

Public key cryptosystems based on algebraic geometry codes

Ruud Pellikaan

g.r.pellikaan@tue.nl

joint work with

Alain Couvreur, Irene Márquez-Corbella

Edgar Martínez-Moro and Diego Ruano

Combinatorial Theory Seminar

Eindhoven, 23 April 2014

- ▶ Error-correcting codes
- ▶ Codes on curves
- ▶ Error-correcting pairs **for** codes on curves
- ▶ Public-key cryptosystems
- ▶ Code-based cryptography
- ▶ Reverse engineering AG codes
- ▶ Error-correcting pairs **from** codes on curves
- ▶ Questions

C linear block code: \mathbb{F}_q -linear subspace of \mathbb{F}_q^n

parameters $[n, k, d]$:

n = length

k = dimension of C

d = minimum distance of C

$$d = \min |\{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}\}|$$

t = error-correcting capacity of C

$$t = \lfloor \frac{d-1}{2} \rfloor$$

The **standard inner product** is defined by

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + \cdots + a_n b_n$$

For two subsets A and B of \mathbb{F}_q^n

$A \perp B$ if and only if $\mathbf{a} \cdot \mathbf{b} = 0$ for all $\mathbf{a} \in A$ and $\mathbf{b} \in B$

Let \mathbf{a} and \mathbf{b} in \mathbb{F}_q^n

The **star product** is defined by coordinatewise multiplication:

$$\mathbf{a} * \mathbf{b} = (a_1 b_1, \dots, a_n b_n)$$

For two subsets A and B of \mathbb{F}_q^n

$$A * B = \langle \{\mathbf{a} * \mathbf{b} \mid \mathbf{a} \in A \text{ and } \mathbf{b} \in B\} \rangle \text{ and } A^{(2)} = A * A$$

Let C be a linear code in \mathbb{F}_q^n

The pair (A, B) of linear subcodes of $\mathbb{F}_{q^m}^n$ is called a **t-error correcting pair (ECP)** over \mathbb{F}_{q^m} for C if

E.1 $(A * B) \perp C$

E.2 $k(A) > t$

E.3 $d(B^\perp) > t$

E.4 $d(A) + d(C) > n$

Let A and B be linear subspaces of \mathbb{F}_q^n

Let $r \in \mathbb{F}_q^n$ be a **received word**

Define the **kernel of error-locator vectors**

$$K(r) = \{ a \in A \mid (a * b) \cdot r = 0 \text{ for all } b \in B \}$$

Lemma

Let C be an \mathbb{F}_q -linear code of length n

Let r be a received word with **error vector e**

So $r = c + e$ for some $c \in C$

If $A * B \subseteq C^\perp$, then

$$K(r) = K(e)$$

Let (A, B) be a t -ECP for C and J a subset of $\{1, \dots, n\}$
Define the subspace of A

$$A(J) = \{ \mathbf{a} \in A \mid a_j = 0 \text{ for all } j \in J \}$$

Set of zeros of error-locator vectors contains the error positions:

Lemma

Let $(A * B) \perp C$

Let \mathbf{e} be an error vector of the received word \mathbf{r}

If $I = \text{supp}(\mathbf{e}) = \{ i \mid e_i \neq 0 \}$, then

$$A(I) \subseteq K(\mathbf{r})$$

If moreover $d(B^\perp) > \text{wt}(\mathbf{e})$, then $A(I) = K(\mathbf{r})$

Theorem

Let C be an \mathbb{F}_q -linear code of length n

Let (A, B) be a t -error-correcting pair over \mathbb{F}_{q^m} for C

Then the basic algorithm corrects t errors
for the code C with complexity $\mathcal{O}((mn)^3)$

Let \mathcal{X} be an **algebraic curve** defined over \mathbb{F}_q of **genus g**
 $\mathcal{X}(\mathbb{F}_q)$ is the set of **\mathbb{F}_q -rational points** of \mathcal{X}

Let $\mathbb{F}_q(\mathcal{X})$ be the vector space of **rational functions** on \mathcal{X} .

Let f be a rational function and P a **place**

$v_P(f)$ is the **valuation** of f at P

$(f) = \sum_P v_P(f)P$ is the **principal divisor** f

Let $E = \sum m_P P$ be a **divisor**, a finite formal sum of places

$\deg(E) = \sum m_P \deg(P)$ is the **degree** of E

Let $L(E) = \{f \in \mathbb{F}_q(\mathcal{X}) \mid (f) \geq -E, f \neq 0\} \cup \{0\}$

Riemann-Roch: $\dim L(E) \geq \deg(E) + 1 - g$

equality holds if $\deg(E) > 2g - 2$

Let \mathcal{X} be an algebraic curve defined over \mathbb{F}_q of genus g

Let $\mathcal{P} = (P_1, \dots, P_n)$ an n -tuple of mutual distinct points of $\mathcal{X}(\mathbb{F}_q)$

(If the support of E is disjoint from \mathcal{P}), then the **evaluation map**

$$\text{ev}_{\mathcal{P}} : L(E) \rightarrow \mathbb{F}_q^n$$

where $\text{ev}_{\mathcal{P}}(f) = (f(P_1), \dots, f(P_n))$, is well defined.

The **algebraic geometry code** $C_L(\mathcal{X}, \mathcal{P}, E)$

is the image of $L(E)$ under the evaluation map $\text{ev}_{\mathcal{P}}$

If $m < n$, then $C_L(\mathcal{X}, \mathcal{P}, E)$ is an $[n, k, d]$ code with

$$k \geq m + 1 - g \text{ and } d \geq n - m$$

$n - m$ is called the **designed minimum distance** of $C_L(\mathcal{X}, \mathcal{P}, E)$

Embedding of \mathcal{X} in **linear system** of E of degree m

Let f_1, f_2, \dots, f_k be a basis of $L(E)$

$$\varphi_E : \mathcal{X} \longrightarrow \mathbb{P}^{k-1}$$

$$P \mapsto (f_1(P) : f_2(P) : \dots : f_k(P))$$

$\mathcal{Y} = \varphi_E(\mathcal{X})$ is a curve of degree m in \mathbb{P}^{k-1}

$\mathcal{Q} = (\varphi_E(P_1), \dots, \varphi_E(P_n))$ **projective system**

$$G_{\mathcal{Q}} = \begin{pmatrix} f_1(P_1) & \cdots & f_1(P_j) & \cdots & f_1(P_n) \\ f_2(P_1) & \cdots & f_2(P_j) & \cdots & f_2(P_n) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ f_k(P_1) & \cdots & f_k(P_j) & \cdots & f_k(P_n) \end{pmatrix} \text{generator matrix}$$

minimum distance $\geq n - m$

Let F and G be divisors

Then there is a well defined linear map

$$L(F) \otimes L(G) \longrightarrow L(F + G)$$

given on generators by

$$f \otimes g \mapsto fg$$

Hence

$$C_L(\mathcal{X}, \mathcal{P}, F) * C_L(\mathcal{X}, \mathcal{P}, G) \subseteq C_L(\mathcal{X}, \mathcal{P}, F + G)$$

Let $C = C_L(\mathcal{X}, \mathcal{P}, E)^\perp$

Choose a divisor F with support disjoint from \mathcal{P}

Let $A = C_L(\mathcal{X}, \mathcal{P}, F)$

Let $B = C_L(\mathcal{X}, \mathcal{P}, E - F)$

Then

- $A * B \subseteq C^\perp$

- If $t + g \leq \deg(F) < n$, then $k(A) > t$

- If $\deg(G - F) > t + 2g - 2$, then $d(B^\perp) > t$

- If $\deg(G - F) > 2g - 2$, then $d(A) + d(C) > n$

Proposition

An algebraic geometry code of designed minimum distance d from a curve over \mathbb{F}_q of genus g has a t -error-correcting pair over \mathbb{F}_q where

$$t = \lfloor \frac{d-1-g}{2} \rfloor$$

Proposition

An algebraic geometry code of designed minimum distance d from a curve over \mathbb{F}_q of genus g has a t -error-correcting pair over \mathbb{F}_{q^m} where

$$t = \lfloor \frac{d-1}{2} \rfloor$$

if

$$m > \log_q (2 \binom{n}{t} + 2 \binom{n}{t+1} + 1)$$

Not constructive!

Majority coset decoding gives a constructive and efficient approach

At the heart of any **public-key cryptosystem** is a **one-way function**

$$y = f(x)$$

that is **easy to evaluate** but for which it is **computationally infeasible** (**one assumes/hopes**) to **find an inverse**

$$x \in f^{-1}(y)$$

PKC systems use **trapdoor one-way functions**

by mathematical problems that are (**supposedly**) **hard**

- RSA, **factoring integers**
- Diffie-Hellman, **discrete-log problem** in finite field
- Elliptic curve PKC, **addition on elliptic curve**
- Lattice-based PKC systems **closest vector problem**
- Code-based PKC systems, **decoding of codes**

Code based PKC systems

Take a class of codes that have an efficient decoding algorithm:

Scramble a generator matrix such that it looks like a random code

- Goppa codes (McEliece)
- with parity check matrix instead of generator matrix (Niederreiter)
- Algebraic geometry codes (Janwa-Moreno)
- subcodes of GRS codes (Berger-Loidreau)
- subfield subcodes of algebraic geometry codes (Janwa-Moreno)

Generic attack – decoding algorithms:

- McEliece 1978
-
- Finiasz-Sendrier 2009
- Bernstein-Lange-Peters 2008-2011
- Becker-Joux-May-Meurer Eurocrypt 2012

Structural attack:

- GRS codes (Sidelnikov-Shestakov)
- subcodes of GRS codes (Wieschebrink, Márquez-Martínez-P)
- Alternant codes: open
- Goppa codes: open
- AG codes: (Faure-Minder, $g \leq 2$)
- VSAG codes: (Couvreur-Márquez-Martínez-P-Ruano, arbitrary g)
- subcodes of AG codes: open

Let \mathcal{X} be an absolutely irreducible and nonsingular curve of genus g over the perfect field \mathbb{F}

Let E be a divisor on \mathcal{X} of degree m

If $m \geq 2g + 1$

then φ_E gives an embedding of \mathcal{X} onto $\mathcal{Y} = \varphi_E(\mathcal{X})$
which is a normal curve in the linear system $|E| = \mathbb{P}^{m-g}$

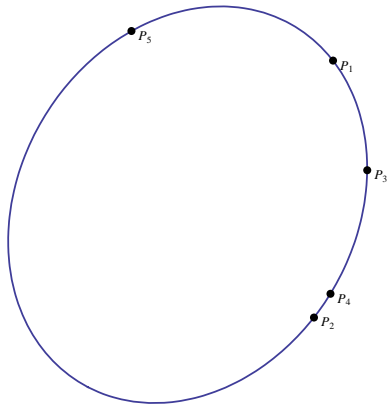
If $m \geq 2g + 2$, then \mathcal{Y} is an intersection of quadrics

More precisely:

$I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$

the set of homogeneous elements of degree two in $I(\mathcal{Y})$

Conic determined by 5 points



Let \mathcal{Y} be a curve embedded in projective r -space of degree m

Let $I(\mathcal{Y})$ be the vanishing ideal of \mathcal{Y}

Let \mathcal{Q} be a subset of \mathcal{Y} of n points

Then

$$I(\mathcal{Y}) \subseteq I(\mathcal{Q})$$

Hence

$$I_2(\mathcal{Y}) \subseteq I_2(\mathcal{Q})$$

Suppose $I(\mathcal{Y})$ is generated by $I_2(\mathcal{Y})$

$$\text{If } n > 2m, \text{ then } I_2(\mathcal{Y}) = I_2(\mathcal{Q})$$

By Bézout's Theorem

$\mathbf{g}_1, \dots, \mathbf{g}_k$ a basis of C

$S^2(C)$ is the **second symmetric power** of C

$S^2(C)$ has basis $\{X_i X_j \mid 1 \leq i \leq j \leq n\}$ and dimension $\binom{k+1}{2}$
with $X_i = \mathbf{g}_i$

$C^{(2)} = C * C$ the **square** of C

Consider the linear map

$$\begin{aligned} \sigma : S^2(C) &\longrightarrow C^{(2)} \\ X_i X_j &\longmapsto \mathbf{g}_i * \mathbf{g}_j \end{aligned}$$

$K_2(C)$ is the **kernel** of this map

Then

$$0 \longrightarrow K_2(C) \longrightarrow S^2(C) \longrightarrow C^{(2)} \longrightarrow 0$$

is an exact sequence and

$$I_2(Q) = K_2(C) := \left\{ \sum_{1 \leq i < j \leq k} a_{ij} X_i X_j \mid \sum_{1 \leq i < j \leq k} a_{ij} g_i * g_j = 0 \right\}$$

Proposition

Let Q be an n -tuple of points in \mathbb{P}^r over \mathbb{F} not in a hyperplane

Then the complexity of the computation of $I_2(Q)$ is at most $\mathcal{O}(n^4)$

C is called **very strong algebraic-geometric (VSAG)**

if $C = C_L(\mathcal{X}, \mathcal{P}, E)$ and the curve \mathcal{X} has **genus g**
 \mathcal{P} consists of **n points** and E has **degree m** such that

$$2g + 2 \leq m < \frac{1}{2}n \quad \text{or} \quad \frac{1}{2}n + 2g - 2 < m \leq n - 4$$

The dual of a VSAG code is again VSAG

Main Theorem

Let C be a VSAG code

Then a VSAG representation of C can be obtained efficiently from its generator matrix

Moreover all VSAG representations of C are strict isomorphic

Shortcut via t -ECP pair (A, B) in \mathbb{F}_q^n

Bypassing computation of triple $(\mathcal{X}, \mathcal{P}, E)$ and Riemann–Roch spaces

	$\mathbb{F}_q(\mathcal{X})$	\mathbb{F}_q^n
		$C = C_L(\mathcal{X}, \mathcal{P}, E)^\perp$
$(\mathcal{X}, \mathcal{P}, E)$	$L(E)$	$C_L(\mathcal{X}, \mathcal{P}, E)$
$(\mathcal{X}, \mathcal{P}, iP_1)$	$L(iP_1)$	$A_i = C_L(\mathcal{X}, \mathcal{P}, iP_1)$
$(\mathcal{X}, \mathcal{P}, E - jP_1)$	$L(E - jP_1)$	$B_j = C_L(\mathcal{X}, \mathcal{P}, E - jP_1)$

In fact, B_j is the space of those code words in C^\perp that are **zero** with **multiplicity** j at P_1

This multiplicity can be controlled since we computed $I_2(Q)$ efficiently

Proposition

$$A_j = \langle B_j * C \rangle^\perp$$

(A_{t+g}, B_{t+g}) is a t -ECP for C with $t = \lfloor (d^* - 1 - g)/2 \rfloor$

Still reference to multiplicities

Circumventing multiplicities:

Let $A_i = C_L(\mathcal{X}, \mathcal{P}, iP_1)$ and $B_j = C_L(\mathcal{X}, \mathcal{P}, E - jP_1)$

Then $B_0 = C_L(\mathcal{X}, \mathcal{P}, E) = C^\perp$

And $B_1 = C_L(\mathcal{X}, \mathcal{P}, E - P_1)$,

the space of those code words in C^\perp that are zero at the first position

So B_0 and B_1 are easily computed for given C

The B_j are obtained as follows **by induction**

Proposition

$$B_{j+1} = \{ z \in B_j \mid z * B_{j-1} \subseteq B_j * B_j \}$$

$$A_i = \langle B_i * C \rangle^\perp$$

(A_{t+g}, B_{t+g}) is a $t - \text{ECP}$ for C with $t = \lfloor (d^* - 1 - g)/2 \rfloor$

