Asymptotically good sequences of curves and codes

Ruud Pellikaan *

Appeared in in *Proc. 34th Allerton Conf. on Communication, Control, and Computing,* Urbana-Champaign, October 2-4, 1996, 276-285.

1 Introduction

The parameters of a linear block code over the finite field F_q of *length* n, *dimension* k and *minimum distance* d will be denoted by $[n, k, d]_q$ or $[n, k, d]$. The quotient k/n is called the *information rate* and denoted by $R = k/n$ and the *relative minimum distance* d/n is denoted by δ.

A sequence of codes $(C_m | m \in N)$ with parameters $[n_m, k_m, d_m]$ over a fixed finite field F_q is called *asymptotically good* if n_m tends to infinity, and d_m/n_m tends to a non-zero constant δ, and k_m/n_m tends to a non-zero constant R for $m \to \infty$. Let $H_q(0) = 0$ and

$$H_q(x) = x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x)$$

for $0 < x \leq (q-1)/q$ be the entropy function. Then there exist asymptotically good sequences of codes attaining the the Gilbert-Varshamov (GV) bound

$$R \geq 1 - H_q(\delta)$$

It was shown by Tsfasman, Vlăduţ and Zink [23, 22] that there exist asymptotically good geometric Goppa codes on modular curves that satisfy the TVZ bound

$$\delta + R \geq 1 - \frac{1}{\sqrt{q} - 1},$$

if q is a square. If moreover $q \geq 49$, then these codes are better than the GV bound. The theory of modular curves is very deep and the construction of these curves and their codes can be done in theory with polynomial complexity [19, 22] but are still too involved for an actual construction.

In this paper the latest results on the effective construction of asymptotically good codes and curves will be surveyed.

*Discrete Mathematics, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands

2 Algebraic geometry codes

Consider a geometric object $\mathcal{X}$ with a subset $\mathcal{P}$ consisting of n points which are enumerated by $P_1, \ldots, P_n$. Suppose that we have a vector space L over F_q of functions on $\mathcal{X}$ with values in F_q. Thus $f(P_i) \in F_q$ for all i and $f \in L$. In this way one has an evaluation map

$$ev_{\mathcal{P}} : L \longrightarrow F_q^n$$

which is defined by $ev_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n))$. If this evaluation map is linear, then its image is a linear code.

The classical example of the above situation is given by *Reed-Solomon (RS)* codes. Here the geometric object $\mathcal{X}$ is the affine line over F_q, the points are n distinct elements of F_q and L is the vector space of polynomials of degree at most $k-1$ and with coefficients in F_q. This code has parameters $[n, k, n-k+1]$ if $k \leq n$. The length of a RS code is at most q.

If we take as geometric object $\mathcal{X}$ the *affine space* of dimension m over F_q, for the set $\mathcal{P}$ all the q^m points of this affine space, and as vector space all polynomials of degree at most r, then we get the *Reed-Muller (RM)* codes of order r in m variables over F_q. The length of these codes is not bounded, but they do not give a sequence of asymptotically good codes.

Let the geometric object be an (affine) *variety*, that is to say the zero set in affine space of some polynomials . Every variety has a *dimension* and a variety of dimension one is called an *algebraic curve*. If $\mathcal{X}$ is an algebraic curve over F_q, $\mathcal{P}$ a set of n distinct points of $\mathcal{X}$ that are defined over F_q, and L a vector space of rational functions with a certain prescribed behaviour of their poles and zeros, then we get the *geometric Goppa* codes also known as *algebraic geometry* (AG) codes. The parameters of these codes are determined by the theorem of *Riemann-Roch*, and they satisfy the following bound

$$k + d \geq n + 1 - g,$$

where g is an invariant of the curve called its *genus*.

3 Asymptotically good sequences of curves and codes

The information rate R and the relative minimum distance δ of an algebraic geometry code on a curve of genus g with n points that are defined over F_q satisfy

$$R + \delta \geq 1 - \frac{g-1}{n}.$$

In order to construct good codes we therefore need curves with low genus and many F_q-rational points. For a curve over F_q of genus g with N F_q-rational

points the Hasse-Weil bound says

$$N \leq q + 1 + 2g\sqrt{q}.$$

Let

$$A(q) = \limsup_{g \to \infty} \frac{N_q(g)}{g},$$

where $N_q(g)$ is the maximal number of F_q-rational points on curves over F_q of genus g. The Hasse-Weil bound implies

$$A(q) \leq 2\sqrt{q}.$$

This has been improved to the Drinfeld-Vlăduţ (DV) bound [3]

$$A(q) \leq \sqrt{q} - 1 \tag{1}$$

When q is a square, Ihara [16] and Tsfasman, Vlăduţ, and Zink [23, 22] showed that

$$A(q) = \sqrt{q} - 1 \ ,$$

by studying the number of rational points on *modular curves* over finite fields. This in turn means that there exists a sequence of codes satisfying TVZ bound which is better than the GV bound when $q \geq 49$ in a certain range, and this fact was the starting point of the current interest in algebraic geometry codes.

Manin and Vlăduţ [19, 22] have shown that the construction of the modular curves and the corresponding codes can be done with polynomial complexity, of degree 20 for classical modular curves and degree 30 for Drinfeld modular curves. The degree for the latter has been reduced to 17 by Lopez [18]. But these constructions have no practical meaning. Others have tried to improve the complexity too [1, 2].

Apart from the fact that the construction has a high degree of complexity, the theory of modular curves is also deep and complex.

The paper of Justesen et al. [17] treats the construction and decoding of the class of codes on plane curves. This was the begining of an active period of research on decoding algorithms that culminated in the paper of Feng-Rao [4] on the majority voting among unknown syndromes. For a survey on this we refer to [15].

This history is mentioned because it gave the start of an elementary treatment of AG codes [5, 6, 8, 14] that is also known under the slogan:

"AG codes without AG".

Suppose that a decoding algorithm guarantees to find a unique closest code word for all received words with t errors, then the minimum distance is at least $2t + 1$. This observation lead to the Feng-Rao bound on the minimum distance, or more generally the order bound on the generalized Hamming weights [13].

4 The effective construction of asymptotically good sequences of codes

Many researchers have tried to construct asymptotically good sequences of curves in an elementary way and in a preprint [7] from 1994 by Feng and Rao the authors claimed to have found asymptotically good codes in an elementary way using what they called *generalized Klein curves* which are defined by the equations

$$X_i X_{i+1}^3 + X_{i+1} + X_i^3 = 0 \text{ for } i = 1, \ldots, m-1$$

over F_8. Notice that the equations are of the following type

$$F(X_i, X_{i+1}) = 0 \text{ for } i = 1, \ldots, m-1,$$

where

$$F(X, Y) = XY^3 + Y + X^3.$$

This is the defining polynomial of the affine Klein quartic. This curve has the property that for every given nonzero element $x \in F_q$ there are exactly 3 nonzero solutions in F_8 of the equation $F(x, Y) = 0$ in Y. Therefore by induction this gives a curve with $7 \cdot 3^{m-1}$ points with nonzero coordinates in F_8. Feng and Rao gave many more examples of polynomials $F(X, Y)$ over a finite field F_q of degree a in Y, and a subset S of F_q, they called a set of *complete selected roots*, such that for any given $x \in$S there exist exactly a distinct $y \in S$ such that $F(x, y) = 0$. In this way they gave a lot of examples of curves with $|S| \cdot a^{m-1}$ rational points.

From the point of view of algebraic geometry one must compute the genus of these curves. A sequence of curves $(\mathcal{X}_m | m \in N)$ is called asymptotically good if $g(\mathcal{X}_m)$ tends to infinity and the following limit exists and

$$\lim_{m \to \infty} \frac{N(\mathcal{X}_m)}{g(\mathcal{X}_m)} > 0,$$

where $g(\mathcal{X})$ is the genus of $\mathcal{X}$ and $N(\mathcal{X})$ is the number of F_q-rational points of $\mathcal{X}$. Frey, Perret and Stichtenoth [9] proved already some negative results in this direction. But this could not be applied to the generalized Klein curves. It was suggested to change the equations of the curves as follows

$$X_i X_{i+1}^2 + X_{i+1} + X_i^2 = 0 \text{ for } i = 1, \ldots, m-1$$

over F_4. It turned out that this gave a tower of so-called Artin-Schreier extensions and Garcia and Stichtenoth [10] generalized it to

$$X_i^{r-1} X_{i+1}^r + X_{i+1} - X_i^r = 0 \text{ for } i = 1, \ldots, m-1,$$

where $q = r^2$, and calculated the genera and the number of F_q-rational points. Thus they proved that the curves were asymptotically good and attain the DV bound.

Now the equations are of the form

$$F(X_i, X_{i+1}) = 0 \text{ for } i = 1, \ldots, m-1$$

with

$$F(X, Y) = X^{r-1}Y^r + Y - X^r.$$

The affine plane curve with equation $F(X, Y) = 0$ has the property that for every given nonzero element $x \in F_q$ there are exactly r nonzero solutions in F_q of the equation $F(x, Y) = 0$ in Y. This is seen by multiplying the equation by X and replacing XY by Z. Then the equation $Z^r + Z = X^{r+1}$ is obtained, which defines the Hermitian curve over F_q. Therefore the curve $\mathcal{C}_2$ has at least $(q-1)r$ points with nonzero coordinates in F_q. Consider the map

$$\pi_m : \mathcal{C}_m \longrightarrow \mathcal{C}_{m-1}$$

defined as $\pi(x_1, \ldots, x_{m-1}, x_m) = (x_1, \ldots, x_{m-1})$. If $(x_1, \ldots, x_{m-1})$ is a given F_q-rational point of $\mathcal{C}_{m-1}$ and $x_{m-1} \neq 0$, then there are exactly r possible nonzero values for $x_m \in F_q$ such that $(x_1, \ldots, x_{m-1}, x_m)$ is a point of $\mathcal{C}_m$. Therefore by induction it is shown that

$$N(\mathcal{C}_m) \geq (q-1)r^{m-1}.$$

The genus of the curve $\mathcal{C}_m$ is computed by induction by applying the formula of Hurwitz-Zeuthen to the covering $\pi_m : \mathcal{C}_m \to \mathcal{C}_{m-1}$. In this case it turns out to be an Artin-Schreier covering [21]. It is easier to view this in terms of function fields. Let F_m be the function field of $\mathcal{C}_m$. Then $F_1 = F_q(Z_1)$ and F_m is obtained from F_{m-1} by adjoining a new element Z_m that satisfies the equation

$$Z_m^r + Z_m = X_{m-1}^{r+1},$$

where $X_{m-1} = Z_{m-1}/X_{m-2} \in F_{m-1}$ for $m \geq 2$, and $X_1 = Z_1$, $X_0 = 1$. The genus g_m of the curve $\mathcal{C}_m$, or equivalently of the function field F_m is equal to

$$g_m = \begin{cases} r^m + r^{m-1} - r^{\frac{m+1}{2}} - 2r^{\frac{m-1}{2}} + 1 & \text{if } m \text{ is odd}, \\ r^m + r^{m-1} - \frac{1}{2}r^{\frac{m+2}{2}} - \frac{3}{2}r^{\frac{m}{2}} - r^{\frac{m-2}{2}} + 1 & \text{if } m \text{ is even}. \end{cases}$$

See [10]. Thus the DV bound is attained. Later it was shown that the tower of the generalized Klein curves is asymptotically bad [11]. This does not mean that the codes on these curves as defined in [7] are not asymptotically good. This has not been decided yet. Joint work with de Boer seems to indicate that they are not asymptotically good. The reason for this lies in the fact that a lot of functions are missing if we restrict ourselves to polynomials. One has to consider *rational functions*.

5 The quest of the missing functions

It turns out that finding bases for the vector spaces involved in the construction of AG codes is difficult. This last part remains to be done in order to make

the codes really constructive. A first step in this direction is made by Voss and Høholdt in [24] for the curve $\mathcal{C}_3$.

If the functions in L have only poles at P, an F_q-rational point not equal to $P_1, \ldots, P_n$, then we can associate to f its pole order at P which we denote by $\rho_P(f)$ or $\rho(f)$. The opposite of this function $-\rho_P(f)$ is equal to the discrete valuation of f at P. Let R be the ring of all rational functions on the curve $\mathcal{X}$ with poles only possibly at P. Then

$$\rho : R \longrightarrow N_0 \cup \{-\infty\}$$

is a function that satisfies the following porperties:

(0) $\rho(f) = -\infty$ if and only if $f = 0$
(1) $\rho(\lambda f) = \rho(f)$ for all nonzero $\lambda \in F$
(2) $\rho(f + g) \leq \max\{\rho(f), \rho(g)\}$
and equality holds when $\rho(f) < \rho(g)$.
(3) $\rho(fg) = \rho(f) + \rho(g)$
(4) If $\rho(f) = \rho(g)$, then there exists a nonzero $\lambda \in F$ such that $\rho(f - \lambda g) < \rho(g)$.

for all $f, g, h \in R$. Here $-\infty + n = -\infty$ for all $n \in N_0$. And is called a *weight function*.

Let $(\rho_l | l \in N)$ be the seqeunce of all nonnegative weights that appear in increasing order. Then $\rho_i < \rho_{i+1}$ for all i, and for all nonzero f there exists an i such that $\rho(f) = \rho_i$.

In this way we get an increasing sequence of codes

$$E(l) = \{ev_{\mathcal{P}}(f) | f \in R, \rho(f) \leq \rho_l\}$$

and a decreasing sequence of codes $C(l) = E(l)^{\perp}$.

The set $S = \{\rho_i | i \in N\}$ forms a *semigroup* . That is to say that $0 \in S$, and if $x, y \in S$, then $x + y \in S$. The parameters of the codes $E(l)$ and $C(l)$ can be determined in terms of properties of this semigroup. The elements of $N_0 \setminus S$ are called *gaps* and the *number of gaps* plays the same role as the genus.

The proof of the existence of a weight function in an elementary way works quite well for so-called plane curves of type I [5, 6] and their generalizations [20].

Consider the F_{64}-algebra

$$R = F_{64}[X, Y]/(X^5 + Y^4 + Y).$$

Assume that R has a weight function ρ. Let x and y be the cosets in R of X and Y, respectively. Then $x^5 = y^4 + y$. Now $y \notin F$, so $\rho(y) > 0$. and $\rho(y^4) = 4\rho(y) > \rho(y)$ by (5). Thus $\rho(y^4 + y) = \rho(y^4)$ by (2). Therefore

$$5\rho(x) = \rho(x^5) = \rho(y^4 + y) = 4\rho(y)$$

Thus the only possible solution is $\rho(x) = 4$ and $\rho(y) = 5$. One can show that there exists a weight function. See [14, 20].

Let $R = F_4[X, Y]/(XY^2 + Y + X^2)$. By the same reasoning as in the example above that $\rho(x) = 2$ and $\rho(y) = 1$ if there exists a weight function ρ on R. But one can show that there exists no weight funktion on R. The reason for this is that the second example has two points at infinity.

Let F be the field of fractions of R, so it is the function field of the curve. If we allowe negative values for the weight function, then we get a weight function ρ_P for every point P of the curve and $v_P = -\rho_P$ is the *discrete valuation* at P. To every rational function is associated its *principal divisor* (f) defined as

$$(f) = \sum_P v_P(f)P.$$

Consider the following ring

$$R = F_4[X_1, \ldots, X_m]/(X_i X_{i+1}^2 + X_i^2 + X_{i+1}, i = 1, 2, ..., m-1).$$

The coset of X_i in R is denoted by x_i. Now the principal divisors of the functions

$$x_m^{e_m} \cdots x_1^{e_1}.$$

will be investigated. Those monomials are tabulated that have only a pole at Q_m for $m = 2, 3$ and 4.

$m=2$:						$m=3$:						
e_2	e_1	Q_0	Q_1	Q_2		e_3	e_2	e_1	Q_0	Q_1	Q_2	Q_3
0	0	0	0	0		0	0	0	0	0	0	0
0	1	1	1	-2		0	0	1	1	1	2	-4
1	1	3	0	-3		0	1	1	3	3	0	-6
0	2	2	2	-4		0	0	2	2	2	4	-8
.	.	.	.	.		1	0	2	6	0	3	-9
0	i	i	i	-2i		0	1	2	4	4	2	-10
1	i	i+2	i-1	-2i-1		1	1	2	8	2	1	-11
.	.	.	.	.		0	0	3	3	3	6	-12
⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮

So the second table says for instance that:

$$(x_3 x_2 x_1^2) = 8Q_0 + 2Q_1 + Q_2 - 11Q_3.$$

If $m = 2$, then the genus is 1 and we have in the list one gap at 1. If $m = 3$, then the genus is 5 and we have in the list 5 gaps at $1, 2, 3, 5$ and 7.

$m = 4$:

e_4	e_3	e_2	e_1	Q_0	Q_1	Q_2	Q_3	Q_4
0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	2	4	-8
0	0	1	1	3	3	6	0	-12
0	0	0	2	2	2	4	8	-16
0	1	0	2	6	6	0	6	-18
0	0	1	2	4	4	8	4	-20
1	0	1	2	12	0	6	3	-21
0	1	1	2	8	8	4	2	-22
0	0	0	3	3	3	6	12	-24
0	1	0	3	7	7	2	10	-26
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

In fact Q_2 is the sum of two points. The genus is 13, but now there are 16 gaps at 1,2,3,4,5,6,7,9,10,11,13,14,15,17,19 and 25. We have to find three more functions. Leonard found the functions

$$X_3(1 + X_1X_4 + X_1X_2), \quad \frac{X_1X_3(1 + X_1X_4)}{X_2}, \quad \text{and } X_1X_3(1 + X_1X_4)$$

that have weights 14, 15 and 19, respectively. Notice that monomials are not sufficient anymore, even if we allow negative exponents. With the help of AXIOM, a computer algebra package, Haché [12] found many more missing functions for $m \geq 4$.

A new tower of function fields T_m is given in [11] over F_q with $q = r^2$, where $T_1 = F_q(X_1)$ and T_m is obtained from T_{m-1} by adjoining a new variable X_m that satisfies the equation:

$$X_m^r + X_m = \frac{X_{m-1}^r}{X_{m-1}^{r-1} + 1}.$$

This sequence of function fields attains the DV bound too.

In joint work with Stichtenoth and Torres we are able to give the missing functions for $m \leq 5$. To give some idea the following function is shown

$$\frac{X_3^2}{X_1} + \frac{\left(X_4 + \frac{X_3^2}{X_2}\right)^2}{X_3} + X_5.$$

References

[1] M. Bronstein, M. Hassner, A. Velasquez and C.J. Williamson, "Computer algebra algorithms for the construction of error correcting codes on algebraic curves," preprint 1991.

[2] W. Burge, M. Hassner, S. Watt and C.J. Williamson, "An efficient AXIOM representation of elliptic modular curves," preprint 1993.

[3] V.G. Drinfeld and S.G. Vlăduţ, "Number of points of an algebraic curve," *Func. Anal.*, vol. 17, pp. 53-54, 1983.

[4] G.-L. Feng and T.R.N. Rao, "Decoding of algebraic geometric codes up to the designed minimum distance," *IEEE Trans. Inform. Theory*, vol. IT-39, pp. 37-45, Jan. 1993.

[5] G.-L. Feng and T.R.N. Rao, "A simple approach for construction of algebraic-geometric codes from affine plane curves," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 1003-1012, July 1994.

[6] G.-L. Feng and T.R.N. Rao, "Improved geometric Goppa codes," Part I: Basic Theory, *IEEE Trans. Inform. Theory*, Nov. 1995, pp. 1678-1693.

[7] G.-L. Feng and T.R.N. Rao, Improved geometric Goppa codes, Part II, Generalized Klein codes, peprint 1994.

[8] G.-L. Feng, V. Wei, T.R.N. Rao and K.K. Tzeng, "Simplified understanding and efficient decoding of a class of algebraic-geometric codes," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 981-1002, July 1994.

[9] G. Frey, M. Perret and H. Stichtenoth, "On the different of Abelian extensions of global fields," In Coding Theory and Algebraic Geometry, Luminy 1991, *Springer Lect. Notes Math.*, vol. 1518, pp. 26-32, 1992.

[10] A. Garcia and H. Stichtenoth, "A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vlăduţ bound", *Invent. Math.*, vol. 121, pp. 211-222, 1995.

[11] A. Garcia and H. Stichtenoth, "On the asymptotic behaviour of some towers of function fields over finite fields," to appear in *Journ. Number Theory.*

[12] G. Haché, *Ph.D. Thesis,* INRIA, Univ. Paris VI, Sept. 1996.

[13] P. Heijnen and R. Pellikaan, "Generalized Hamming weights of q-ary Reed-Muller codes," preprint June 1996.

[14] T. Høholdt, J.H. van Lint and R. Pellikaan, "Algebraic geometry codes," preprint 1996.

[15] T. Høholdt and R. Pellikaan, "On the Decoding of Algebraic-Geometric Codes," *IEEE Trans. Inform. Theory,* vol. 41, pp. 1589-1614, nov. 1995.

[16] Y. Ihara, "Some remarks on the number of rational points of algebraic curves of finite fields," *Journ. Fac. Sc. Univ. Tokyo* IA, vol. 28, pp. 721-724, 1981.

[17] J. Justesen, K.J. Larsen, H. Elbrønd Jensen, A. Havemose and T. Høholdt, "Construction and decoding of a class of algebraic geometric codes," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 811-821, July 1989.

[18] B. López Jiménez, "Plane models of Drinfeld modular curves," *Ph.D. Thesis,* Unver. Complutense, Madrid, March 1996.

[19] Yu.I. Manin and S.G. Vlăduţ, "Linear codes and modular curves," *Journ. Sov. Math.* vol. 30, pp.2611-2643, 1985.

[20] R. Pellikaan, "On the existence of order functions," In *Proc. 2nd Shanghai Conference on Designs, Codes and Finite Geometry* 1996.

[21] H. Stichtenoth, *Algebraic function fields and codes*, Universitext, Springer-Verlag, Berlin 1993.

[22] M.A. Tsfasman and S.G. Vlăduţ, *Algebraic-geometric codes*, Mathematics and its Applications vol. 58, Kluwer Acad. Publ., Dordrecht 1991.

[23] M.A. Tsfasman, S.G. Vlăduţ and T. Zink, "Modular curves, Shimura curves and Goppa codes, better than Varshamov-Gilbert bound," *Math. Nachrichten*, vol. 109, pp. 21-28, 1982.

[24] C. Voss and T. Høholdt, "An explicit construction of a sequence of codes attaining the Tsfasman-Vlăduţ-Zink bound. The first steps," to appear in *IEEE Trans. Inform. Theory.*