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

Universitat Rovira i Virgili, Tarragona
10 November 2014

Introduction and content

- Error correcting pair
- Algebraic geometry codes (AGC)
- Error correcting pair for AG codes
- Code based public key crypto system
- Reverse engineering AG codes
- Error correcting pair from AG codes
- Error correcting pair from subcodes of AG codes
- Questions

Error correcting codes

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

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

Inner product

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 $\mathbf{b} \in B$

Dual code:

$$C^\perp = \{ \mathbf{x} \mid \mathbf{x} \cdot \mathbf{c} = 0 \text{ for all } \mathbf{c} \in C \}$$

Star product

Let $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{F}_q^n$
The star product is defined by coordinatewise multiplication:

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

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

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

$$A^{(2)} = A * A$$

Generalized Reed-Solomon codes - 1

Let $\mathbf{a} = (a_1, \ldots, a_n)$ be an n-tuple of mutually distinct elements of $\mathbb{F}_q$
Let $\mathbf{b} = (b_1, \ldots, b_n)$ be an n-tuple of nonzero elements of $\mathbb{F}_q$

Evaluation map:

$$\mathrm{ev}_{\mathbf{a},\mathbf{b}}(f(X)) = (f(a_1)b_1, \ldots, f(a_n)b_n)$$

$$GRS_k(\mathbf{a}, \mathbf{b}) = \{\ \mathrm{ev}_{\mathbf{a},\mathbf{b}}(f(X)) \mid f(X) \in \mathbb{F}_q[X], \deg(f(X) < k\ \}$$

Parameters: $[n, k, n-k+1]$ if $k \leq n$

Generalized Reed-Solomon codes - 2

Furthermore

$$\mathrm{ev}_{\mathbf{a},\mathbf{b}}(f(X)) * \mathrm{ev}_{\mathbf{a},\mathbf{c}}(g(X)) = \mathrm{ev}_{\mathbf{a},\mathbf{b}*\mathbf{c}}(f(X)g(X))$$

$$GRS_k(\mathbf{a},\mathbf{b}) * GRS_l(\mathbf{a},\mathbf{c}) = GRS_{k+l-1}(\mathbf{a},\mathbf{b}*\mathbf{c})$$

Error correcting pair - 1

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 a 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 $d(A) + d(C) > n$

t-ECP for $GRS_{n-2t}(\mathbf{a}, \mathbf{b})$

Let $C^{\perp} = GRS_{n-2t}(\mathbf{a}, \mathbf{b})$, has parameters: $[n, 2t, n-2t+1]$
Then $C = GRS_{2t}(\mathbf{a}, \mathbf{c})$ for some $\mathbf{c}$
has parameters: $[n, n-2t, 2t+1]$

Let $A = GRS_{t+1}(\mathbf{a}, \mathbf{1})$ and $B = GRS_{t}(\mathbf{a}, \mathbf{b})$
Then $(A * B) \subseteq C^{\perp}$

A has parameters $[n, t+1, n-t]$
B has parameters $[n, t, n-t+1]$
So $B^{\perp}$ has parameters $[n, n-t, t+1]$

Hence (A, B) is a t-error correcting pair for C

Error correcting pair - 2

Let A and B be linear subspaces of $\mathbb{F}_{q^m}^n$
Let $\mathbf{r} \in \mathbb{F}_q^n$ be a received word
Define the kernel of error locator vectors

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

Lemma
Let C be an $\mathbb{F}_q$-linear code of length n
Let $\mathbf{r}$ be a received word with error vector $\mathbf{e}$
So $\mathbf{r} = \mathbf{c} + \mathbf{e}$ for some $\mathbf{c} \in C$
If $A * B \subseteq C^{\perp}$, then

$$K(\mathbf{r}) = K(\mathbf{e})$$

Error correcting pair - 3

Let (A, B) be a t-ECP for C and J a subset of $\{1, \ldots, n\}$
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 $A(I) = K(\mathbf{r})$

Error correcting pair - 4

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 $\mathcal{O}((mn)^3)$

Algebraic geometry codes - 1

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 of f

Algebraic geometry codes - 2

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

Riemann-Roch space

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

Riemann-Roch:

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

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

Algebraic geometry codes - 3

Let $\mathcal{X}$ be an algebraic curve defined over $\mathbb{F}_q$ of genus g
Let $\mathcal{P} = (P_1, \ldots, P_n)$ 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

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

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

The algebraic geometry code $C_L(\mathcal{X}, \mathcal{P}, E)$
is the image of $L(E)$ under the evaluation map $\mathrm{ev}_{\mathcal{P}}$
If $m < n$, then $C_L(\mathcal{X}, \mathcal{P}, E)$ is an $[n, k, d]$ code with

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

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

Algebraic geometry codes - 4

Embedding of $\mathcal{X}$ in linear system of E of degree m
Let $f_1, f_2, \ldots, f_k$ be a basis of $L(E)$

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

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

$\mathcal{Y} = \varphi_E(\mathcal{X})$ is a curve of degree m in $\mathbb{P}^{k-1}$
$\mathcal{Q} = (\varphi_E(P_1), \ldots, \varphi_E(P_n))$ projective system

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

minimum distance $\geq n - m$

Algebraic geometry codes - 5

Let ω be a differential form
with a simple pole and residue 1 at P_j for all $j = 1, \ldots, n$
Let K be the canonical divisor of ω

Then

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

where $E^{\perp} = P_1 + \cdots + P_n + K - E$
and $\deg(E^{\perp}) = n - m + 2g - 2$

minimum distance is at least
$d^* = m - 2g + 2$
the designed minimum distance

ECP for AG codes - 1

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

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

Equality holds if $\deg(F) \geq 2g$ and $\deg(G) \geq 2g+1$

ECP for AG codes - 2

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

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

ECP for AG codes - 4

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

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

Not constructive!

Coset decoding

Feng-Rao, Duursma

Let C be a code for which we need a decoding algorithm
Let D be a subcode for which we have a decoding algorithm

Coset decoding is an algorithm
Input: $\mathbf{x}$ such that $\mathbf{x} = \mathbf{e} + \mathbf{c}$ and $\mathbf{c} \in C$
Output: $\mathbf{y}$ such that $\mathbf{y} = \mathbf{e} + \mathbf{d}$ and $\mathbf{d} \in D$

Solution:
– Majority voting of unknown syndromes
– Majority coset decoding
– Error correcting array

Code based public key crypto (PKC) system

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)

Reverse engineering AG codes - 1

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

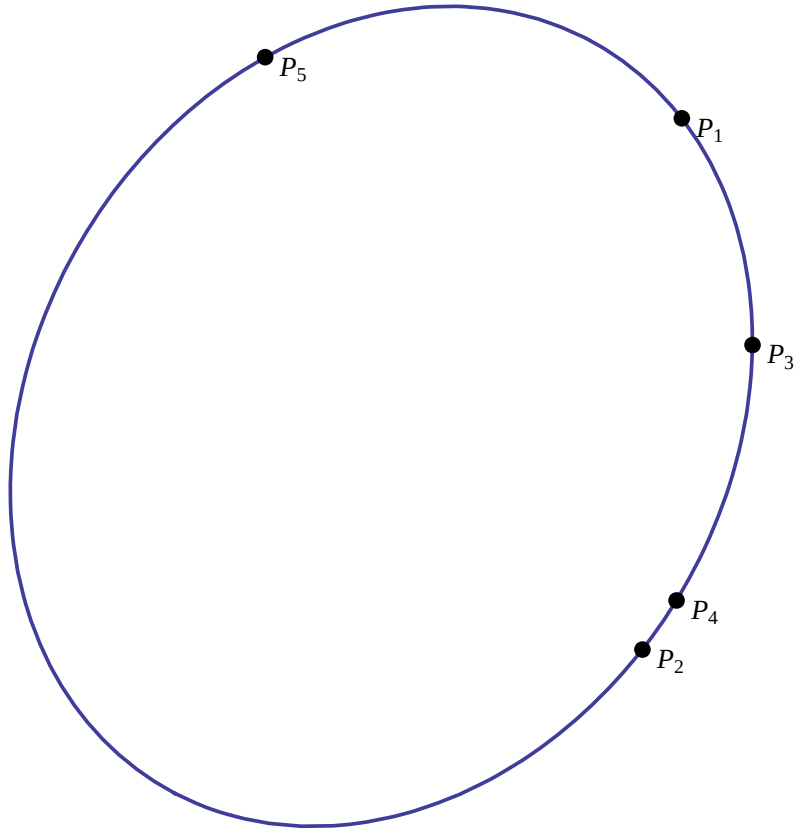
Let E be a divisor on $\mathcal{X}$ of degree m

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 $I(\mathcal{Y})$

Reverse engineering AG codes - 2

Conic determined by 5 points



Reverse engineering AG codes - 3

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

Reverse engineering AG codes - 4

$\mathbf{g}_1, \ldots, \mathbf{g}_k$ 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

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

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

Reverse engineering AG codes - 5

Then

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

is an exact sequence and

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

Proposition

Let $\mathcal{Q}$ be an n-tuple of points in $\mathbb{P}^r$ over $\mathbb{F}$ not in a hyperplane
Then the complexity of the computation of $I_2(\mathcal{Q})$ is at most $\mathcal{O}(n^4)$

Reverse engineering AG codes - 6

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

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

The dual of a VSAG code is again VSAG

Reverse engineering AG codes - 7

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

Error correcting pair from VSAG code - 1

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

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

Error correcting pair from VSAG code - 2

In fact, D_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 $I_2(\mathcal{Q})$ efficiently

Proposition
Let $A_i := \langle D_i * C \rangle^\perp$, then
(A_{t+g}, D_{t+g}) is a t-ECP for C with $t = \lfloor (d^* - 1 - g)/2 \rfloor$

Still reference to multiplicities

Error correcting pair from VSAG code - 3

Circumventing multiplicities altogether :

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

Then $D_0 = C_L(\mathcal{X}, \mathcal{P}, E) = C^{\perp}$

And $D_1 = C_L(\mathcal{X}, \mathcal{P}, E - P_1)$,

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

So D_0 and D_1 are easily computed for given C

The D_j are obtained as follows by induction

Proposition

$$D_{j+1} = \{ \mathbf{z} \in D_j \mid \mathbf{z} * D_{j-1} \subseteq D_j^{(2)} \}$$

$$A_i = \langle D_i * C \rangle^{\perp}$$

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

Subcodes of AG codes - 1

Let $C \subset \mathbb{F}_q^n$ be a code and $t \geq 2$ be an integer
The t-closure of C is defined by

$$\overline{C}^t = \left\{ \mathbf{a} \in \mathbb{F}_q^n \mid \mathbf{a} * C^{(t-1)} \subseteq C^{(t)} \right\}.$$

The code C is said to be t-closed if $\overline{C}^t = C$

Proposition
Let E be a divisor and $2g + 1 \leq \deg(E) \leq \frac{n-2}{2}$
Then

$$\overline{C_L(\mathcal{X}, \mathcal{P}, E)}^2 = C_L(\mathcal{X}, \mathcal{P}, E)$$

Subcodes of AG codes - 2

$$\text{Let } 2g+1 \leq \deg(E) \leq \frac{n-1}{2}, \quad k = \deg(E) + 1 - g$$

The public key consists of the subcode

$$C \subseteq C_L(\mathcal{X}, \mathcal{P}, E) \text{ and } l := \dim C$$

Assume

$$2k - 1 + g \leq \binom{l+1}{2}$$

Then with a high probability we have that

$$C^{(2)} = C_L(\mathcal{X}, \mathcal{P}, E) \text{ and } \overline{C}^2 = C_L(\mathcal{X}, \mathcal{P}, E)$$

Conclusion – Questions

– Algebraic geometry codes are not suitable for a McEliece PKC
– Also certain subcodes of AG are not suitable for a McEliece PKC

– What about subfield subcodes of AG codes?

– What about codes from varieties of dimension lager than 1?
– What about Reed-Muller and order domain codes?