Error-correcting Pairs for a Public-key Cryptosystem

Ruud Pellikaan
g.r.pellikaan@tue.nl
joint work with
Irene Márquez-Corbella

Code-based Cryptography Workshop 2012
Lyngby, 9 May 2012
Introduction and content

- Error-correcting pair
  - Generalized Reed-Solomon codes
  - Alternant codes
  - Goppa codes
- $t$-error-correcting pair corrects $t$-errors
- Algebraic geometry codes
- Code-based cryptography
Error-correcting codes

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

parameters $[n, k, d]$:
- $n = \text{length}$
- $k = \text{dimension of } C$
- $d = \text{minimum distance of } C$

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

$t = \text{error-correcting capacity of } C$

$$t = \left\lfloor \frac{d(C) - 1}{2} \right\rfloor$$
Inner and star product

The standard inner product is defined by

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

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

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

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

The star product is defined by coordinatewise multiplication:

$$a \ast b = (a_1 b_1, \ldots, a_n b_n)$$

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

$$A \ast B = \{a \ast b \mid a \in A \text{ and } b \in B\}$$
Error-correcting pairs

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 $(A \ast B) \perp C$
E.2 $k(A) > t$
E.3 $d(B^\perp) > t$
E.4 $d(A) + d(C) > n$
Generalized Reed-Solomon codes

Let \( a = (a_1, \ldots, a_n) \) be an \( n \)-tuple of \textit{mutually distinct} elements of \( \mathbb{F}_q \).

Let \( b = (b_1, \ldots, b_n) \) be an \( n \)-tuple of \textit{nonzero} elements of \( \mathbb{F}_q \).

**Evaluation map:**

\[
ev_{a,b}(f(X)) = (f(a_1)b_1, \ldots, f(a_n)b_n)
\]

\( GRS_k(a, b) = \{ \ev_{a,b}(f(X)) \mid f(X) \in \mathbb{F}_q[X], \deg(f(X)) < k \} \)

**Parameters:** \([n, k, n - k + 1]\) if \( k \leq n \)

Furthermore

\[
ev_{a,b}(f(X)) \ast ev_{a,c}(g(X)) = ev_{a,b}(f(X)g(X)) \ast c
\]

\[
\langle GRS_k(a, b) \ast GRS_l(a, c) \rangle = GRS_{k+l-1}(a, b \ast c)
\]
**$t$-ECP for $\text{GRS}_{n-2t}(a, b)$**

Let $C = \text{GRS}_{n-2t}(a, b)$

Then $C$ has parameters: $[n, n - 2t, 2t + 1]$

and $C^\perp = \text{GRS}_{2t}(a, c)$ for some $c$

Let $A = \text{GRS}_{t+1}(a, 1)$ and $B = \text{GRS}_{t}(a, c)$

Then $A \ast B \subseteq C^\perp$

$A$ 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 $C$

Conversely an $[n, n - 2t, 2t + 1]$ code that has a $t$-ECP is a GRS code
Alternant codes

Let $a$ be an $n$-tuple of mutually distinct elements of $\mathbb{F}_{q^m}$
Let $b$ be an $n$-tuple of nonzero elements of $\mathbb{F}_{q^m}$

Let $GRS_k(a, b)$ be the GRS code over $\mathbb{F}_{q^m}$ of dimension $k$

The **alternant code** $ALT_r(a, b)$ is the $\mathbb{F}_q$-linear restriction

$$ALT_r(a, b) = \mathbb{F}_q^n \cap (GRS_r(a, b))^\perp$$

Then $ALT_r(a, b)$ has parameters $[n, k, d]_q$ with

$$k \geq n - mr \text{ and } d \geq r + 1$$

Every linear code of minimum distance at least 2 is an alternant code!
Let $C = ALT_{2t}(a, b)$
Then $C$ has minimum distance $d \geq 2t + 1$
and $C \subseteq (GRS_{2t+1}(a, b))^\perp$

Let $A = GRS_{t+1}(a, 1)$ and $B = GRS_{t}(a, b)$
Then $A \ast B \subseteq GRS_{2t+1}(a, b)$
Then $(A \ast B) \perp C$

$A$ 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 over $\mathbb{F}_{q^m}$ for $C$
Goppa codes

Let $L = (a_1, \ldots, a_n)$ be an $n$-tuple of $n$ distinct elements of $\mathbb{F}_{q^m}$
Let $g$ be a polynomial with coefficients in $\mathbb{F}_{q^m}$ such that

$$g(a_j) \neq 0 \text{ for all } j$$

Then $g$ is called **Goppa polynomial** with respect to $L$

Define the $\mathbb{F}_q$-linear **Goppa code** $\Gamma(L, g)$ by

$$\Gamma(L, g) = \left\{ c \in \mathbb{F}_q^n \mid \sum_{j=1}^{n} \frac{c_j}{X - a_j} \equiv 0 \text{ mod } g(X) \right\}$$
Goppa codes are alternant codes

Let $L = a = (a_1, \ldots, a_n)$
Let $g$ be a Goppa polynomial of degree $r$

Let $b_j = 1/g(a_j)$
Then

$$\Gamma(L, g) = \text{ALT}_r(a, b)$$

Hence $\Gamma(L, g)$ has parameters $[n, k, d]_q$ with

$$k \geq n - mr \text{ and } d \geq r + 1$$

and has an $\lfloor r/2 \rfloor$-error-correcting pair
Binary Goppa codes

Let $L = a = (a_1, \ldots, a_n)$

Let $g$ be a Goppa polynomial with coefficients in $\mathbb{F}_{2^m}$ of degree $r$

Suppose moreover that $g$ has no square factor

Then

$$\Gamma(L, g) = \Gamma(L, g^2)$$

Hence $\Gamma(L, g)$ has parameters $[n, k, d]_q$ with

$$k \geq n - mr \text{ and } d \geq 2r + 1$$

and has an $r$-error-correcting pair
Theory of error-correcting pairs

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 $(A \ast B) \perp C$
E.2 $k(A) > t$
E.3 $d(B^\perp) > t$
E.4 $d(A) + d(C) > n$

Let $(A, B)$ be linear subcodes of $\mathbb{F}_{q^m}^n$ that satisfy E.1, E.2, E.3 and

E.5 $d(A^\perp) > 1$
E.6 $d(A) + 2t > n$

Then $d(C) \geq 2t + 1$ and $(A, B)$ is a $t$-ECP for $C$
Kernel of a received word

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

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

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 $A \ast B \subseteq C^\perp$, then

$$K(r) = K(e)$$
Kernel for a GRS code

Let $A = GRS_{t+1}(a, 1)$ and $B = GRS_t(a, 1)$ and $C = \langle A \ast B \rangle^\perp$

Let

$a_i = ev_{a, 1}(X^{i-1})$ for $i = 1, \ldots, t + 1$
$b_j = ev_{a, 1}(X^j)$ for $j = 1, \ldots, t$
$h_l = ev_{a, 1}(X^l)$ for $l = 1, \ldots, 2t$

Then

$a_1, \ldots, a_{t+1}$ is a basis of $A$
$b_1, \ldots, b_t$ is a basis of $B$
$h_1, \ldots, h_{2t}$ is a basis of $C^\perp$

Furthermore

$a_i \ast b_j = ev_{a, 1}(X^{i+j-1}) = h_{i+j-1}$
Matrix of syndromes for a GRS code

Let \( r \) be a received word and \( s = rH^T \) its syndrome. Then

\[(b_j \ast a_i) \cdot r = s_{i+j-1}.\]

To compute the kernel \( K(r) \) we have to compute the null space of the matrix of syndromes

\[
\begin{pmatrix}
s_1 & s_2 & \cdots & s_t & s_{t+1} \\
s_2 & s_3 & \cdots & s_{t+1} & s_{t+2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
s_t & s_{t+1} & \cdots & s_{2t-1} & s_{2t}
\end{pmatrix}
\]
Error location

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

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

Lemma
Let \((A \ast B) \perp C\)
Let \(e\) be an error vector of the received word \(r\)
If \(I = \text{supp}(e) = \{ i \mid e_i \neq 0 \}\), then

\[
A(I) \subseteq K(r)
\]

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

Let \((A, B)\) be a \(t\)-ECP for \(C\) with \(d(C) \geq 2t + 1\)
Suppose that \(c \in C\) is the code word sent and \(r = c + e\) is the received word for some error vector \(e\) with \(\text{wt}(e) \leq t\)

The basic algorithm for the code \(C\):
- Compute the kernel \(K(r)\)
  This kernel is nonzero since \(k(A) > t\)
- Take a nonzero element \(a\) of \(K(r)\)
  \(K(r) = K(e)\) since \((A \ast B) \perp C\)
- Determine the set \(J\) of zero positions of \(a\)
  \(\text{supp}(e) \subseteq J\) since \(d(B^\perp) > t\)
  \(|J| < d(C)\) since \(d(A) + d(C) > n\)
- Compute the error values by erasure decoding
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

Let $X$ be an algebraic variety over $\mathbb{F}_q$ with a subset $P$ of $X(\mathbb{F}_q)$ enumerated by $P_1, \ldots, P_n$.

Suppose that we have a vector space $L$ over $\mathbb{F}_q$ of functions on $X$ with values in $\mathbb{F}_q$.

So $f(P_i) \in \mathbb{F}_q$ for all $i$ and $f \in L$.

In this way we have an evaluation map

$$ ev_P : L \longrightarrow \mathbb{F}_q^n $$

defined by $ev_P(f) = (f(P_1), \ldots, f(P_n))$.

This evaluation map is linear, so its image is a linear code.
The classical example: 
**Generalized Reed-Solomon codes**

The geometric object $X$ is the affine line over $\mathbb{F}_q$
The points are $n$ distinct elements of $\mathbb{F}_q$
$L$ is the vector space of polynomials of degree at most $k - 1$
and with coefficients in $\mathbb{F}_q$

This vector space has dimension $k$
Such polynomials have at most $k - 1$ zeros
so nonzero codewords have at least $n - k + 1$ nonzeros

This code has parameters $[n, k, n - k + 1]$ if $k \leq n$
Let $\mathcal{X}$ be an algebraic curve over $\mathbb{F}_q$ of genus $g$

$\mathbb{F}_q(\mathcal{X})$ is the function field of the curve $\mathcal{X}$ with field of constants $\mathbb{F}_q$

Let $f$ be a nonzero rational function on the curve

The divisor of zeros and poles of $f$ is denoted by $(f)$

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

Then

$$L(E) = \{ f \in \mathbb{F}_q(\mathcal{X}) \mid f = 0 \text{ or } (f) \geq -E \}$$

The dimension of the space $L(E)$ is denoted by $l(E)$

Then $l(E) \geq m + 1 - g$ and equality holds if $m > 2g - 2$

by the Theorem of Riemann-Roch
Codes on curves

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

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

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

The algebraic geometry code $C_L(\mathcal{X}, \mathcal{P}, E)$ is the image of $L(E)$ under the evaluation map $\text{ev}_{\mathcal{P}}$

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 $C_L(\mathcal{X}, \mathcal{P}, E)$
Information rate

Information rate $R = k / n$
Relative minimum distance $\delta = d / n$
Singleton $R + \delta \leq 1$
Gilbert-Varshamov $R \geq 1 - H_q(\delta)$
q-ary entropy function $H_q$
Goppa for AG codes $R + \delta \geq 1 - \gamma$
Relative genus $\gamma = g / n$
Ihara-Tsfasman-Vladut-Zink $\gamma = \frac{1}{\sqrt{q-1}}$
Figuur: Bounds on $R$ as a function of $\delta$ for $q = 49$ and $\gamma = \frac{1}{6}$. 
Let $\omega$ be a differential form with a simple pole at $P_j$ with residue 1 for all $j = 1, \ldots, n$

Let $K$ be the canonical divisor of $\omega$
Let $m$ be the degree of the divisor $E$ on $X$
with disjoint support from $P$

Let $E^\perp = D - E + K$ and $m^\perp = \deg(E^\perp)$
Then $m^\perp = 2g - 2 - m + n$ and

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

$m - 2g + 2$ is called the designed minimum distance of $C_L(X, P, E)^\perp$
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

$$C_L(\mathcal{X}, \mathcal{P}, F) \ast C_L(\mathcal{X}, \mathcal{P}, G) \subseteq C_L(\mathcal{X}, \mathcal{P}, F + G)$$
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
- $A \ast B \subseteq C^\perp$
- If $t + g \leq \text{deg}(F) < n$, then $k(A) > t$
- If $\text{deg}(G - F) > t + 2g - 2$, then $d(B^\perp) > t$
- If $\text{deg}(G - F) > 2g - 2$, 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

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

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

if

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

By randomnization - Not constructive!
Koblitz:

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

\[ y = f(x) \]

that is easy to evaluate but for which is computationally infeasible (one hopes) to find the inverse

\[ x = f^{-1}(y) \]
PKC systems use **trapdoor one-way functions**

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

RSA, **factoring integers**: given $n = pq$ find $(p, q)$

Diffie-Hellman, **discrete-log problem** in $\mathbb{F}_q$: given $b = a^n$ find $n$

Elliptic curve PKC, **addition on elliptic curve**: given $Q = nP$, find $n$

**Code based PKC systems, decoding of codes**

McEliece (Goppa codes)

Niederreiter with parity check matrix instead of generator matrix

Janwa-Moreno (Algebraic geometry codes)
Decoding up to half the minimum distance

Decoding arbitrary linear codes

Exponential complexity $\approx q^{e(R)n}$

$x$-axis: information rate $R = k/n$

$y$-axis: complexity exponent $e(R)$
**McEliece:**
Let $C$ be a class of codes that have efficient decoding algorithms correcting $t$ errors with $t \leq (d - 1)/2$

**Secret key:** $(S, G, P)$  
$S$ an invertible $k \times k$ matrix  
$G$ a $k \times n$ generator matrix of a code $C$ in $C$.  
$P$ an $n \times n$ permutation matrix

**Public key:** $G' = SGP$

**Message:** $m$ in $\mathbb{F}_q^k$  
**Encryption:** $y = mG' + e$ with random chosen $e$ in $\mathbb{F}_q^n$ of weight $t$  
**Decryption:** $yP^{-1} = mSG + eP^{-1}$ and $eP^{-1}$ has weight $t$  
Decoder gives $c = mSG$ as closest codeword
Code based PKC systems - 2

$G$, $S$ and $P$ are kept secret
$G' = SGP$ is public

The (trapdoor) one-way function of the McEliece public cryptosystem is given by

$$x = (m, e) \mapsto y = mG' + e$$

where $m \in \mathbb{F}_q^k$ is the plaintext
$e \in \mathbb{F}_q^n$ is a random error vector with hamming weight at most $t$
Let $C_{ECP}$ be the set of pairs $(A, B)$ that satisfy $E.2$, $E.3$, $E.5$ and $E.6$

The McEliece cryptosystem on codes $C \subseteq (A \ast B) \perp$ with $(A, B)$ in $C_{ECP}$ is based on the inherent tractability of finding an inverse on the one-way function

$$x = (A, B) \mapsto y = (A \ast B)$$

where $(A, B)$ is in $C_{ECP}$
State of the art

- GRS codes: solved by Sidelnikov-Shestakov
- Alternant codes: open
- Goppa codes: open
- AG codeds: work in progress by

Irene Márquez-Corbella
Edgar Martínez-Moro
Ruud Pellikaan
Diego Ruano