Best and near-best polynomial approximation I

Alex T.

Let $\mathcal{P}_n$ = space of polynomials of degree $\le n$.

Let $f : [-1,1] \longrightarrow \mathbb{C}$ be a continuous function.

AIM: Replace f by a proxy $p_n \in \mathcal{P}_n$. [Do further computation with p_n as a replacement for f.]

Many different ways to approximation f:

- Best L_∞, $p_n^{L_\infty}$, $\quad p_n^{L_\infty} = \underset{q \in \mathcal{P}_n}{\operatorname{argmin}} \|f-q\|_\infty$, $\quad \|f\|_\infty = \max_{x \in [-1,1]} |f(x)|$.
- Best L_2, $p_n^{L_2}$, $\quad p_n^{L_2} = \underset{q \in \mathcal{P}_n}{\operatorname{argmin}} \|f-q\|_2$, $\quad \|f\|_2^2 = \int_{-1}^{1} |f(x)|^2 dx$.
- Best L_1, $p_n^{L_1}$, $\quad p_n^{L_1} = \underset{q \in \mathcal{P}_n}{\operatorname{argmin}} \|f-q\|_1$, $\quad \|f\|_1 = \int_{-1}^{1} |f(x)|\, dx$.
- ChebT interp, p_n^{chebT}, $\quad p_n^{chebT}(x_i) = f(x_i)$, $0 \le i \le n$, $\quad x_i = \cos\left(\frac{\pi(i+\frac{1}{2})}{n+1}\right)$
- Legendre intep, p_n^{leg}, $\quad p_n^{leg}(x_i) = f(x_i)$, $0 \le i \le n$, $\quad P_{n+1}(x_i) = 0$.
- ChebU interp, p_n^{chebU}, $\quad p_n^{chebU}(x_i) = f(x_i)$, $0 \le i \le n$, $\quad x_i = \cos\left(\frac{i+1}{n+2}\pi\right)$.
- Leja interp., p_n^{leja}, $\quad p_n^{leja}(x_i) = f(x_i)$, $0 \le i \le n$, $\quad x_0, \dots, x_n$ Leja pts
- GE interp., p_n^{GE}, $\quad p_n^{GE}(x_i) = f(x_i)$, $0 \le i \le n$, $\quad x_0, \dots, x_n$ chosen greedily.

First observation: $p_n^{L_1}, p_n^{L_2}, p_n^{L_\infty}$ are similar:

$$\|f-p_n^{L_1}\|_1 \le \|f-p_n^{L_2}\|_1 \le \sqrt{2}\,\|f-p_n^{L_2}\|_2 \le \sqrt{2}\,\|f-p_n^{L_\infty}\|_2 \le 2\|f-p_n^{L_\infty}\|_\infty$$

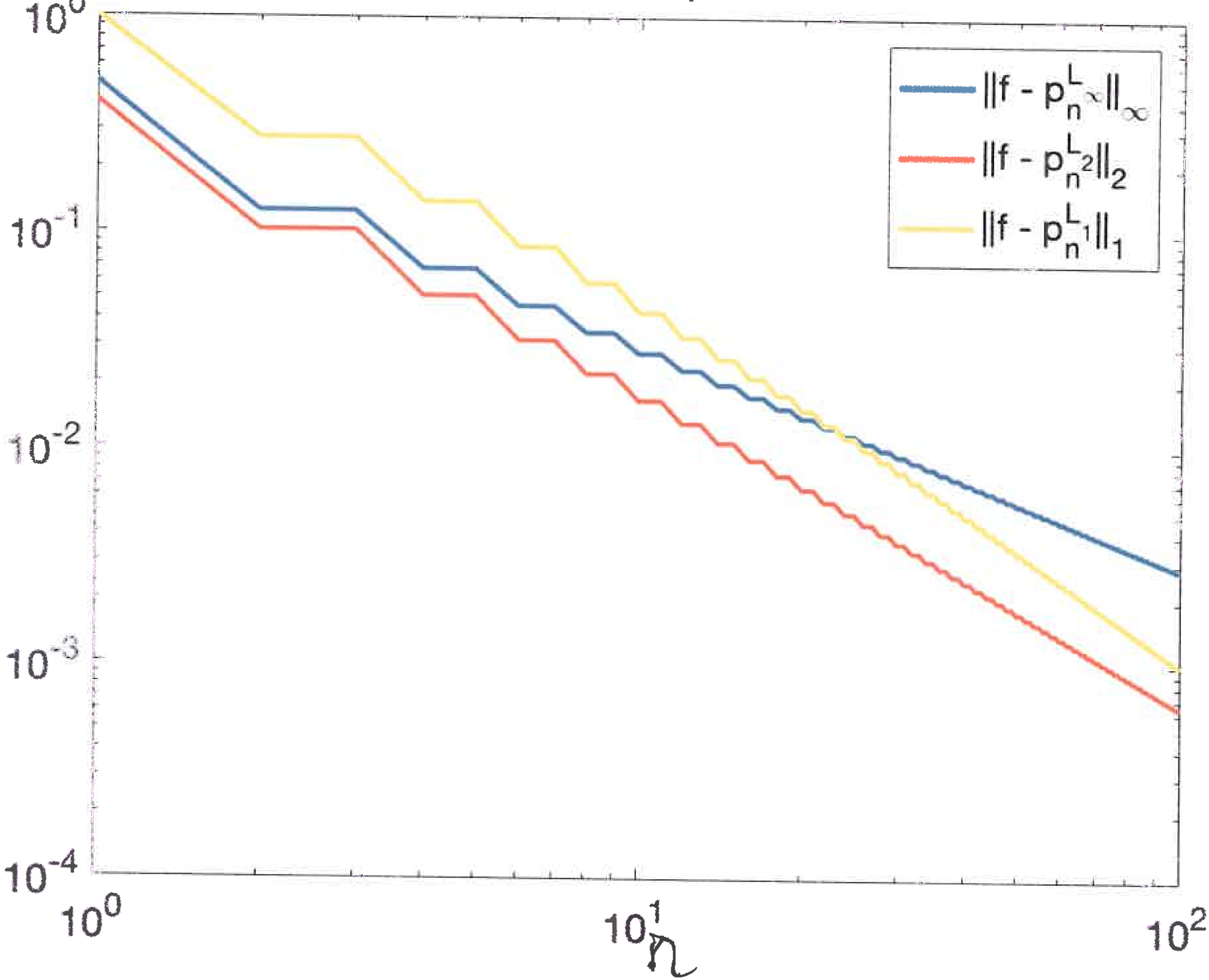
If f = piecewise polynomial of degree $\le n$:

$$\|f-p_n^{L_\infty}\|_\infty \le \|f-p_n^{L_2}\|_\infty \le C_1 n \|f-p_n^{L_2}\|_2 \le C_1 n \|f-p_n^{L_1}\|_2 \le C_2 n^2 \|f-p_n^{L_1}\|_1$$

Nikolskii inequalities.

*

$f = |x|$



10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^0
10^1
10^2
n
$\|f - p_n^{L_\infty}\|_\infty$
$\|f - p_n^{L_2}\|_2$
$\|f - p_n^{L_1}\|_1$

If $\|f - p_n^{L_\infty}\|_\infty$ decays geometrically, then so does $\|f - p_n^{L_t}\|_t$ for $t = 1, 2$.

The main differences are revealed when f is not particularly smooth.

Best L_∞ approximation :

Equioscillation (certificate for optimality). *

$$p_n^{L_\infty} \text{ is best} \iff (f - p_n^{L_\infty})(\xi_i) = \pm(-1)^i \|f - p_n^{L_\infty}\|_\infty$$

for $\xi_0, \ldots, \xi_{n+1} \in [-1,1]$

(e.g. Maximum error is attain $\geq n+2$ in $[-1,1]$.)

Lower bounds on $\|f - p_n^{L_\infty}\|_\infty$: (De La Vallée-Poussin theorem)

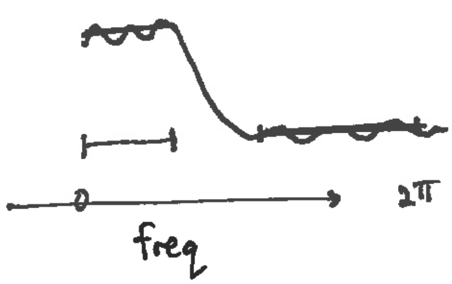
For any $q \in \mathcal{P}_n$, if $(f-q)(\xi_i) = -(f-q)(\xi_{i+1})$ for $\xi_0, \ldots, \xi_{n+1} \in [-1,1]$ *

then $\cancel{\|f - p_n^{L_\infty}\|_\infty}$

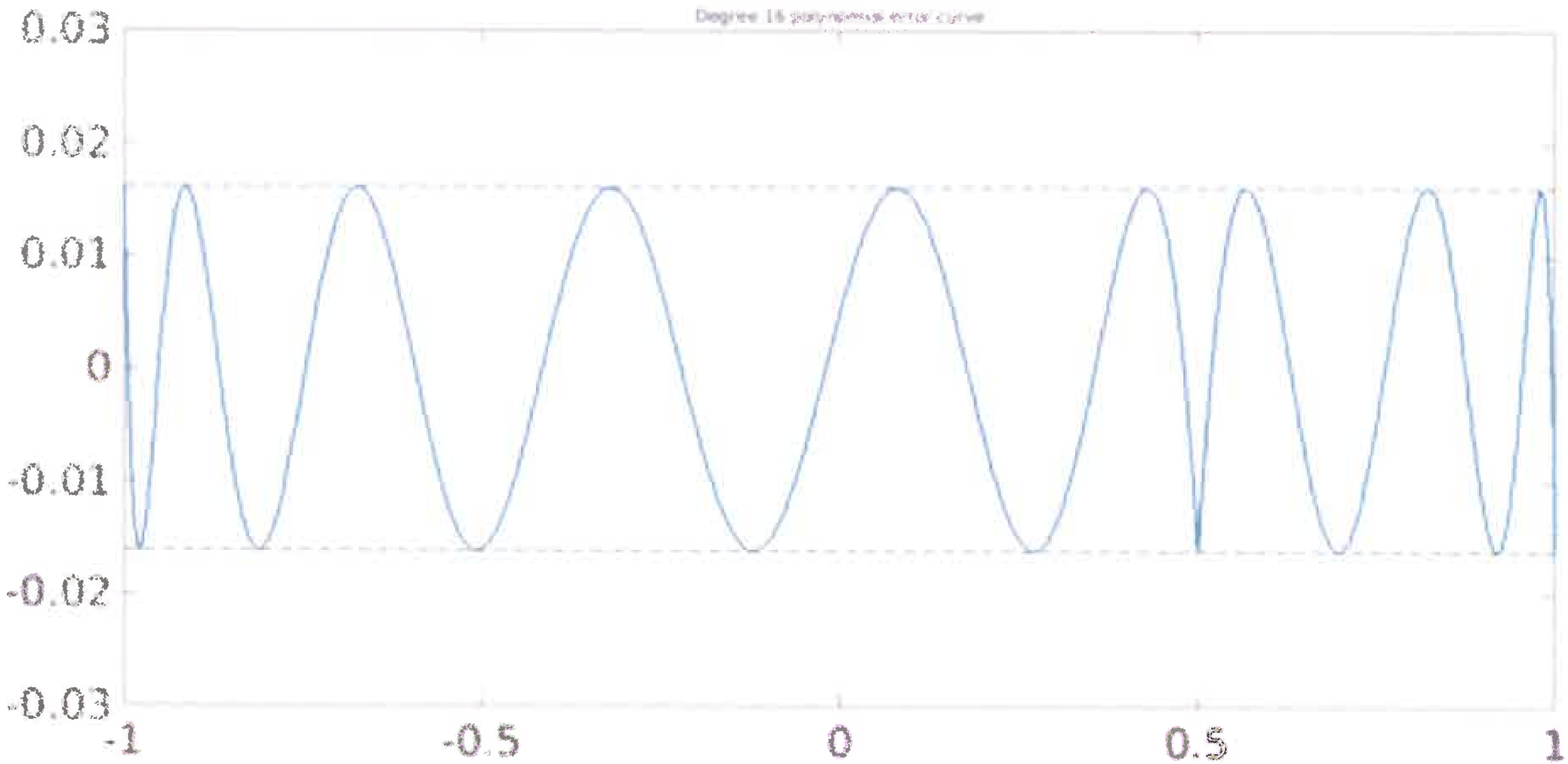
$$\min_{0 \leq i \leq n+1} |f(\xi_i) - q(\xi_i)| \leq \|f - p_n^{L_\infty}\|_\infty \leq \|f - q\|_\infty$$

$\uparrow$ very important lower bound!

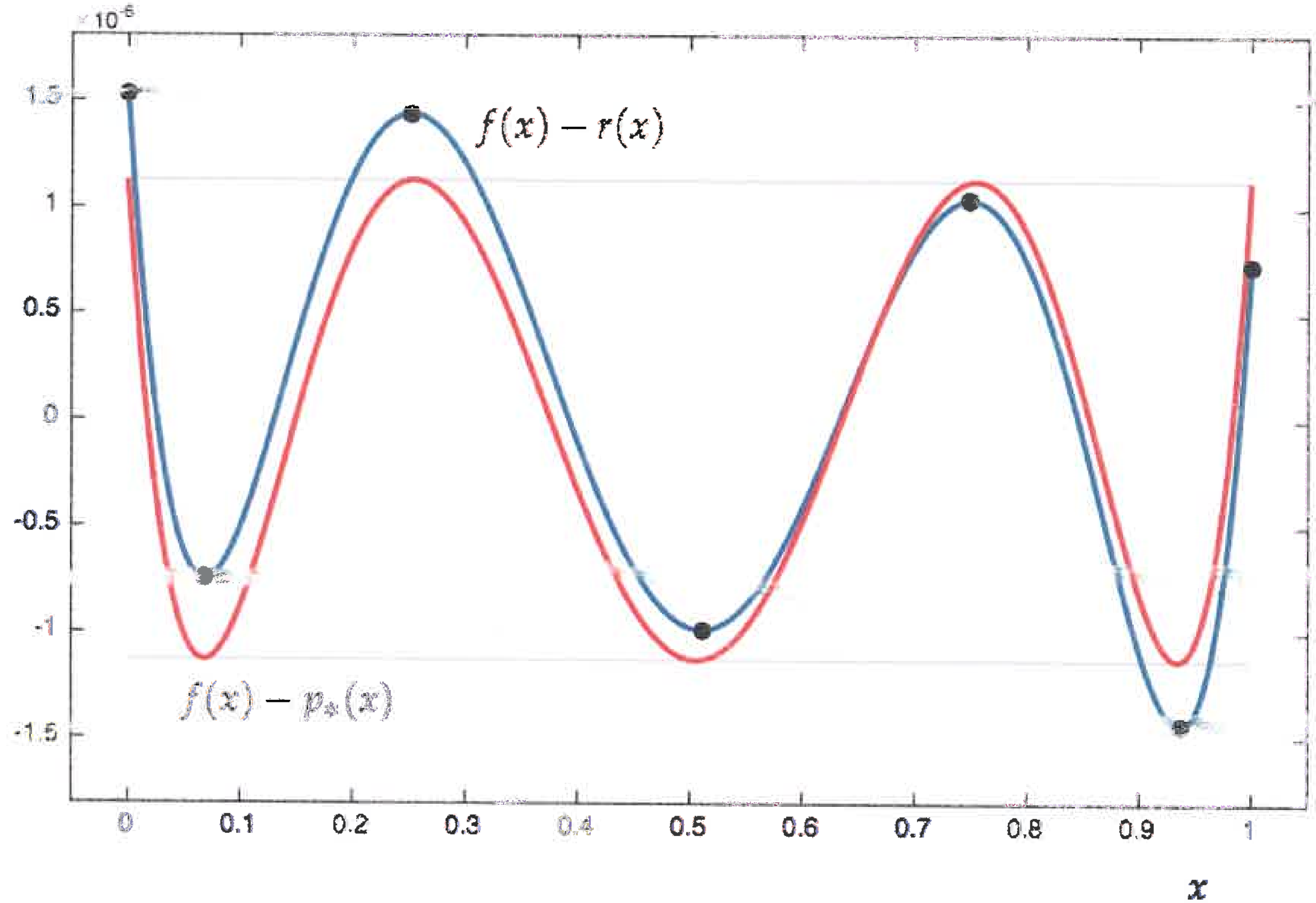
Unfortunately, $p_n^{L_\infty}$ is difficult to compute (done by Remez algorithm) and its applications are restricted to filtering, where "two intervals".



Equioscillation theorem.



De La Vallée - Poussin theorem



<u>Near-best</u>: $x^{n+1} - p_n^{L_\infty}(x) = 2^{-n} T_{n+1}(x)$, T_{n+1} = chebyshev polyonomial.

That means $p_n^{L_\infty}(x) = p_n^{chebT}(x)$ for x^{n+1}. *

<u>In fact</u>: $\|f - p_n^{chebT}\|_\infty \le (1+\sigma_n)\|f - p_n^{\cancel{b}L_\infty}\|_\infty$

$\sigma_n \le 1 + \frac{2}{\pi}\log(n+1)$ [Erdős, 1962]

$$\boxed{\begin{array}{l}\text{The problem: Solve} \\ \qquad p_n^{chebT}(x_i) = f(x_i), \quad x_i = \cos\left(\frac{\pi(i+\frac{1}{2})}{n+1}\right) \\ \qquad \text{where } p_n^{chebT}(x) = \sum_{i=0}^{n} c_i T_i(x)\end{array}}$$

is equivalent to an $(n+1)\times(n+1)$ linear system: $A\underline{c} = \underline{f}$.

The matrix A = DCT-II (up to diagonal scaling)

A^{-1} = DCT-III (up to diagonal scaling).

Therefore, $A\underline{c} = \underline{f}$ can be solved in $O(n\log n)$ operation via ~~a~~ a FFT [Gentleman, 1970].

<u>Best L_2 approximation</u>: $p_n^{L_2} = \underset{p\in P_n}{\operatorname{argmin}} \|f - p\|_2$

Let $Q = [\tilde{P}_0(x) | \dots | \tilde{P}_n(x)]$, where $\int_{-1}^{1} \tilde{P}_i(x)\tilde{P}_j(x)\,dx = \begin{cases}1, & i=j \\ 0, & o/w\end{cases}$.

and degree$(\tilde{P}_i(x)) = i$. Known as ~~a~~ normalized ~~Lob~~ Legendre polynomials.

Then, $p = Q\underline{c}$ for all $p \in P_n$.

Therefore, $p_n^{L_2} = \underset{\underline{c}\in\mathbb{C}^{n+1}}{\operatorname{argmin}} \|f - \overset{\infty\times(n+1)}{Q}\underline{c}\|_2$ This is a least squares problem.

$$\text{Solution:} \quad Q^TQ\underline{c} = Q^T\underline{f} \implies \underbrace{\underline{c} = Q^T f}_{\text{projection of } f \text{ onto } \mathcal{P}_n}$$

$p_n^{L_2}$ is computed by calculating $c_k = \int_{-1}^{1} f(x)\tilde{P}_k(x)\,dx$ for $0 \le k \le n$.

Legendre interpolation: ~~Solve~~ $\cancel{\tilde{\mathbb{P}}_n \underline{c} = \underline{f}}$

$$\text{Solve} \quad p_n^{leg}(x_i) = f(x_i), \quad \tilde{P}_{n+1}(x_i) = 0$$

$$p_n^{leg}(x) = \sum_{k=0}^{n} c_k \tilde{P}_k(x)$$

$$\implies \text{Solve:} \quad \tilde{\mathbb{P}}_n \underline{c} = \underline{f}, \quad (\tilde{\mathbb{P}}_n)_{jk} = \tilde{P}_k(x_j).$$

By Gauss quadrature: $\{x_i, w_i\}$

$$\int_{-1}^{1} \tilde{P}_j(x)\tilde{P}_k(x)\,dx = \sum_{i=0}^{n} w_i \tilde{P}_j(x_i)\tilde{P}_k(x_i) = \begin{cases} 1, & j=k, \\ 0, & j\neq k, \end{cases}$$

$$\therefore \quad \tilde{\mathbb{P}}_n D_w \tilde{\mathbb{P}}_n^T = I$$

$$\therefore \quad \tilde{\mathbb{P}}_n \underline{c} = \underline{f} \implies \underline{c} = D_w \mathbb{P}_n^T \underline{f} \implies c_k = \sum_{i=0}^{n} w_i P_k(x_i) f(x_i)$$

Near-best:

$$\|f - p_n^{leg}\|_2^2 = \|f - p_n^{L_2}\|_2^2 + \|p_n^{L_2} - p_n^{leg}\|_2^2 \qquad \left(\langle f - p_n^{L_2}, q\rangle = 0 \text{ for any } q \in \mathcal{P}_n\right)$$

$$\|p_n^{L_2} - p_n^{leg}\|_2^2 = \sum_{i=0}^{n} w_i\left(p_n^{L_2}(x_i) - f(x_i)\right)^2 \approx \int_{-1}^{1}(f - p_n^{L_2})(x)\,dx.$$

Gauss quadrature error

Since f is continuous:

$$\|f - p_n^{leg}\|_2 \le \left(2 + \frac{c}{n}\right)^{\frac{1}{2}} \|f - p_n^{L_2}\|_2 \qquad \text{[Xiang \& Bornemann 2012]}$$

Since $\underline{c} = D_w \mathbb{P}_n^T \underline{f}$, the polynomial p_n^{leg} can be computed in $O(n\log n)$ operations via a FFT-based transform [Hale & T. 2015].

Noisy function approximation

Suppose one only knows $f : [-1,1] \to \mathbb{R}$ up to a noise level s^2.

That is, $\hat{f}(x_i) = f(x_i) + \xi_i \qquad \xi_i \sim N(0, s^2)$

$$\mathbb{E}\left[\|\hat{\underline{c}} - \underline{c}\|_2^2\right] = \mathbb{E}\left[\|D_w \mathbb{P}_n^T \underline{\xi}\|_2^2\right] \le \|D_w^{\frac{1}{2}}\|_2^2 \, \mathbb{E}\left[\|\underline{\xi}\|_2^2\right]$$

$$\le \frac{\pi}{n+1} \cdot (n+1) s^2 = \pi s^2.$$

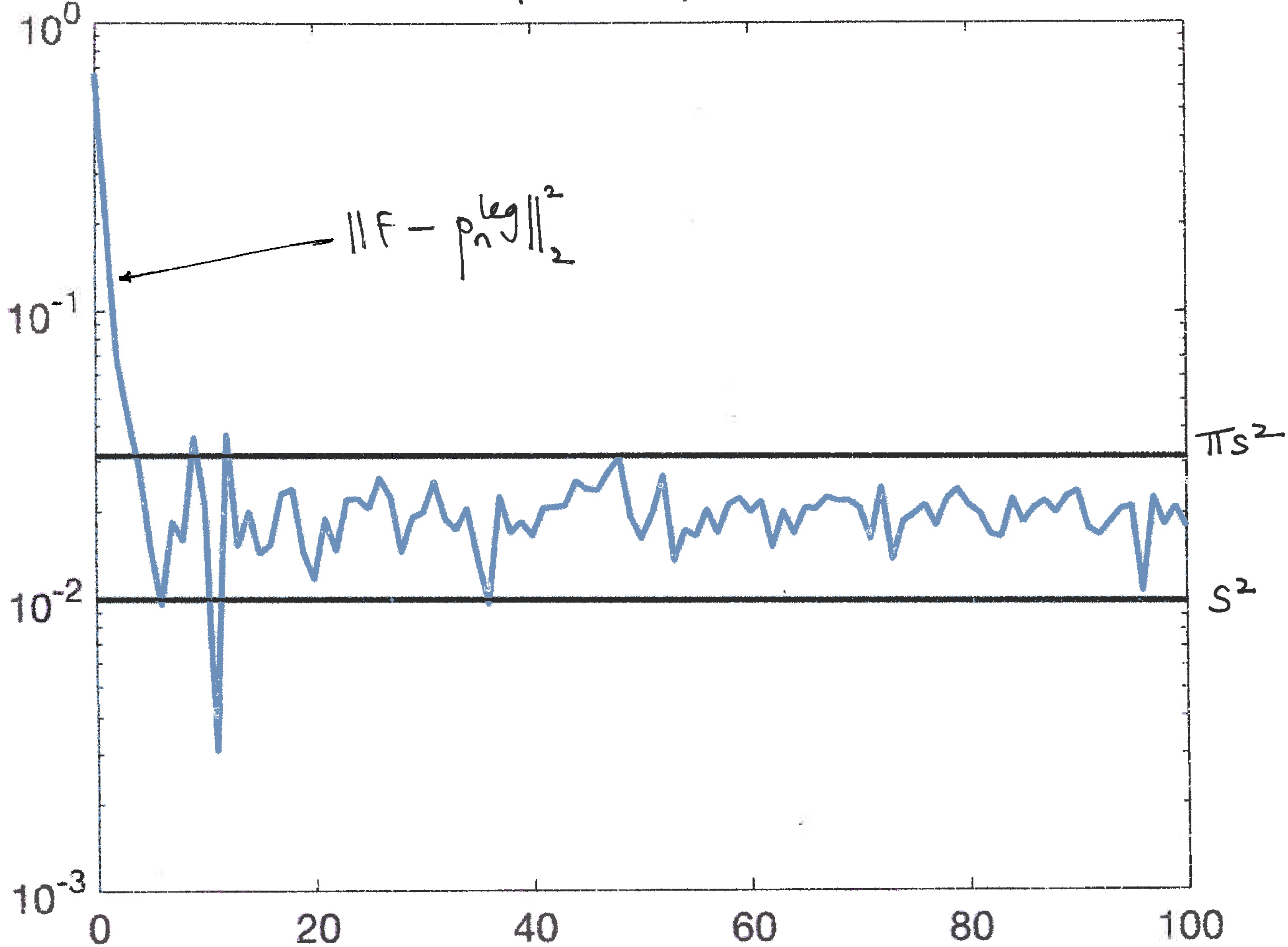
Moreover,

$$\mathbb{E}\left[\|\hat{p}_n^{leg} - p_n^{leg}\|_2^2\right] = \mathbb{E}\left[\|\hat{\underline{c}} - \underline{c}\|_2^2\right] \le \pi s^2.$$

Recovery of noisy f, up to the noise level. This can be performed/computed in $O(n\log n)$ operations.

Conclusion: Legendre Interpolation is ideal for approximating noisy functions.

$f = |x| + \text{noise}$



Best and near-best polynomial approximation II

Best L_1 approximation

(Characterization): If $f(x) - p_n^{L_1}(x)$ has finite many zeros, then

$$p_n^{L_1} \text{ best} \quad \Longleftrightarrow \quad \int_{-1}^{1} \operatorname{sign}\left(f(x) - p_n^{L_1}(x)\right) q(x)\, dx = 0 \text{ for } \forall q \in \mathcal{P}_n$$

While this isn't as visual as equioscillation theorem, it is a remarkable result. For example; it tells us that:

$$\left|\{x \in [-1,1] : f(x) - p_n^{L_1}(x) > 0\}\right| = \left|\{x \in [-1,1] : f(x) - p_n^{L_1}(x) < 0\}\right|$$

<u>Corollary</u>: If $f(x) - p_n^{L_1}(x)$ has exactly $n+1$ zeros (which it does generically / usually) in $[-1,1]$, then

$$p_n^{L_1}(x) = p_n^{\text{chebU}}(x)$$

Here, $p_n^{\text{chebU}}(x)$ is the polynomial interpolant of f at the zeros of $U_{n+1}(x)$. U_{n+1} = the Chebyshev polynomial of the second kind

$$= \frac{\sin((n+2)\cos^{-1}(x))}{\sin(\cos^{-1}(x))}$$

We conclude that $p_n^{\text{chebU}}(x)$ is <u>very</u> often the best L_1 polynomial approximant. It can be computed via an FFT (more precisely a DST) in $O(n \log n)$ operations.

For example: $x^{n+1} - p_n^{L_1}(x) = b_n U_{n+1}(x)$ (b_n to make $b_n U_{n+1}$ monic.)

Proof: $U_{n+1}(x)$ has precisely $n+1$ zeros.

Fiedler and Jurkat (1990) gave a large class of $f(x)$ such that $p_n^{L_1}(x) = p_n^{chebU}(x)$. Including: $|x|$, $\sqrt{1 \mp x^2}$, $\frac{1}{x^2+a^2}$, $\sin^{-1}(x)$, ...

Near-best: I believe that:

$$\|f - p_n^{chebU}\|_1 \le (1 + \tau_n)\|f - p_n^{L_1}\|_1$$

$$\tau_n \le 1 + \frac{2}{\pi}\log(n+2)$$

[A ~~very~~ closely related result in Freilich & Mason 1971.]

Error concentration/localization

Recall that $\|f - p_n^{L_1}\|_1 \le 2\|f - p_n^{L_\infty}\|_\infty \le 2C_2 n^2 \|f - p_n^{L_1}\|_1$ if f = piecewise poly. of degree $\le n$.

Therefore, $\|f - p_n^{L_1}\|_1$ can converge to zero at a faster rate than $\|f - p_n^{L_\infty}\|_\infty$. But, $\|f - p_n^{L_1}\|_\infty \ge \|f - p_n^{L_\infty}\|_\infty$.

If $\|f - p_n^{L_1}\|_1 << \|f - p_n^{L_\infty}\|_\infty$, then the only choice is for $f(x) - p_n^{L_1}(x)$ to have localized error.

If $\Omega_n = \left\{ x \in [-1,1] : f(x) - p_n^{L_1}(x) > \frac{1}{2}\left(f(x) - p_n^{L_\infty}(x)\right) \right\}$

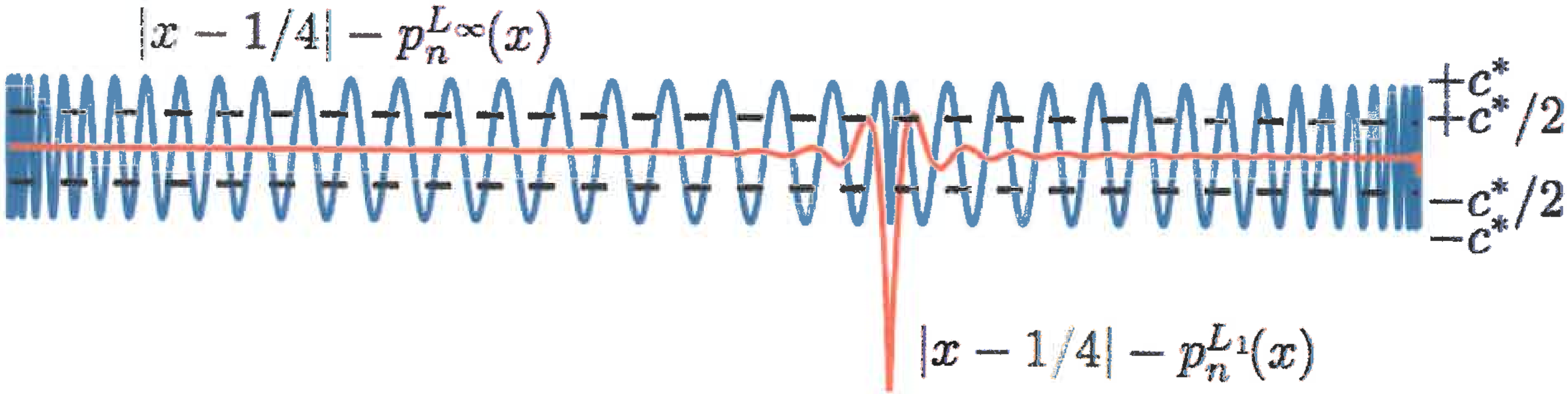
then

$$|\Omega_n| \le \frac{2\|f - p_n^{L_1}\|_1}{\|f - p_n^{L_\infty}\|_\infty}$$

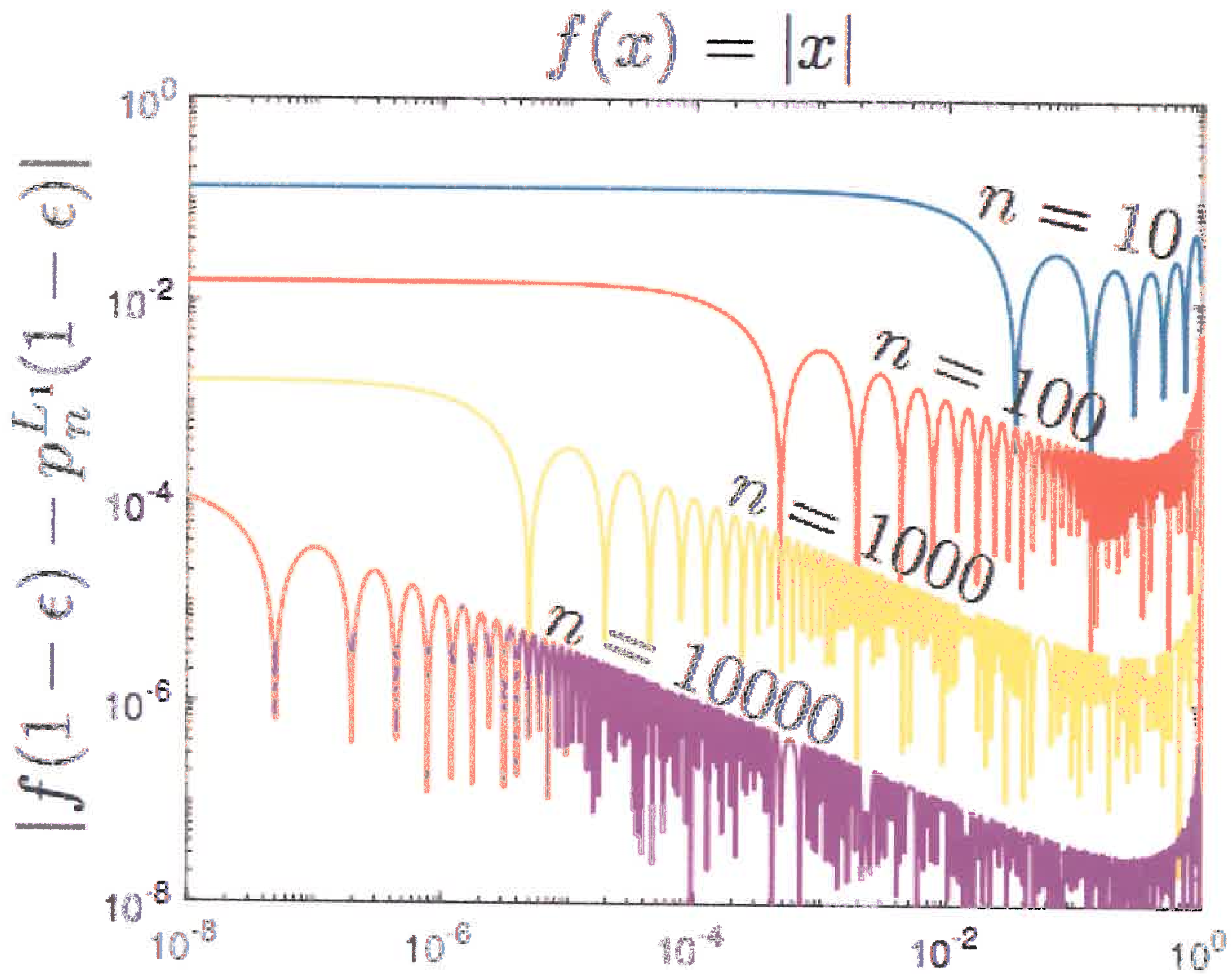
For example, if $f(x) = |x|$:

- $\|f - p_n^{L_1}\|_1 \approx \frac{\pi^2}{4n^2}$
- $\|f - p_n^{L_\infty}\|_\infty \sim \frac{0.28017}{2n}$

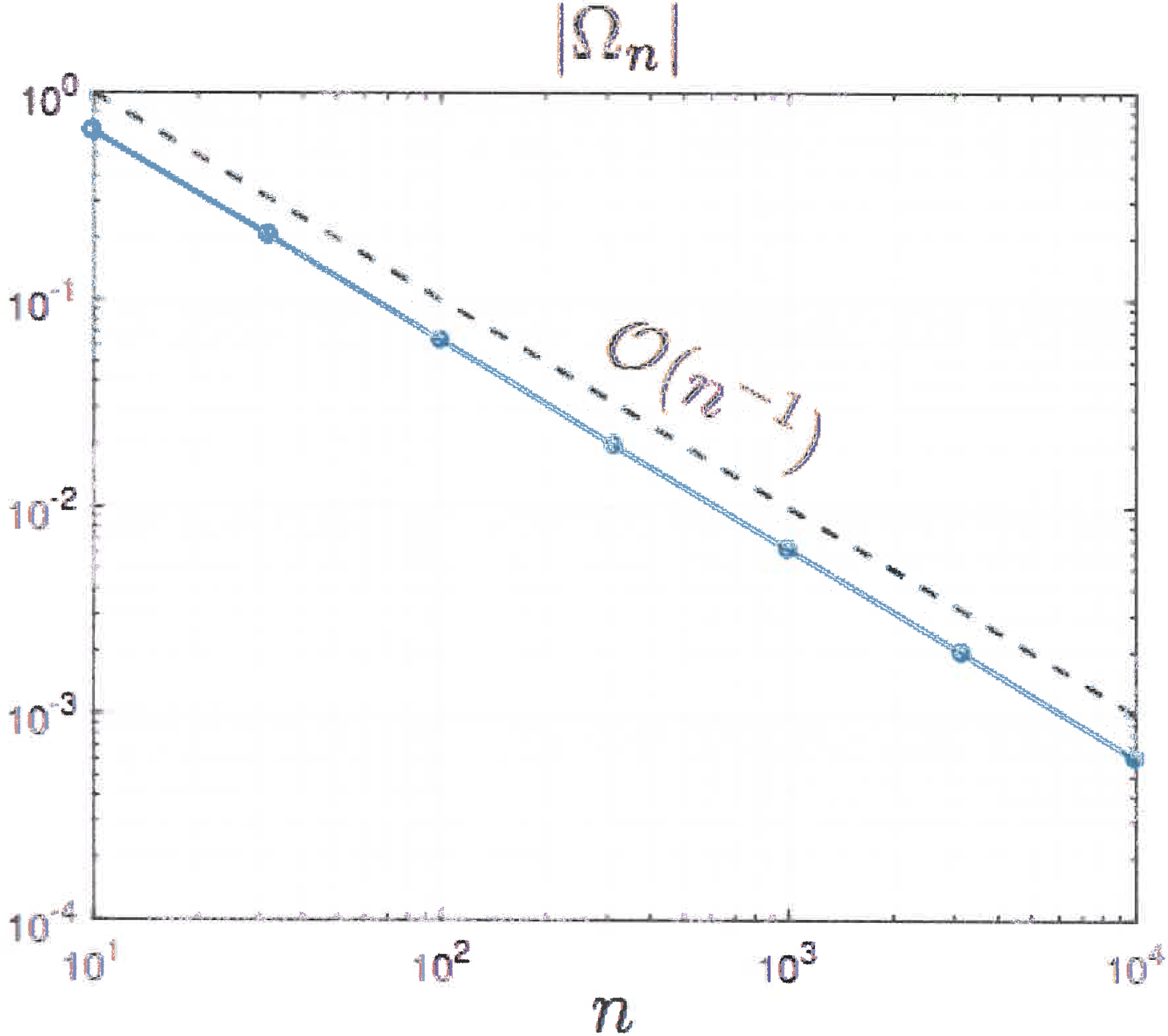
$\Rightarrow |\Omega_n| \lesssim \mathcal{O}\left(\frac{1}{n}\right)$.



|x − 1/4| − p_n^{L∞}(x)
+c*
+c*/2
−c*/2
−c*
|x − 1/4| − p_n^{L1}(x)



$f(x) = |x|$
$|f(1-\epsilon) - p_n^{L_1}(1-\epsilon)|$
ϵ
$n = 10$
$n = 100$
$n = 1000$
$n = 10000$
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}



$|\Omega_n|$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
$\mathcal{O}(n^{-1})$
10^1
10^2
10^3
10^4
n

The error localization properties make $p_n^{L_1}$ ideal for employing on functions that have been corrupted on small intervals.

~~Fejer int~~

<u>GE on functions</u>: If a function is very expensive to evaluate, then $p_n^{L_\infty}$, $p_n^{L_2}$, $p_n^{L_1}$ or their faster variants p_n^{chebT}, p_n^{leg}, p_n^{chebU} may not be suitable:

- Requires lots of function evaluation to compute,
- If, for a fixed n, one has not resolved f sufficiently, then ~~theres is also~~ p_{n+1} costs another $n+1$ function evaluation.

<u>Gaussian elimination:</u>

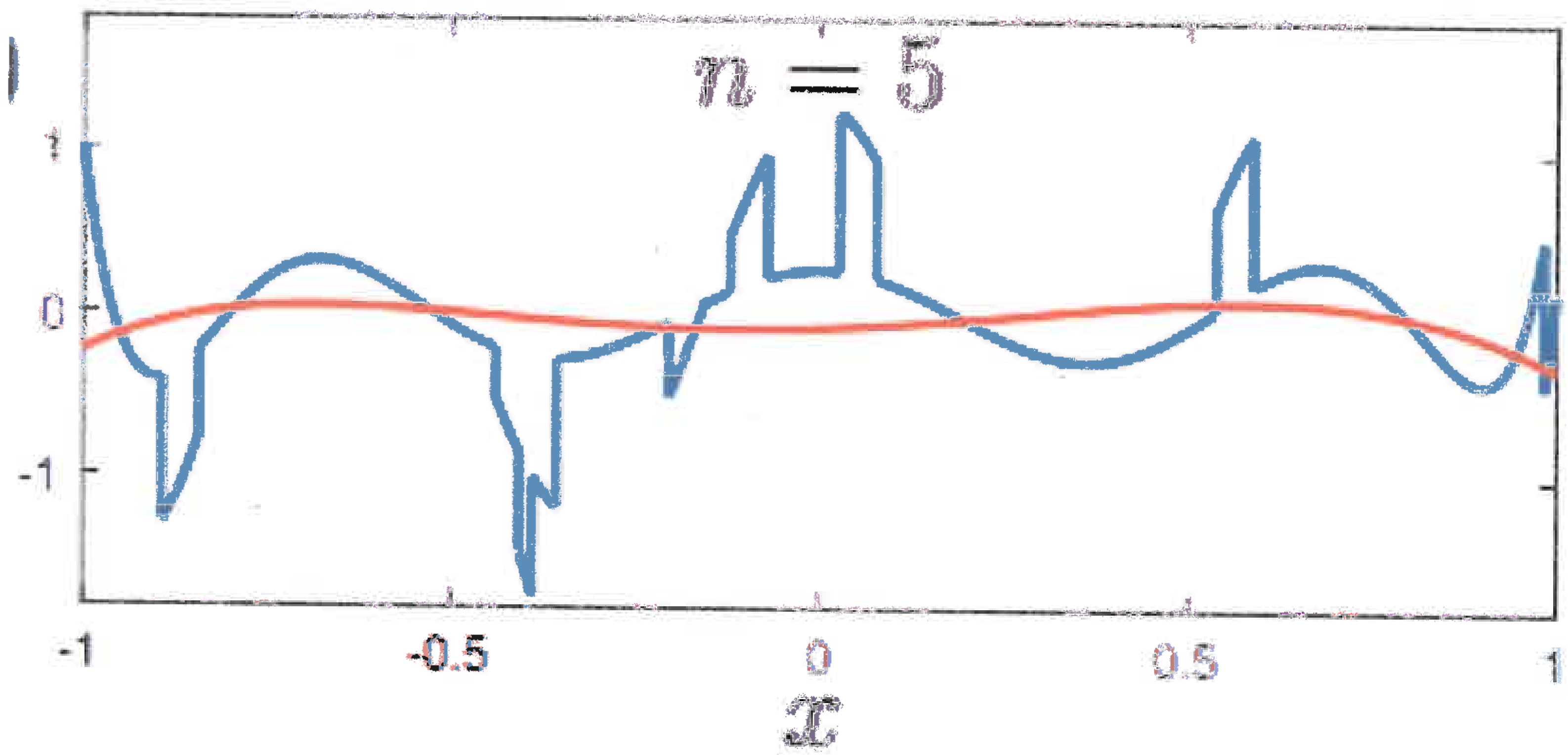
0. Set $p(x) = 0$. Set $\underline{x} = \{\ \}$ empty set.
1. Select $x_* = \operatorname{argmax} |f(x) - p(x)|$
2. ~~Select~~ update $\underline{x} \leftarrow \{\underline{x}, x_*\}$ and set $p(x)$ to be the polynomial interpolant of f at $\underline{x}$.
3. Repeat 1 & 2.

<u>Conjecture / observation</u>: For reasonable functions f, the point set $\underline{x}$ is distributed (in some sense) like the chebyshev points.

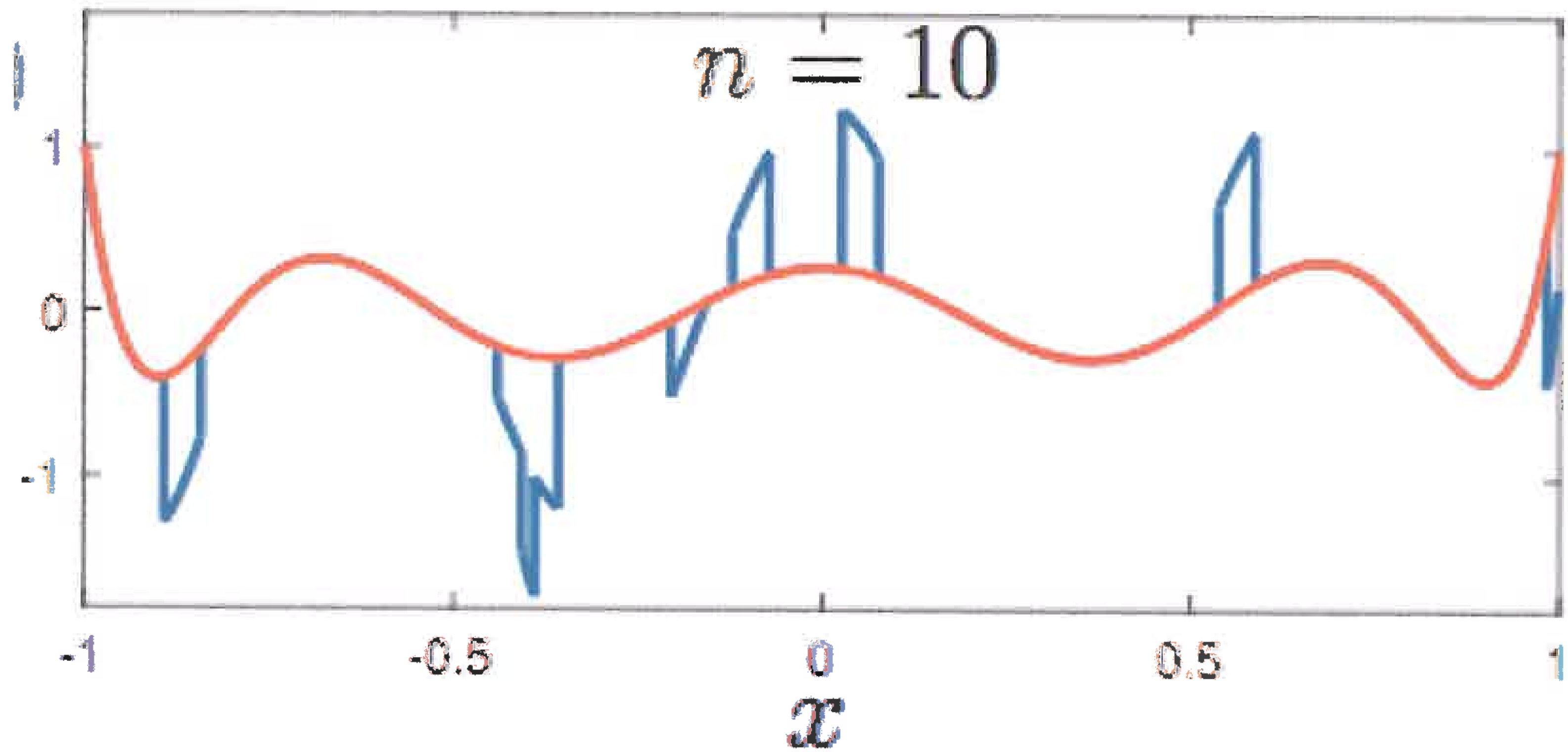
<u>Convergence</u>: All that is known: If $f : [-1,1] \to \mathbb{R}$ is analytic in a Stadium of radius $\beta > 2$ and bounded on its closure, then

$$\| f - p_n^{GE} \|_\infty \leq C \left(\frac{2}{\beta} \right)^{-n}.$$

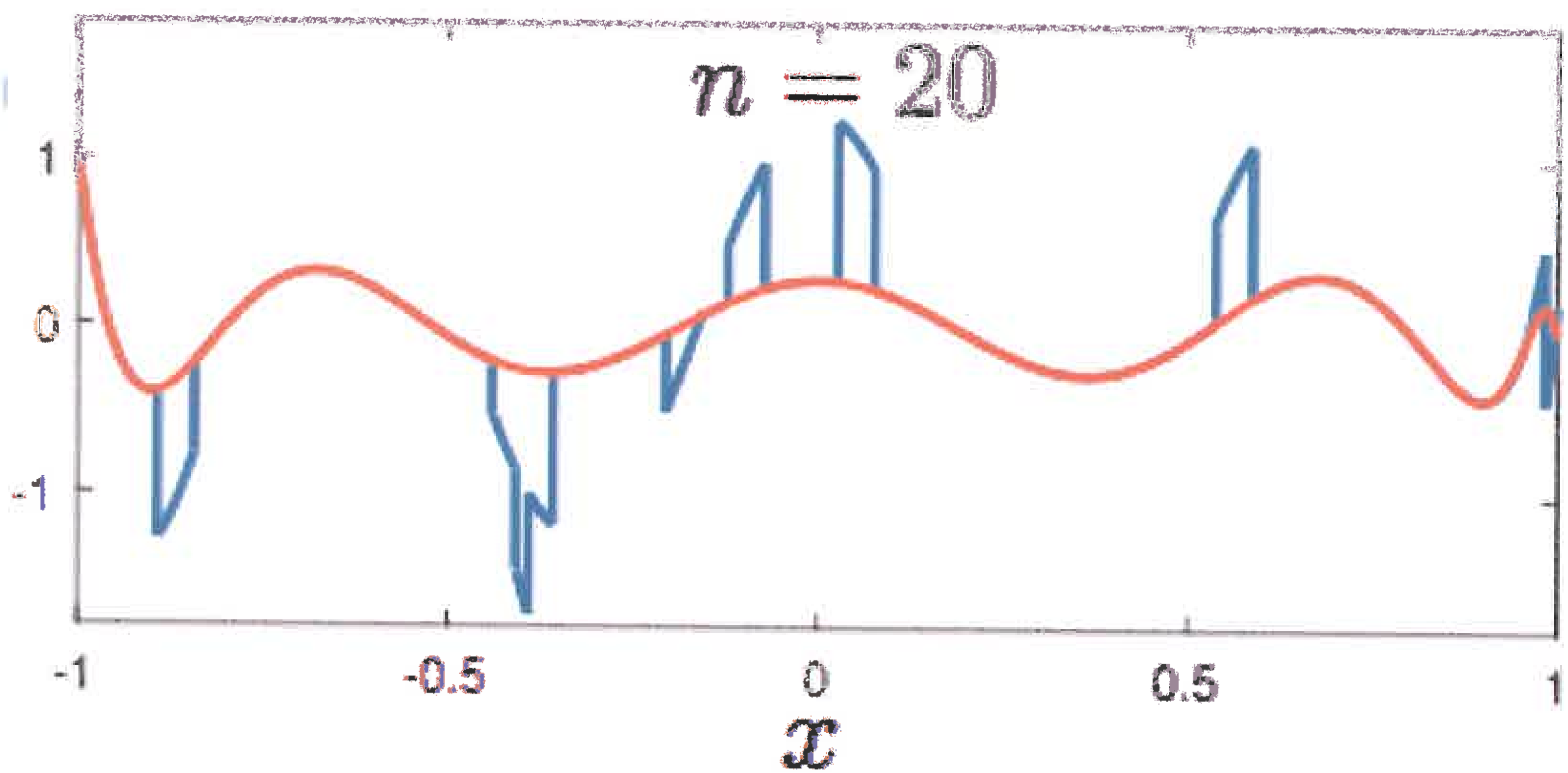
i.e. geometrically fast.



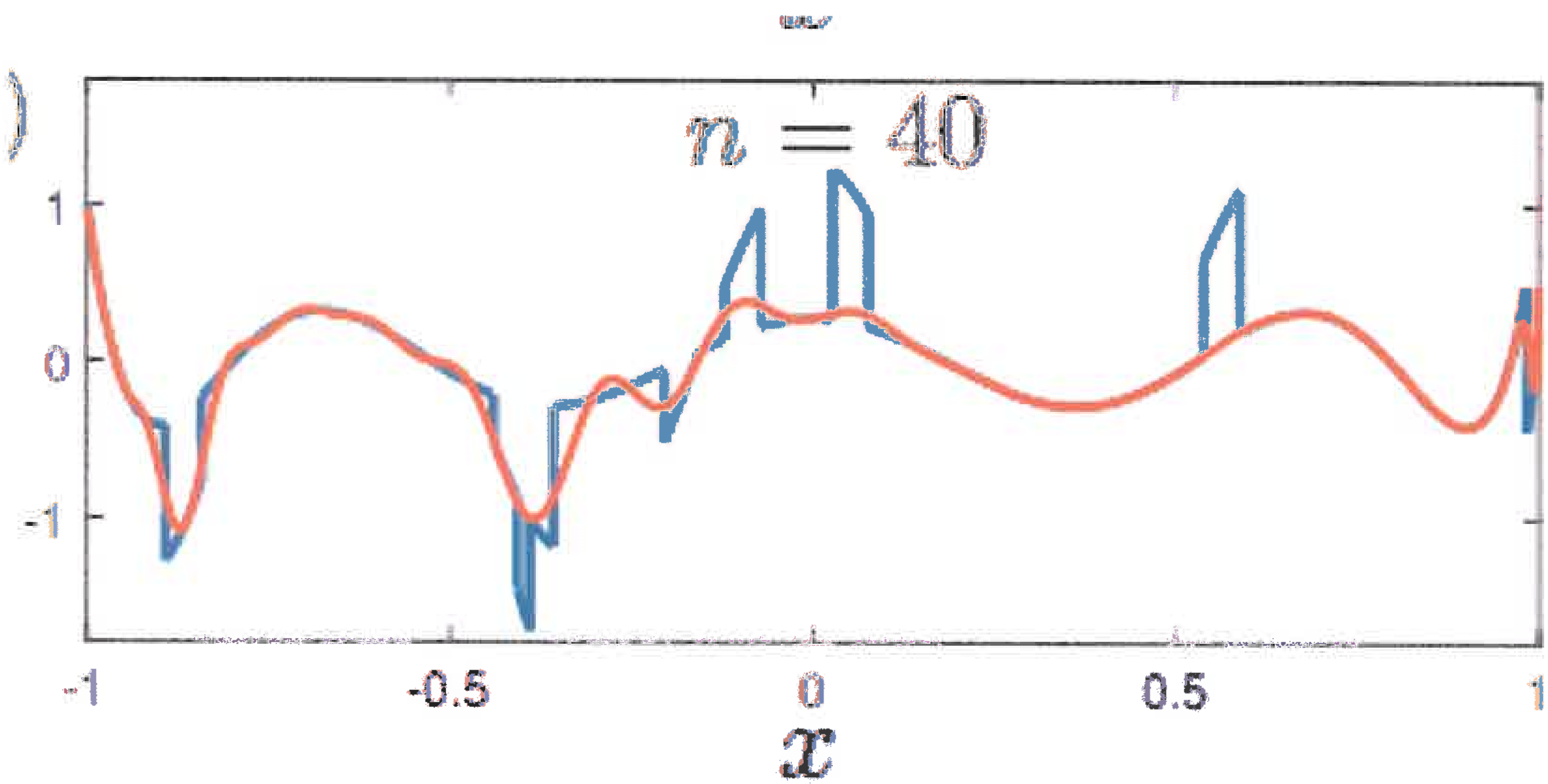
$n = 5$
1
0
-1
-1
-0.5
0
0.5
1
x



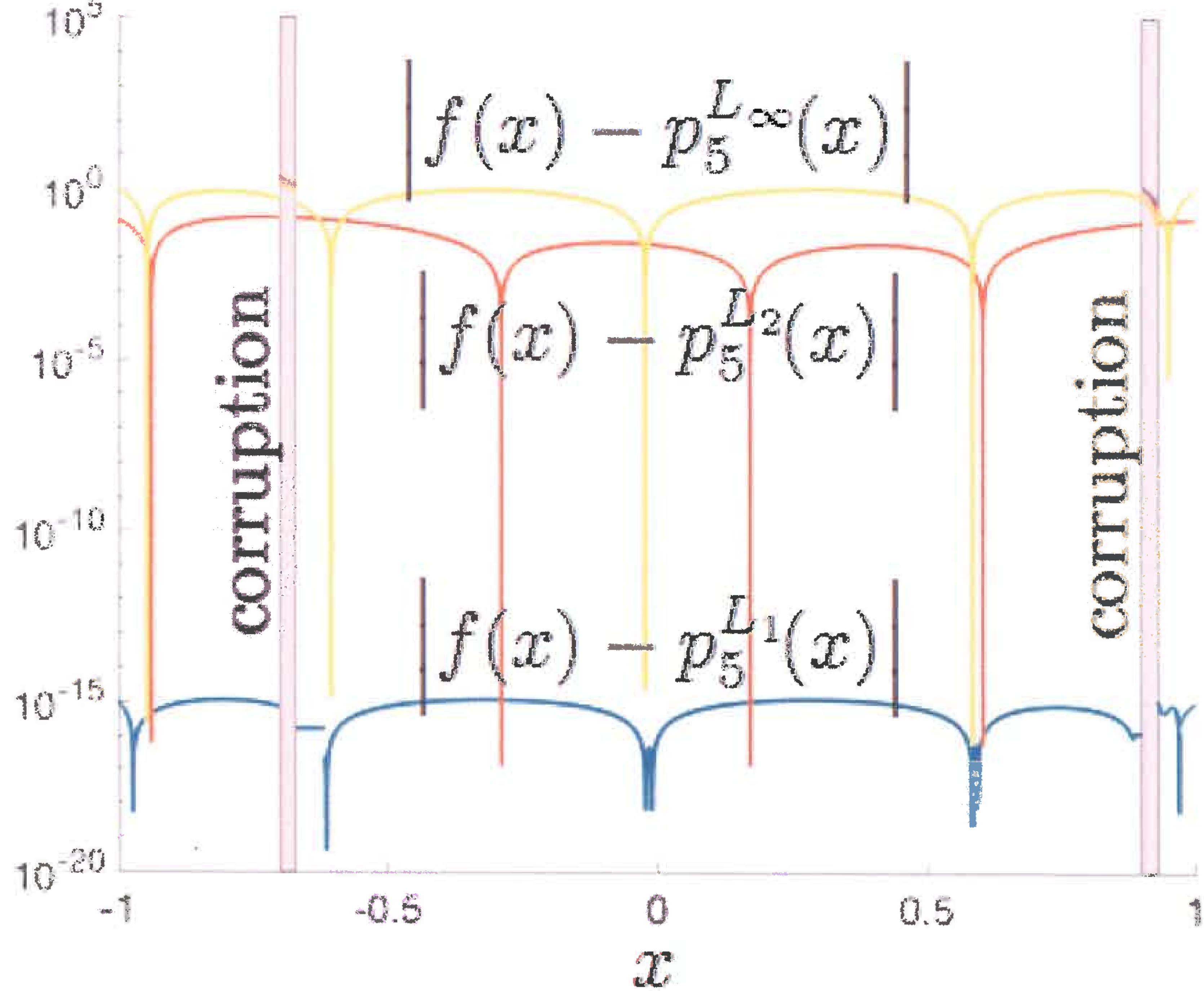
$n = 10$
1
0
-1
-1
-0.5
0
0.5
1
x



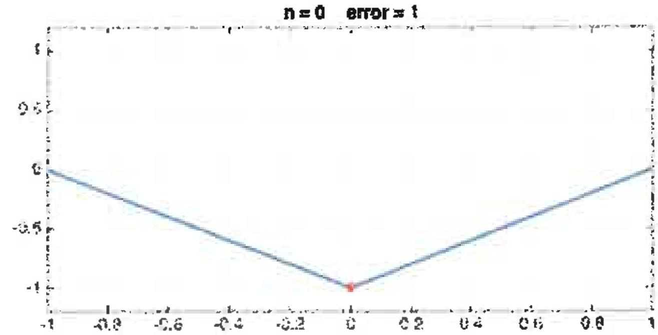
$n = 20$
1
0
-1
-1
-0.5
0
0.5
1
x

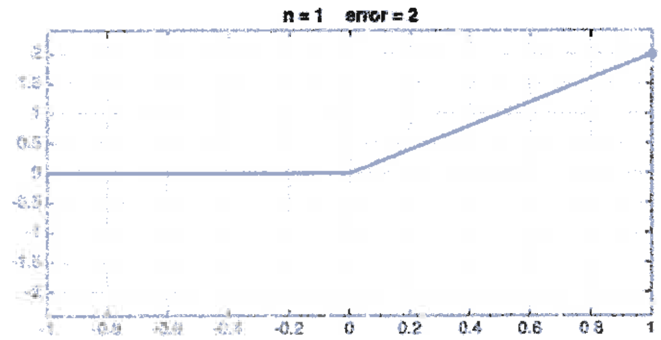


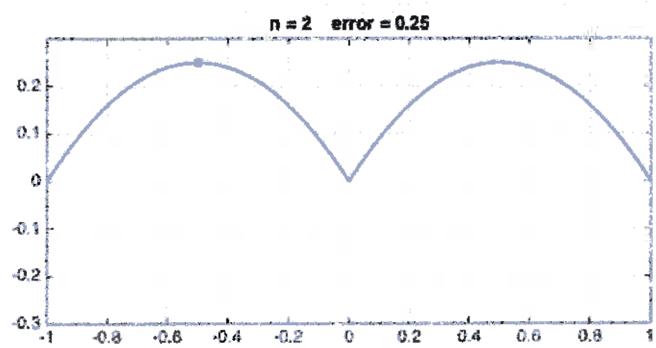
$n = 40$
1
0
-1
-1
-0.5
0
0.5
1
x

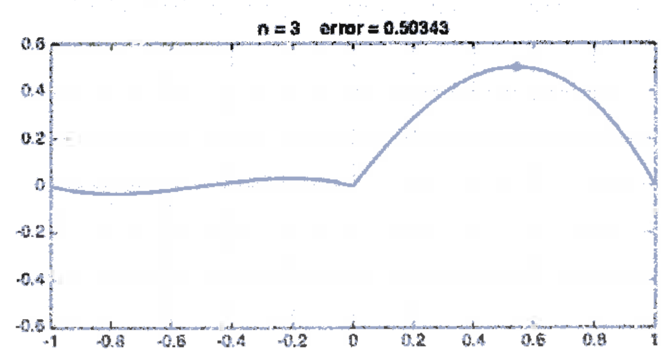


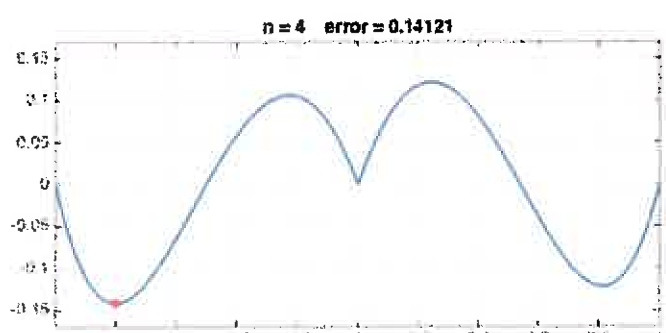
$|f(x) - p_5^{L_\infty}(x)|$
$|f(x) - p_5^{L_2}(x)|$
$|f(x) - p_5^{L_1}(x)|$
corruption
corruption
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
10^{-20}
-1
-0.5
0
0.5
1
x

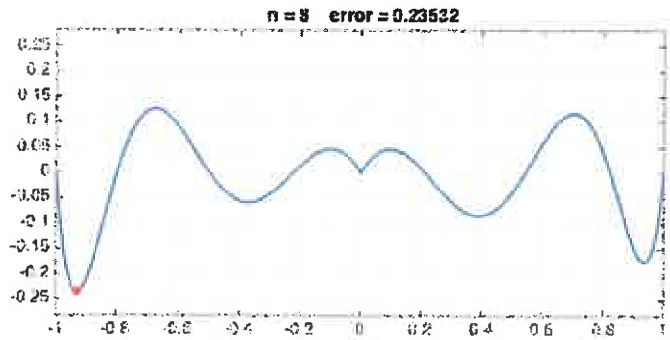

n = 0 error = 1

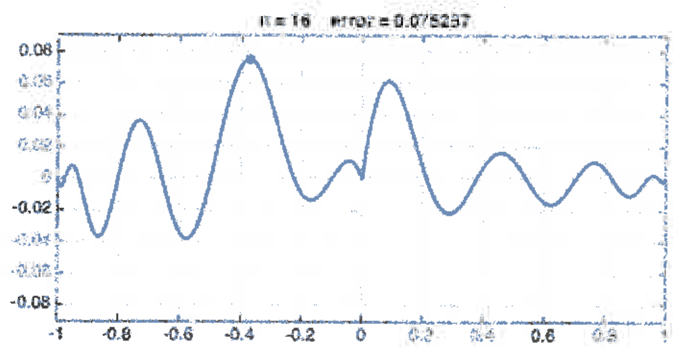

n = 1 error = 2

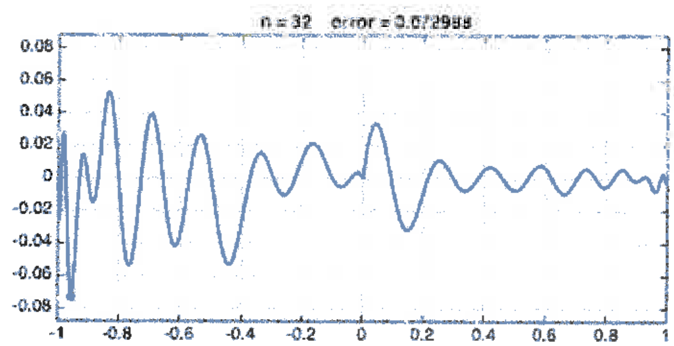

n = 2 error = 0.25

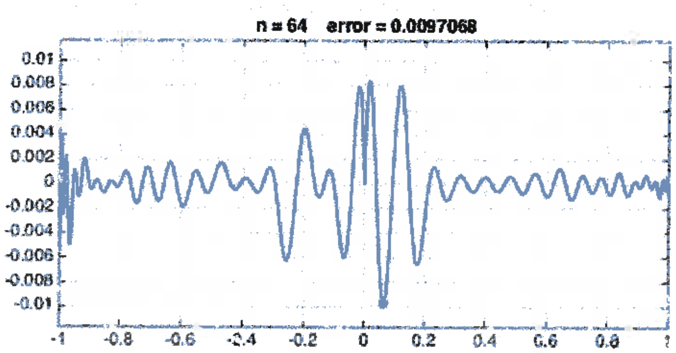

n = 3 error = 0.50343

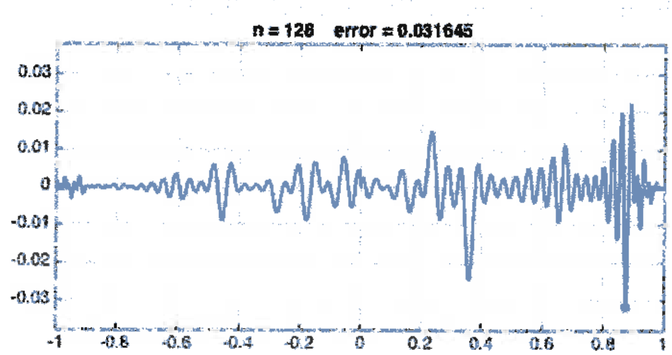

n = 4 error = 0.14121

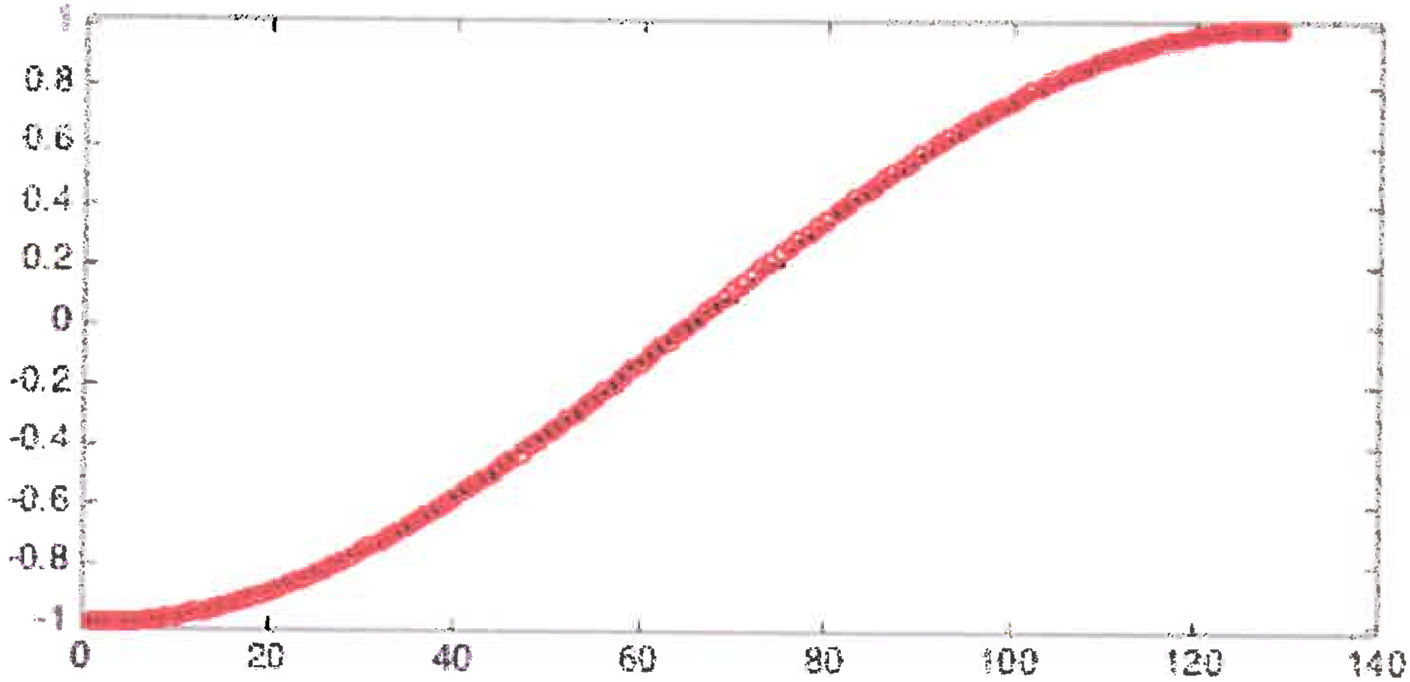

n = 8 error = 0.23532


n = 16 error = 0.075257

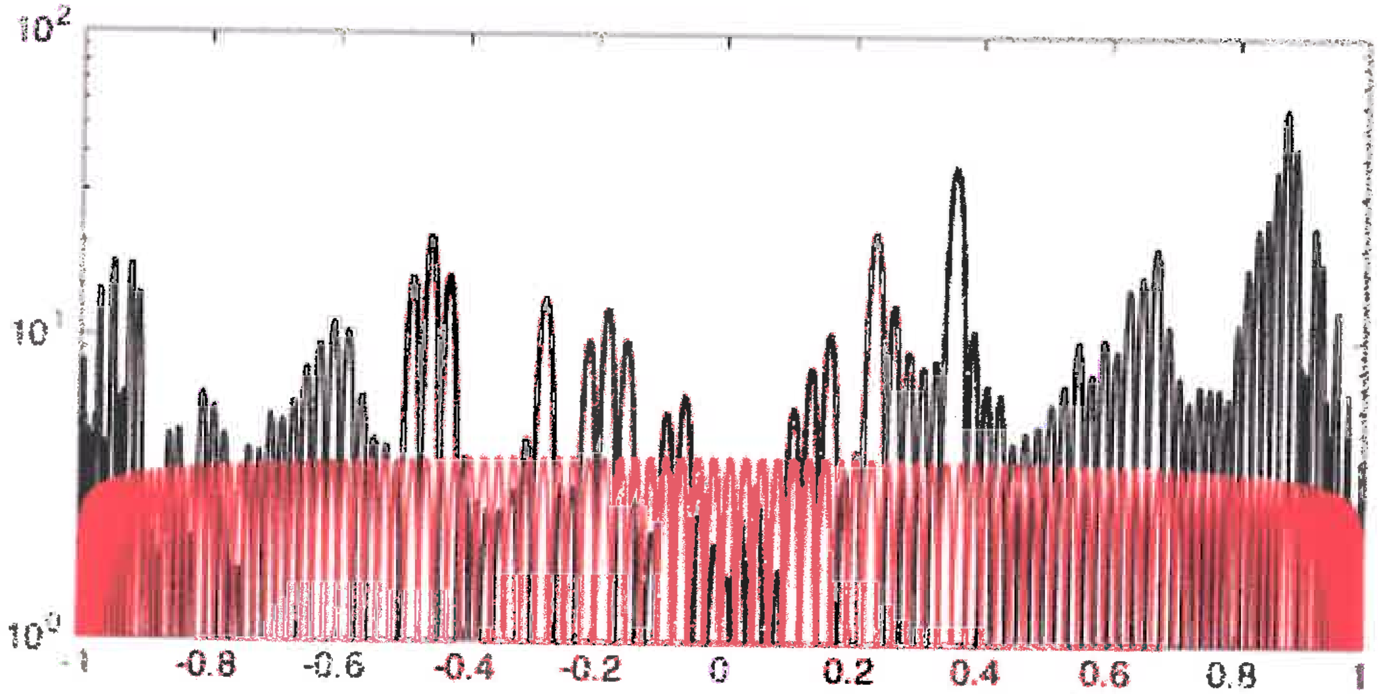

n = 32 error = 0.072988


n = 64 error = 0.0097068


n = 128 error = 0.031645



0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
20
40
60
80
100
120
140


10^2
10^1
10^0
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1