A differential equation for decoding derivative codes a.k.a Hermite interpolation with outliers

Daniel Augot
INRIA
Joint work with Alain Couvreur, Emmanuel Hallouin, Thierry Hénocq and Marc Perret

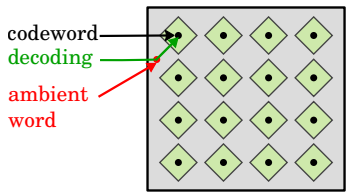
FELIM 2024

Coding Theory, Redundancy and Decoding

- A correcting code C encodes a message of length k into a codeword of length $n > k$

$$m \mapsto c \in C \subset \mathbb{F}_q^n$$

- Some noise modifies the codeword
- Decoding: given an "erroneous codeword", recover the message



- Message should be close (Hamming distance)
- d is the minimum distance
- less than $(d-1)/2$ errors $\implies$ unique decoding

combinatorics A good code has many codewords and a large minimum distance

algorithmics And an associated decoding algorithm

Difficult Tasks

- Difficult to build good codes
- Decoding is intractable

Reed-Solomon codes

Definition

Given $L = \{x_1, \ldots, x_n\} \subset \mathbb{F}_q$, and $k \leq n$
the Reed-Solomon codes RS$[L, k]$ is the code

$$\{(f(x_1), \ldots, f(x_n)) | \deg f(X) < k\} \subset \mathbb{F}_q^n$$

It has minimum distance $d = n - k + 1$ (*Singleton bound*: best distance achievable)

- $k = 1$ (constants) $\frac{n}{2}$ errors: majority vote
- $k = 2$ (lines) $\frac{n-1}{2}$ errors.
- $\frac{d-1}{2} = \frac{n-k}{2}$ errors can be corrected with a unique solution
- $f(X)$ coincides in $t = \frac{n+k}{2}$ positions

Lagrange interpolation with outliers

Problem

Given $x_1, \ldots, x_n$, $y = (y_1, \ldots, y_n) \in \mathbb{F}_q^n$, k and $t \leq n$, find

$$\mathcal{L}_y^t = \{f(X) \in \mathbb{F}[X], \deg f(X) < k, |C(f(X), y)| \geq t\}$$

- $t = n$: unique solution
- $t \geq \frac{n+k}{2}$: unique solution
- $t = k$: $\binom{n}{k}$ solutions
- $t < k$: problem is ill-founded
- $k < t < \frac{n+k}{2}$: "list-decoding"

Unique decoding: Berlekamp-Welch

Code: $\mathbb{F}$, $k \leq n$, $x_1, \ldots, x_n \in \mathbb{F}$, $x_i \neq x_j$

Decoding Radius: $t = \frac{n+k}{2}$

"Received word": $y = (y_1, \ldots, y_n) \in \mathbb{F}^n$

Bivariate Interpolation

Find $A(X, Y) = A_0(X) + A_1(X)Y$ such that

- $A(x_i, y_i) = 0$, $i = 1, \ldots, n$
- $\deg A_0(X) < t$
- $\deg A_1(X) < t - (k-1)$

Root-Finding

return $f(X) = A_1(X)/A_0(X)$ if

- $f(X)$ is a polynomial of degree $< k$
- $C(f(X), y) \geq t$

$A(X, Y) = A_0(X) + A_1(X)Y$ is such that

- $A(x_i, y_i) = 0,\ i = 1, \ldots, n$
- $\deg A_0(X) < t$
- $\deg A_1(X) < t - (k-1)$

Correctness

Let $f(X)$ with $C(f(X), y) \geq t$ and $\deg f(X) < k$

1. $\deg A_0(X) + A_1(X)f(X) < t$
2. $f(x_i) = y_i \implies A(x_i, f(x_i)) = 0$
3. $C(f(X), y) \geq t \implies A(X, f(X))$ has at least t roots $\implies A(X, f(X)) = 0$

Non triviality

1. linear system with n equations
2. number of unknowns $t + t - (k-1) = 2t + 1 - k \geq 2\frac{n+k}{2} + 1 - k = n + 1$
3. $\implies$ there exists a non zero $A(X, Y)$

Sudan

- Look for
$$A(X, Y) = A_0(X) + A_1(X)Y + \cdots + A_\ell(X)Y^\ell$$
with $\mathrm{wdeg}_{1,k-1} A(X, Y) < t$ such that
$$A(x_i, y_i) = 0, \quad i = 1, \ldots, n$$
- Finding $A(X, Y)$ is solving a homogoneous linear system
- Solve for $f(X)$ solution to $A(X, f(X)) = 0$

The existence of a non zero $A(X, Y)$ is ensured by

$$\text{number of unknowns} > \text{number of equations}$$

$$\implies t \geq \sqrt{2n(k-1)},\ |\mathcal{L}| \leq \sqrt{2/R}$$

Guruswami-Sudan

- Look for
$$A(X, Y) = A_0(X) + A_1(X)Y + \cdots + A_\ell(X)Y^\ell$$
with $\mathrm{wdeg}_{1,k-1} A(X, Y) < t$ such that
$$A(x_i, y_i) = 0, \quad i = 1, \ldots, n$$
with multiplicity s
- Solve for $f(X)$ solution to $A(X, f(X)) = 0$

The existence of a non zero $A(X, Y)$ is ensured by

$$\text{number of unknowns} > \text{number of equations}$$

$$\implies t \geq \sqrt{n(k-1)(1+\tfrac{1}{s})} \text{ and } |\mathcal{L}| \leq \sqrt{ns(s+1)/k}$$

"Capacity" of list decoding for general codes

With the notation $\tau = (n - t)/n$, and ℓ the maximum list size

Existence result (ELias 1957)

There exists a family of (τ, ℓ) list-decodable codes over $\mathbb{F}_q$ of rate

$$R \geq 1 - H_q(\tau) - 1/\ell$$

with

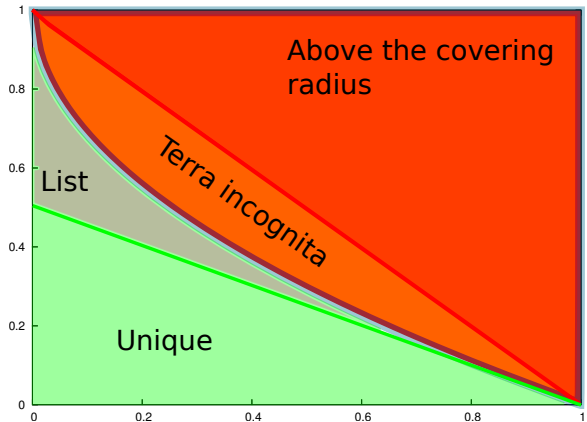
$$H_q(\tau) = -\tau \log_q(\tau) - (1 - \tau) \log_q(1 - \tau) + \tau \log_q(q - 1)$$

Using $\ell = O(\frac{1}{\varepsilon})$, and $q = \exp(O(1/\varepsilon))$ gives

$$R \geq 1 - \tau - \varepsilon$$

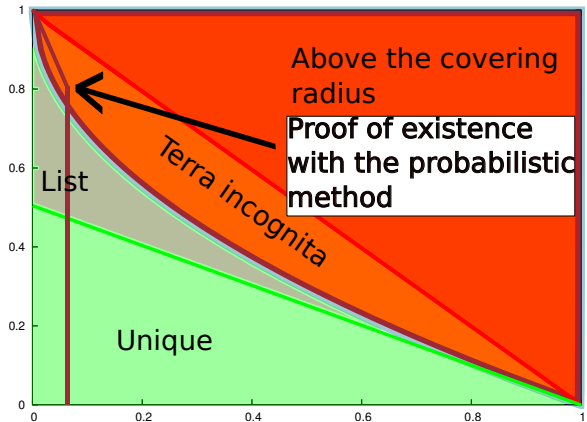
Same result for linear codes (Mosheiff, Resch, Ron-Zewi, Silas,Wootters 2019)

Case of Reed-Solomon codes



- Decoding of Reed-Solomon is NP-hard (Guruswami-Vardy 2004)
- Existence of Reed-Solomon codes with better list decoding radius (Goldberg-Shangguan-Tamo 2021, Ferber-Kwan-Sauermann 2020, Wooters 2014)

Case of Reed-Solomon codes



- Decoding of Reed-Solomon is NP-hard (Guruswami-Vardy 2004)
- Existence of Reed-Solomon codes with better list decoding radius (Goldberg-Shangguan-Tamo 2021, Ferber-Kwan-Sauermann 2020, Wooters 2014)

Derivative codes (Guruswami-Wang 2011)

For $m \leq k \leq nm$, the derivative code $\mathrm{DRS}^m[n,k]$ has "codewords" given by

$$\begin{pmatrix} f(x_1) & \dots & f(x_n) \\ f^{(1)}(x_1) & \dots & f^{(1)}(x_n) \\ \vdots & & \vdots \\ f^{(m-1)}(x_1) & \dots & f^{(m-1)}(x_n) \end{pmatrix}$$

where $f(X)$ is a polynomial of degree less than k

Beware!

- "messages" have alphabet $\mathbb{F}_q$
- "codewords" have alphaet $\mathbb{F}_q^m$
- The rate is $R = (k-1)/(nm)$
- the code

$$C \subset \left(\mathbb{F}_q^m\right)^n$$

is $\mathbb{F}_q$-linear

Hermite interpolation with errors

Problem

Given $x_1, \ldots, x_n$, k and $t \leq n$,

$$y = \begin{pmatrix} x_1 & \ldots & x_n \\ y_{01} & \ldots & y_{0n} \\ y_{11} & \ldots & y_{1n} \\ \vdots & & \vdots \\ y_{m-1,1} & \ldots & y_{m-1,n} \end{pmatrix}$$

find

$$\mathcal{L}_y^t = \{f(X) \in \mathbb{F}[X], \deg f(X) < k, |C(f(X), y)| \geq t\}$$

where $C(f(X), y)$ is the set of agreeing columns

Linear-algebraic list decoding, first version

Given a received word y, find

$$A(X, Y_1, \ldots, Y_{m-1}) = B(X) + A_0(X)Y_0 + \cdots + A_{m-1}(X)Y_{m-1}$$

such that

1. $\deg B(X) < t$
2. $\deg A_i(X) < t - (k-1)$
3. for $i = 1, \ldots, n$

$$A(x_i, y_{0,i}, \ldots, y_{m-1,i}) = 0$$

(Correctness) For any $f \in \mathcal{L}$

$$A(X, f(X), \ldots, f^{(m-1)}(X)) = 0$$

Analysis

(Decoding radius) The "more unknowns than equations" analysis gives

$$\begin{aligned} t &> \frac{n}{m+1} + \frac{m}{m+1} \cdot (k-1) \\ \frac{t}{n} &> \frac{1}{m+1} + \frac{m(k-1)}{n(m+1)} \\ &\sim \frac{1}{m+1} + mR \end{aligned}$$

With $\tau = 1 - t/n$:

$$\tau \sim 1 - \varepsilon - \varepsilon^{-1}R$$

We aim at

$$\tau \sim 1 - \varepsilon - R$$

Linear-algebraic list decoding, improved version

Derivation

$$D(p(X)Y_i) = p'(X)Y_i + p(X)Y_{i+1},$$

Given a received word y, find **for an auxiliary** s

$$A(X, Y_1, \ldots, Y_s) = B(X) + A_0(X)Y_0 + \cdots + A_{s-1}(X)Y_{s-1}$$

such that

1. $\deg B(X) < (m - s + 1)t$
2. $\deg A_i(X) < (m - s + 1)t - (k - 1)$
3. for $i = 1, \ldots, n$

$$\left(D^{(j)}A\right)(x_i, y_{0,i}, \ldots, y_{m-1,i}) = 0$$

for $j = 0, \ldots, m - s$

Reaching capacity

(Decoding radius) The "more unknowns than equations" analysis gives

$$t > \frac{n}{s+1} + \frac{s}{s+1} \cdot \frac{k-1}{m-s+1}$$

Optimisation

- $s = \Omega(\varepsilon^{-1})$
- $m = \Omega(\varepsilon^{-2})$

$$\frac{t}{n} \geq R + \varepsilon$$

Finite dimensional space

We are left with solving **a linear differential equation with polynomial coefficients**

$$B(X) + A_0(X)f(X) + \cdots + A_{s-1}(X)f^{(s-1)}(X) = 0 \quad (*)$$

- **looking for polynomial solutions**
- **of given bounded degree $k-1$**
- **which have t initial values in the received word y**

Case of a finite field

Proposition

If $A_{s-1}(X) \neq 0$, *all solution to* $(*)$ *belong to an affine space of dimension* $\leq s-1$

Exhaustive search when $\mathbb{F} = \mathbb{F}_q$ is finite $\implies$ list-size at most $q^s = q^{1/\varepsilon}$

Problem (reformulated)

Given

$$y = \begin{pmatrix} x_1 & \dots & x_n \\ y_{01} & \cdots & y_{0n} \\ y_{11} & \cdots & y_{1n} \\ \vdots & & \vdots \\ y_{m-1,1} & \cdots & y_{m-1,n} \end{pmatrix}$$

and

$$B(X) + A_0(X)f(X) + \cdots + A_{s-1}(X)f^{(s-1)}(X) = 0 \quad (*)$$

Find all solutions $f(X)$ which coincide with y in a at least t positions

Issue

$f(X)$ may be such that $A_{s-1}(x_i) = 0$ for $i \in C(f(X), y)$

Case $s = m$

$$\deg A_{s-1}(X) < (m - s + 1)t - (k - 1) = t - (k - 1) < t$$

Proposition

If $A_{s-1}(X) \neq 0$, for any $f(X) \in \mathcal{L}$, there exists $i \in [n]$ such that

$$i \in C(f(X), y) \text{ and } A_{s-1}(x_i) \neq 0$$

Proof. Suppose there is an $f(X) \in \mathcal{L}$ such that this is not true.

$i \in C(f(X), y) \implies (X - x_i) | A_{s-1}(X)$

$$\prod_{i \in C(f(X), y)} (X - x_i) \mid A_{s-1}(X)$$

But $|C(f(X), y)| \geq t$

List size

Over any field, when $s = m$

- Any $f(X) \in \mathcal{L}$ has a correct index i such that $A_{s-1}(x_i) \neq 0$
- Any $f(X)$ can be found from its Taylor series at this position

$\implies |\boldsymbol{\mathcal{L}}| \leq n$

More generally for $s \neq m$

When $A_{s-1}(X)$ has less than t zeros in $\{x_1, \ldots, x_n\}$

- any $f(X) \in \mathcal{L}$ has a correct index i such that $A_{s-1}(x_i) \neq 0$

$\implies |\boldsymbol{\mathcal{L}}| \leq \boldsymbol{n}$

Case $s = 2$, m general

First order linear equation

$$B(X) + A_0(X)f(X) + A_1(X)f'(X) = 0$$

w.l.o.g $\gcd(A_1(X), A_0(X)) = 1$

Let

$$Z(X) = \prod_{i \in I}(X - x_i)$$

- If $\deg Z(X) < t$
  Any $f(X) \in \mathcal{L}$ is determined by Taylor expansion at some point $x_i \implies |\mathcal{L}| \leq n$
- If $\deg Z(X) \geq k$, then
  $$B(X) + A_0(X)f(X) = \text{mod } Z(X)$$
  which uniquely determines $f(X) \implies |\mathcal{L}| \leq 1$
- $t \leq \deg Z(X) < k$?

First order approximation of $f(X)$

For x_i such that $Z(x_i) = 0$

$$B(x_i) + A_0(x_i)f(x_i) = 0$$
$$B(x_i) + A_0(x_i)y_{0i} = 0$$

thus

$$A_0(x_i)(y_{0i} - f(x_i)) = 0$$

and $f(x_i) = y_{0i}$ since $A_0(X)$ is coprime to $Z(X)$.

$\implies$ $f(X)$ is known modulo $Z(X)$

Write

$$f(X) = f_0(X) + f_1(X)Z(X)$$

with $f_1(X)$ to be determined

If $Z(x_i) \neq 0$
from $f(X) = f_0(X) + f_1(X)Z(X)$

$$f_1(x_i) = \frac{f(x_i) - f_0(x_i)}{Z(x_i)},$$

If $Z(x_i) = 0$
from $f'(X) = f_0'(X) + f_1'(X)Z(X) + f_1(X)Z'(X)$

$$f_1(x_i) = \frac{f'(x_i) - f_0'(x_i)}{Z'(x_i)}$$

We do not know $f(x_i)$ neither $f'(x_i)$, only y_{i0} and y_{i1}

If $Z(x_i) \neq 0$
from $f(X) = f_0(X) + f_1(X)Z(X)$

$$f_1(x_i) = \frac{f(x_i) - f_0(x_i)}{Z(x_i)},$$

If $Z(x_i) = 0$
from $f'(X) = f_0'(X) + f_1'(X)Z(X) + f_1(X)Z'(X)$

$$f_1(x_i) = \frac{f'(x_i) - f_0'(x_i)}{Z'(x_i)}$$

We do not know $f(x_i)$ neither $f'(x_i)$, only y_{i0} and y_{i1}

Set the first row of $y^{(1)}$ as

$$y_{0i}^{(1)} = \frac{y_{0i} - f_0(x_i)}{Z(x_i)}$$

$$y_{0i}^{(1)} = \frac{y_{1i} - f_0'(x_i)}{Z'(x_i)}$$

If $Z(x_i) \neq 0$
from $f(X) = f_0(X) + f_1(X)Z(X)$

$$f_1(x_i) = \frac{f(x_i) - f_0(x_i)}{Z(x_i)},$$

If $Z(x_i) = 0$
from $f'(X) = f_0'(X) + f_1'(X)Z(X) + f_1(X)Z'(X)$

$$f_1(x_i) = \frac{f'(x_i) - f_0'(x_i)}{Z'(x_i)}$$

We do not know $f(x_i)$ neither $f'(x_i)$, only y_{i0} and y_{i1}

Set the first row of $y^{(1)}$ as

$$y_{0i}^{(1)} = \frac{y_{0i} - f_0(x_i)}{Z(x_i)}$$

$$y_{0i}^{(1)} = \frac{y_{1i} - f_0'(x_i)}{Z'(x_i)}$$

If $Z(x_i) \neq 0$
from $f(X) = f_0(X) + f_1(X)Z(X)$

$$f_1(x_i) = \frac{f(x_i) - f_0(x_i)}{Z(x_i)},$$

If $Z(x_i) = 0$
from $f'(X) = f_0'(X) + f_1'(X)Z(X) + f_1(X)Z'(X)$

$$f_1(x_i) = \frac{f'(x_i) - f_0'(x_i)}{Z'(x_i)}$$

We do not know $f(x_i)$ neither $f'(x_i)$, only y_{i0} and y_{i1}

Set the first row of $y^{(1)}$ as

$$y_{0i}^{(1)} = \frac{y_{0i} - f_0(x_i)}{Z(x_i)}$$

$$y_{0i}^{(1)} = \frac{y_{1i} - f_0'(x_i)}{Z'(x_i)}$$

Other rows of $y^{(1)}$: iterative procedure

Suppose $f^{(u)}(x_i)$ known for $u < j$

Leibniz' rule on $f(X) = f_0(X) + f_1(X)Z(X)$

$$f^{(j)}(x_i) = f_0^{(j)}(x_i) + f_1^{(j)}(x_i)Z(x_i) + \sum_{u=0}^{j-1} \binom{j}{u} f_1^{(u)}(x_i) Z^{(j-u)}(x_i)$$

For $i \notin I$, replacing with received symbols

$$\frac{y_{ji}^{(1)} = y_{ji} - f_0^{(j)}(x_i) - \sum_{u=0}^{j-1} \binom{j}{v} y_{u,i}^{(1)} Z^{(j-u)}(x_i)}{Z(x_i)}$$

Other rows of $y^{(1)}$: iterative procedure

Suppose $f^{(u)}(x_i)$ known for $u < j$

Leibniz' rule on $f(X) = f_0(X) + f_1(X)Z(X)$

$$f^{(j)}(x_i) = f_0^{(j)}(x_i) + f_1^{(j)}(x_i)Z(x_i) + \sum_{u=0}^{j-1} \binom{j}{u} f_1^{(u)}(x_i) Z^{(j-u)}(x_i)$$

For $i \notin I$, replacing with received symbols

$$y_{ji}^{(1)} = \frac{y_{ji} - f_0^{(j)}(x_i) - \sum_{u=0}^{j-1} \binom{j}{v} y_{u,i}^{(1)} Z^{(j-u)}(x_i)}{Z(x_i)}$$

Other rows of $y^{(1)}$: iterative procedure

Suppose $f^{(u)}(x_i)$ known for $u < j$

Leibniz' rule on $f(X) = f_0(X) + f_1(X)Z(X)$

$$f^{(j)}(x_i) = f_0^{(j)}(x_i) + f_1^{(j)}(x_i)Z(x_i) + \sum_{u=0}^{j-1} \binom{j}{u} f_1^{(u)}(x_i) Z^{(j-u)}(x_i)$$

For $i \notin I$, replacing with received symbols

$$y_{ji}^{(1)} = \frac{y_{ji} - f_0^{(j)}(x_i) - \sum_{u=0}^{j-1} \binom{j}{v} y_{u,i}^{(1)} Z^{(j-u)}(x_i)}{Z(x_i)}$$

For $i \in I$, the derivative at order $j+1$

$$f^{(j+1)}(x_i) = f_0^{(j+1)}(x_i) + f_1^{(j+1)}(x_i)Z(x_i) + f_1^{(j)}(x_i)Z'(x_i) + \dots$$

$$y_{ji}^{(1)} = \frac{y_{j+1,i} - f_0^{(j+1)}(x_i) - \dots}{Z'(x_i)}$$

Other rows of $y^{(1)}$: iterative procedure

Suppose $f^{(u)}(x_i)$ known for $u < j$

Leibniz' rule on $f(X) = f_0(X) + f_1(X)Z(X)$

$$f^{(j)}(x_i) = f_0^{(j)}(x_i) + f_1^{(j)}(x_i)Z(x_i) + \sum_{u=0}^{j-1} \binom{j}{u} f_1^{(u)}(x_i) Z^{(j-u)}(x_i)$$

For $i \notin I$, replacing with received symbols

$$y_{ji}^{(1)} = \frac{y_{ji} - f_0^{(j)}(x_i) - \sum_{u=0}^{j-1} \binom{j}{v} y_{u,i}^{(1)} Z^{(j-u)}(x_i)}{Z(x_i)}$$

For $i \in I$, the derivative at order $j+1$

$$f^{(j+1)}(x_i) = f_0^{(j+1)}(x_i) + f_1^{(j+1)}(x_i)Z(x_i) + f_1^{(j)}(x_i)Z'(x_i) + \dots$$

$$y_{ji}^{(1)} = \frac{y_{j+1,i} - f_0^{(j+1)}(x_i) - \dots}{Z'(x_i)}$$

Other rows of $y^{(1)}$: iterative procedure

Suppose $f^{(u)}(x_i)$ known for $u < j$

Leibniz' rule on $f(X) = f_0(X) + f_1(X)Z(X)$

$$f^{(j)}(x_i) = f_0^{(j)}(x_i) + f_1^{(j)}(x_i)Z(x_i) + \sum_{u=0}^{j-1} \binom{j}{u} f_1^{(u)}(x_i) Z^{(j-u)}(x_i)$$

For $i \notin I$, replacing with received symbols

$$y_{ji}^{(1)} = \frac{y_{ji} - f_0^{(j)}(x_i) - \sum_{u=0}^{j-1} \binom{j}{v} y_{u,i}^{(1)} Z^{(j-u)}(x_i)}{Z(x_i)}$$

For $i \in I$, the derivative at order $j+1$

$$f^{(j+1)}(x_i) = f_0^{(j+1)}(x_i) + f_1^{(j+1)}(x_i)Z(x_i) + f_1^{(j)}(x_i)Z'(x_i) + \dots$$

$$y_{ji}^{(1)} = \frac{y_{j+1,i} - f_0^{(j+1)}(x_i) - \dots}{Z'(x_i)}$$

$\rightarrow$ Iterative computation of the rows of $y^{(1)}$

New instance of the problem

A new received word $y^{(1)}$ and a new code $\mathrm{DRS}^{(m-1)}[n, k-t]$

A correct position of $f(X)$ w.r.t y is correct for $f_1(X)$ w.r.t $y^{(1)}$

The new t_1 for $m-1$ and $\deg f_1(X) < k-1-t$ is

$$t_1 = \frac{n}{3} + \frac{2}{3} \cdot \frac{k-1-t}{m-1-s+1} = \frac{n}{3} + \frac{2}{3}\frac{k-1-t}{m-2}$$

Then a miracle!

$$t_1 - t = \frac{2}{3} \cdot \frac{k-1-t(m-1)}{(m-2)(m-1)} \leq \mathbf{0}$$

$\rightarrow$ recursive algorithm with decreasing m
$\rightarrow$ until $m = 2$ $(= s = 2)$

Conclusion

- Hermite interpolation at order m with errors reduces to solving a differential equation
- The differential equation has order s, with auxiliary s
- Initial condition are "given"
- Linear list size for $s = 2$ and $s = m$
- No clue about the truth for other value of s