the possibilities of the zeta function and showing that zeta functions in this list imply a decomposition of the Jacobian as a product of principally polarized abelian varieties, which contradicts the irreducibility of the theta divisor of the curve, cf. [Se4], [L2], [L3].

Furthermore, a result of Serre says that for a curve of genus ≥ 3 with $\#X(\mathbf{F}_q) < q+1+g[2\sqrt{q}]$ one has $\#X(\mathbf{F}_q) \leq q-1+g[2\sqrt{q}]$.

Sometimes one can rule out that $N_q(g)$ equals the Serre bound (6) by a specific argument, like Galois descent. This works e.g. for $(q=27, g=3)$ and $(q=8, g=4)$, cf. [Se4], [L3].

We give a sample of results thus obtained.

Proposition 8.1. *One has the following explicit results:*
1) $N_2(7)=10$*;*
2) $N_3(5) \leq 13$ *and* $N_3(7)=16$*;*
3) $N_4(4)=15$ *and* $N_4(9)=30$.
4) $N_{27}(3)=56$.

It seems not unlikely that with such methods many more results can be obtained, maybe even of a general nature.

A look at tables for $N_q(g)$ (see [G-V3]) suggests the following question.

Question 8.2. *Is the function* $N_q(g)$ *a non decreasing function of* g *for fixed* q*?*

Instead of studying the maximum value of $\#X(\mathbf{F}_q)$ for all curves of genus g one could restrict to curves of a specific type. For example, for hyperelliptic curves one has the obvious bound $\#X(\mathbf{F}_q) \leq 2(q+1)$. One could try to generalize this to bounds for the maximum number of points on a curve of genus g with given gonality. The gonality vector $\gamma(X)=(\gamma_1,\gamma_2,\dots)$ of a curve X over an algebraically closed field k is given by

$$\gamma_r(X)=\min\{d : 1 \leq d \leq g-1, \text{ there exists a } \mathfrak{g}_d^r \text{ on } X\},$$

where $\mathfrak{g}_d^r$ stands for a linear system of degree d and dimension r. A curve of genus g admits a map of degree $\leq [(g+3)/2]$ to the projective line, so the geometric gonality $\gamma_1(X)$ is bounded by $\leq [(g+3)/2]$. But this map need not be defined over our ground field. For example, a non-hyperelliptic curve of genus 4 over $\mathbf{F}_q$ has a map of degree 3 to $\mathbb{P}^1$ over $\mathbf{F}_{q^2}$, but not necessarily over $\mathbf{F}_q$, depending on whether the two rulings of the quadric containing the canonical curve are defined over $\mathbf{F}_q$ or not.

Question 8.3. *What is the maximum number of rational points on a curve of genus g and gonality γ defined over $\mathbf{F}_q$?*

In [St-V] Stöhr and Voloch present a sort of answer to this question, namely an upper bound for the maximum number of points on a curve over $\mathbf{F}_q$ which does not only depend on the genus g, but also on a given linear system defined over $\mathbf{F}_q$. It uses an infinitesimal approach which counts points on a curve, embedded in projective space with such a linear system, such that the Frobenius image of a point lies in the osculating hyperplane of the curve at that point. As a special case this provides a new proof of the Hasse-Weil bound.

Quite a lot of people have tried to construct curves with many points in order to test how good the upper bounds on $N_q(g)$ are. A variety of methods have been used for this, like methods from class field theory (Serre, Schoof, Lauter, Niederreiter, Xing and Auer), methods from Drinfeld modules (Niederreiter and Xing), fibre products of Artin-Schreier curves (Van der Geer and Van der Vlugt, Shabat), Kummer curves (Van der Geer and Van der Vlugt) and various other methods. We refer to [G-V1] and [G-V3] for a summary of the methods and results.

9. Another relationship between curves and codes

A relationship between curves and codes quite different from the one discovered by Goppa, but not less important arises from the observation that many classical codes are trace codes of the form $\mathrm{Tr}(C')$, where C' is a code over a field $\mathbf{F}_{q^s}$ obtained from evaluating functions f at points P_i of the affine line and $\mathrm{Tr} = \mathrm{Tr}_{q^s/q}$ is the trace from $\mathbf{F}_{q^s}$ to $\mathbf{F}_q$. To give an example, the classical dual Melas codes $M(q)^{\perp}$ over $\mathbf{F}_p$ are codes of length $q-1$ with words of the form

$$c_{a,b} = (\mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}(ax + b/x))_{x \in \mathbf{F}_q^*} \quad \text{with} \quad a, b \in \mathbf{F}_q$$

Since $\mathrm{Tr}_{\mathbf{F}_q/\mathbf{F}_p}(u) = 0$ if and only if there exists an element $v \in \mathbf{F}_q$ with $v^p - v = u$ we see that the weight of a word $c_{a,b}$ equals

$$q - 1 - \frac{1}{p}(\#X_{a,b}(\mathbf{F}_q) - 2),$$

where $X_{a,b}$ is the (smooth projective) curve given by $y^p - y = ax + b/x$.

More generally, we consider a Goppa code $C' = C(D, P)$ over $\mathbf{F}_{q^s}$ associated to some triple (X, D, P) as above. Since the words of $\mathrm{Tr}_{q^s/q}(C')$ are of the form

$$\mathrm{Tr}(f(P_i))_{i=1}^n \quad \text{for} \quad f \in L(D)$$

we see that there is a simple relation between the weight of a word $\mathrm{Tr}(\alpha(f)) = \mathrm{Tr}(f(P_i))_{i=1}^n$ and the number of rational points on the curve defined by $y^q - y = f$. In this way we obtain a one-to-one correspondence between words in a trace code of a Goppa code and curves in a k-dimensional family of curves. Linear subspaces correspond then to fibre products of curves. The weight distribution in the code is directly related to the distribution of the number of points in this family. For example, the weight distribution of the quadratic Reed-Muller code $R_q(2, r)$ is related to a family of supersingular curves given by equations of the form $y^p - y = xR(x)$ with R running through a family of linearized polynomials of the form $R = \sum_{i=0}^h a_i x^{p^i}$. These curves are thus related to quadric hypersurfaces and the fact that we are dealing with quadrics enables one to determine the weight distribution.

This relationship between codes and families of curves leads to difficult questions on the behavior of the zeta function in families of curves.

10. Distribution of traces of Frobenius and weights

Although most of the attention so far focused on determining or bounding the function $N_q(g)$ one may ask more generally:

Question 10.1. *For given pair (q, g) which values can the number of points on a smooth projective irreducible curve of genus g over $\mathbf{F}_q$ assume?*

In a given family of curves over $\mathbf{F}_q$ we can ask for the frequencies with which a given number of points is assumed. Here we count the contribution of each curve X defined over $\mathbf{F}_q$ with multiplicity $1/\#\mathrm{Aut}_{\mathbf{F}_q}(X)$. The most basic families are the universal families over a cover of the moduli space M_g of smooth irreducible complete curves of genus g.

Question 10.2. *Determine the frequencies of the number of points in such (universal) families.*

Note that this is a difficult question in general since the information given by these frequencies suffices to determine the number of points on $M_{g,n}(\mathbf{F}_q)$, where $M_{g,n}$ is the universal n-pointed curve for all n. In joint work with S. del Baño and C. Faber we have determined the frequencies for the genus 2 moduli spaces $M_2 \otimes \mathbf{F}_p$ for all primes $p \leq 181$.

11. Strata on the moduli spaces

To an abelian variety in characteristic $p > 0$ we can associate the characteristic polynomial of Frobenius (acting on $H^1_{et}(X, \mathbf{Q}_\ell)$ with $p \neq \ell$) and the Newton polygon of this characteristic polynomial. As the Newton polygon goes up under specialization this leads to a stratification on moduli spaces of (polarized) abelian varieties. For an abelian variety over a finite field the isogeny class of X is determined by the characteristic polynomial as Tate showed and by work of Tate and Honda it is known which characteristic polynomials can occur. We restrict ourselves to principally polarized abelian varieties. The two extreme cases in this stratification on the moduli space $A_g \otimes \mathbf{F}_p$ of principally polarized abelian varieties of dimension g are the ordinary case and the supersingular case; the corresponding strata have dimension $g(g+1)/2$ and $[g^2/4]$, respectively, see [L-O]. Intermediate cases are the strata given by the condition that the p-rank equals r. These strata have codimension $g - r$ in A_g.

By associating to a curve its Jacobian variety this induces also a stratification on the moduli space $M_g \otimes \mathbf{F}_p$ of curves in characteristic $p > 0$. But unlike the case of abelian varieties it is not known which Newton polygons can occur. We know that the generic curve is ordinary, but for example for general p it is not known whether for every genus g there exists a curve defined over a finite field of characteristic p which is supersingular. For $p = 2$ it was proven in [G-V4] that for every g there exists a supersingular curve of genus g defined over $\mathbf{F}_2$. A similar construction in characteristic $p > 2$ gives supersingular curves for genera whose p-adic expansion uses 0 and $(p-1)/2$ only. This motivates the following questions, which can also be specialized to finite fields, but one may ask them as well in a more general setting.

Question 11.1. *For which genera does there exists a supersingular curve in characteristic p?*

Question 11.2. *Which Newton polygons occur for Jacobians?*

Since the characteristic polynomial of Frobenius is an isogeny invariant we can ask more generally the following important question that belongs to the folklore of the field:

Question 11.3. *Which isogeny classes of abelian varieties contain a jacobian variety?*

It is proved in [F-G] that the stratum of curves of p-rank r has codimension $g - r$ in the moduli space of curves $M_g \otimes \mathbf{F}_p$.

It might well be that the following final question admits unexpected answers over finite fields.

Question 11.4. *Can one give an effective procedure for deciding whether or not a polarized abelian variety is a jacobian?*

References

[Au] R. Auer: Ray class fields of global function fields with many rational places. Report University of Oldenburg, 1998.

[E1] N. Elkies: Explicit Modular Towers. *Proceedings of the Thirty -Fifth Annual Allerton Conference on Communication, Control and Computing* (1997, T. Basar and A. Vardy, eds), Univ. of Illinois at Urbana-Champaign 1998, p. 23–32.

[E-Z] N. Elkies, A. Kresch, B. Poonen, J. Wetherell, M. Zieve: Curves of every genus with many points II: asymptotically good families, in preparation.

[F-T] R. Fuhrmann, F. Torres: The genus of curves over finite fields with many rational points. *Manuscripta Math.* **89** (1996), p. 103–106.

[F-G-T] R. Fuhrmann, A. Garcia, F. Torres: On maximal curves. *J. Number Theory* **67** (1997), p. 29–51.

[F-G] C. Faber, G. van der Geer: Complete subvarieties of the moduli space of curves. Manuscript in preparation.

[G-S] F. Garcia, H. Stichtenoth: A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound. *Invent. Math.* **121** (1995), pp. 211-22.

[G] C.F. Gauss: Disquisitiones Arithmeticae 1801.

[G-V1] G. van der Geer, M. van der Vlugt: How to construct curves over finite fields with many points. In: *Arithmetic Geometry*, (Cortona 1994), F. Catanese Ed., Cambridge Univ. Press, Cambridge, 1997, p. 169-189.

[G-V2] G. van der Geer, M. van der Vlugt: On generalized Reed-Muller codes and curves with many points. *J. of Number Theory* **72** (1998), p. 257–268.

[G-V3] G. van der Geer, M. van der Vlugt: Tables for the function $N_q(g)$. Math. of Computation. Regularly updated tables at: http://www.science.uva.nl ~/geer

[G-V4] G. van der Geer, M. van der Vlugt: On the existence of supersingular curves of given genus. *Journal für die Reine und angewandte Math.*, **458** (1998), p. 53–61.

[G-V5] G. van der Geer, M. van der Vlugt: manuscript in preparation.

[Go1] V.D. Goppa: Codes associated with divisors. (Russian) *Problemy Peredavci Informacii* **13** (1977), p. 33–39.

[Go2] V.D. Goppa: Codes on algebraic curves. (Russian) *Dokl. Akad. Nauk SSSR* **259** (1981), p. 1289–1290.

[I1] Y. Ihara: Congruence relations and Shimura curves. II. *J. Fac. Sci. Univ. Tokyo* **25** (1979), p. 301–361.

[I2] Y. Ihara: Some remarks on the number of points of algebraic curves over finite fields. *J. Fac. Sci. Tokyo* **28** (1982), p. 721–724.

[K-W-Z] A. Kresch, J. Wetherell, M. Zieve: Curves of every genus with many points, I: Abelian and toric families. Preprint 2000.

[L1] K. Lauter: Ray class field constructions of curves over finite fields with many rational points. In: *Algorithmic Number Theory* (Talence 1996), H. Cohen Ed., Lecture Notes in Computer Science 1122, Springer, Berlin, 1996, p. 187-195.

[L2] K. Lauter: Non-existence of a curve over $\mathbf{F}_3$ of genus 5 with 14 rational points. *Proc. Amer. Math. Soc.* **128** (2000), p. 369–374.

[L3] K. Lauter: Improved upper bounds for the number of rational points on algebraic curves over finite fields. *Comptes Rend. Acad. Sci. Paris Sér. I Math.* **328** (1999), p. 1181–1185.

[L-O] K.-Z. Li, F. Oort: Moduli of supersingular abelian varieties. Springer Lecture Notes in Mathematics, 1680. Springer-Verlag, Berlin, 1998.

[vL] J. van Lint: Introduction to Coding Theory. Graduate Texts in Mathematics. Springer Verlag, 1998.

[M-S] F. MacWilliams, N. Sloane: The theory of error-correcting codes. North-Holland Publishing Company 1977.

[Ma] Yu. I. Manin: What is the maximum number of points on a curve over $\mathbf{F}_2$? *J. Fac. Sci. Tokyo* **28** (1981), pp. 715–720.

[N-X1] H. Niederreiter, C. P. Xing: Drinfeld modules of rank 1 and algebraic curves with many rational points II. *Acta Arithm.* **81** (1997), p. 81-100.

[N-X2] H. Niederreiter, C. P. Xing: Global function fields fields with many rational points over the ternary field. *Acta Arithm.* **83** (1998), p. 65–86.

[N-X3] H. Niederreiter, C. P. Xing: Algebraic curves with many rational points over finite fields of characteristic 2. To appear in: *Proc. Number Theory Conference* (Zakopane 1997), de Gruyter, Berlin.

[N-X4] H. Niederreiter, C. P. Xing: A general method of constructing global function fields with many rational places. To appear in: *Algorithmic Number Theory* (Portland 1998), Lecture Notes in Comp. Science, Springer, Berlin.

[Sch] R. Schoof: Algebraic curves and coding theory. UTM 336, University of Trento, 1990.

[Se1] J.P. Serre: Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini. Comptes Rendus Acad. Sci. Paris, **296**, (1983), pp. 397–402.

[Se2] J-P. Serre : Nombre de points des courbes algébriques sur $\mathbf{F}_q$. Sém. de Théorie des Nombres de Bordeaux, 1982/83, exp. no. 22. (= Oeuvres III, No. 129, p. 664-668).

[Se3] J-P. Serre : Quel est le nombre maximum de points rationnels que peut avoir une courbe algébrique de genre g sur un corps fini $\mathbf{F}_q$? Résumé des Cours de 1983-1984. (=Oeuvres III, No. 132, p. 701-705).

[Se4] J-P. Serre: Rational points on curves over finite fields. Notes of lectures at Harvard University 1985.

[Sh] V. Shabat: Unpublished manuscript, University of Amsterdam, 1997/2000.

[St-V] K.O. Stöhr, J.F. Voloch: Weierstrass points and curves over finite fields. *Proc. London Math. Soc.* **52** (1986), p. 1–19.

[T-V-Z] M.A. Tsfasman, S.G. Vladuts, Th. Zink: On Goppa codes which are better than the Varshamov-Gilbert bound. Math. Nachrichten **109**, (1982), pp. 21–28.

[V-D] S.G. Vladuts, V.G. Drinfeld: Number of points of an algebraic curve. Funct. Anal. **17**, (1983), pp. 68–69.

[Wei] V.K. Wei: Generalized Hamming weights for linear codes. *IEEE Trans. Inform. Theory* **37** (1991), p. 1412–1418.

[Z] Th. Zink: Degeneration of Shimura surfaces and a problem in coding theory. Fundamentals of computation theory (Cottbus, 1985), 503–511, Lecture Notes in Comput. Sci., 199, Springer, Berlin-New York, 1985.

Korteweg-de Vries Instituut,
Universiteit van Amsterdam,
Plantage Muidergracht 24,
1018 TV Amsterdam, The Netherlands.
E-mail address: `geer@science.uva.nl`