Bernstein Bases are Optimal,

but, sometimes,

Lagrange Bases are Better

Robert M. Corless Stephen M. Watt

Ontario Research Center for Compuer Algebra
University of Western Ontario
London Canada

Conditioning

- *Conditioning* of a problem measures the sensitivity to changes in the problem data.
- This is an intrinsic property of the problem and has nothing to do with the method by which it is solved.
- Useful concept in determining sensitivity of problem solution to changes in the problem formulation.
- We are interested in how the choice of basis affects the condition of polynomial rootfinding.

Polynomial Representation

- Polynomials in x of degree d with coefficients in k form a vector space $k_d[x]$ of dimension $d+1$.
- A *basis* is a set of $d+1$ polynomials $b_i(x) \in k_d[x]$ such that for every polynomial f there are c_i such that

$$f = \sum_{i=0}^{d} c_i \; b_i(x)$$

- Choice of basis

$$20 \; x^3 + 16 \; x^2 + 120 \; x^1 + 319 \; x^0$$

$$5 \; T_3(x) + 8 \; T_2(x) + 135 \; T_1(x) + 327 \; T_0(x)$$

$$\frac{5}{2} \; H_3(x) + 4 \; H_2(x) + 75 \; H_1(x) + 327 \; H_0(x)$$

Root Finding Condition Number

- Perturb polynomial coefficients

$$f + \Delta f = \sum_{i=0}^{d} (c_i + \delta_i c_i)\ b_i(x)$$

- How much do roots of $f + \Delta f$ differ from those of f?

- For each root r of f, there will be a corresponding root $r + \Delta r$ of $f + \Delta f$

- We wish to determine the condition $|\Delta r|/||\delta||_\infty$ of the root r to perturbations in the polynomial f represented in base b.

Root Finding Condition Number

- Let $\Omega \subset \mathbb{R}$ be the region of interest and $T \subset \Omega$ characterize the non-negativity of the basis b.

- We define

$$
\begin{aligned}
C_b(f,x) &:= \frac{1}{|f'(x)|} \sum_{i=0}^{n} |c_i b_i(x)| \qquad (1) \\
\text{cond}(b; f,x) &:= \frac{C_b(f,x)}{\|f\|_\infty} \qquad (2) \\
\text{cond}_T(b,f) &:= \sup_{x \in T} \text{cond}(b; f,x) \qquad (3)
\end{aligned}
$$

- Then with $\varepsilon = ||\delta||_\infty$ we have as $\varepsilon \to 0$

$$|\Delta r| = C_b(f,r)\ \varepsilon + O(\varepsilon^2)$$

Bernstein Bases

- The Bernstein basis for polynomials of degree d are

$$B_i^d(x) = \begin{pmatrix} d \\ i \end{pmatrix} x^i \ (1-x)^{d-i}, \qquad\qquad i = 0, ..., d$$

- Each member of the basis is of degree d, e.g. for $d = 3$

$$B_0^3(x) = (1-x)^3, \quad B_1^3(x) = 3x(1-x)^2, \quad B_2^3(x)x^2(1-x), \quad B_3^3(x) = x^3$$

- Well suited to applications in CAGD.

- Previous authors (e.g. Farouki, Lyche, Winkler...) have examined Bernstein bases for evaluation and root finding, and have shown that in a certain sense they are optimal.

Lagrange Bases

- A Lagrange basis for polynomials of degree d

$$L_i^X(x) = \prod_{\substack{j=0 \\ j \neq i}}^{d} \frac{x - x_j}{x_i - x_j}, \qquad i = 0, ..., d$$

  uses values at $d + 1$ disctinct points $X = \{x_i\}_{i=0}^{d}$.

- Knowing the values y_i of a polynomial at the points X lets us write

$$y(x) = \sum_{i=0}^{d} y_i \; L_i^X(x)$$

- Used in algorithms for computing polynomial algebra by values.

Extending Bases

- We consider polynomial basis families where the basis for $k_{d-1}[x]$ is not necessarily a subset of the basis for $k_d[x]$
- This is not so strange, e.g. the basis $\{[1,0],[0,1]\}$ of $\mathbb{R}^2$ is not a subset of the basis $\{[1,0,0],[0,1,0],[0,0,1]\}$ of $\mathbb{R}^3$.
- Examples of extending polynomial bases

Lift $i = 0...,d-1$	**New** d
$T_i(x) = T_i(x)$	$T_d(x) = 2xT_{d-1}(x) - T_{d-2}(x)$
$B_i^d(x) = \frac{d}{d-1}B_i^{d-1}(x)$	$B_d^d(x) = x^d$
$L_i^{X\cup\{x_d\}}(x) = \frac{x-x_d}{x_i-x_d}L_i^X(x)$	$L_d^{X\cup\{x_d\}} = \prod_{j=0}^{d-1} \frac{x-x_j}{x_i-x_j}$

Generalized Companion Matrix Root Finding Method

- Let $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$
- Consider the *pair* of matrices

$$C_0 = \begin{bmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{bmatrix} \qquad C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a_3 \end{bmatrix}.$$

- It is easy to see that $\det(xC_1 - C_0) = p(x)$.
- C_0 and C_1 are called the “companion matrix pencil” for $p(x)$.
- **The generalized eigenvalues of the pencil are the roots of** $p(x)$.

GCM in Lagrange Basis

- Consider for expository purposes the degree 3 polynomial on 4 points x_0, x_1, x_2, and x_3:

$$p(x) = p_0 L_0(x) + p_1 L_1(x) + p_2 L_2(x) + p_3 L_3(x) \,.$$

- In the Lagrange basis, we have [Corless 2003]

$$\mathbf{C_0} = \begin{bmatrix} x_0 & & & & p_0 \\ & x_1 & & & p_1 \\ & & x_2 & & p_2 \\ & & & x_3 & p_3 \\ -\ell_0 & -\ell_1 & -\ell_2 & -\ell_3 & 0 \end{bmatrix}$$

where $\ell_0 = 1/((x_0 - x_1)(x_0 - x_2)(x_0 - x_3))$ is the first Lagrange polynomial coefficient, ℓ_1 is the second, and so on.

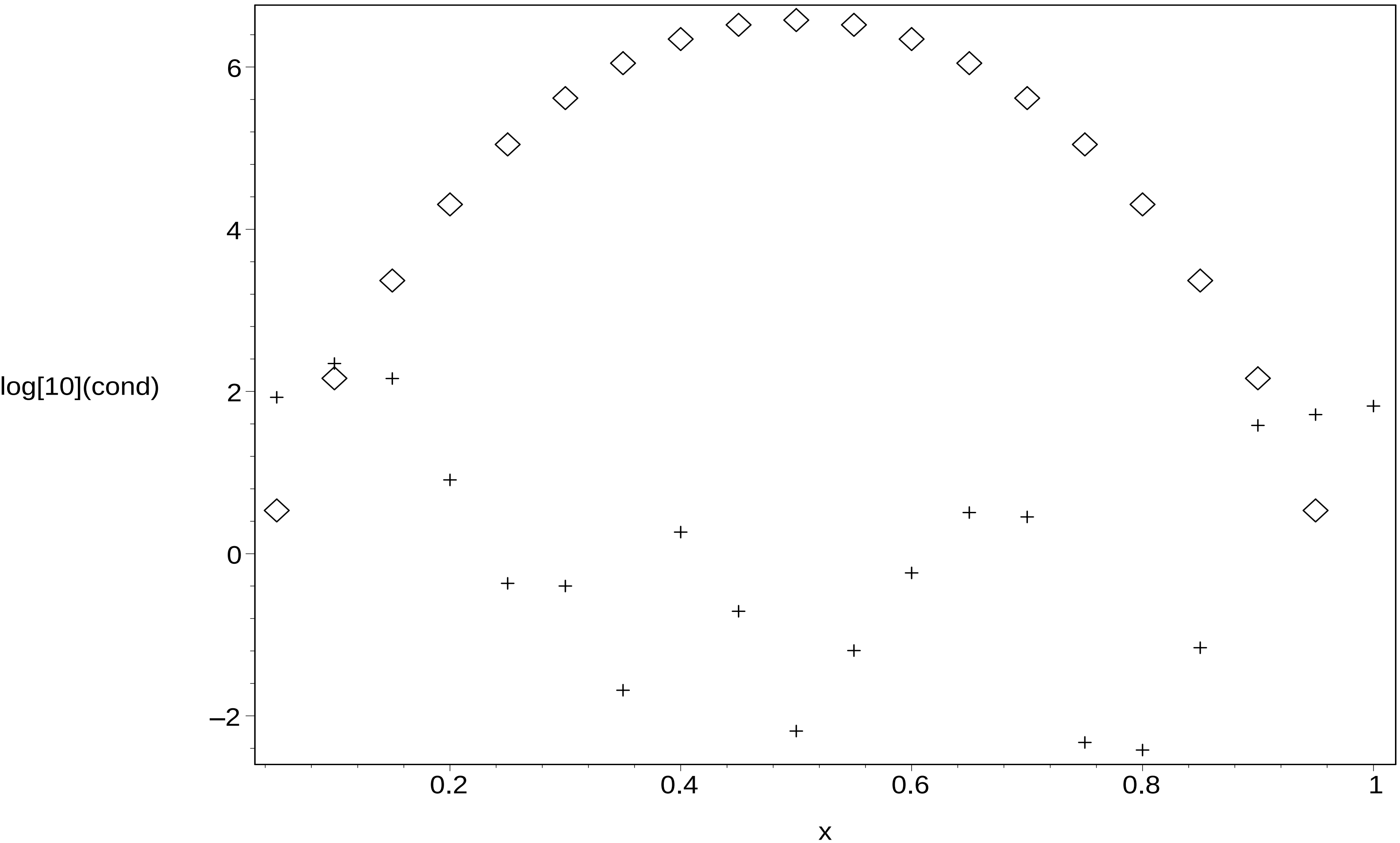
- Again, $\mathbf{C_1}$ is the identity matrix except that the final entry is zero.

Example 1 : The Wilkinson Polynomial

- Use the Wilkinson polynomial example of Farouki and Goodman

$$W_1 = \prod_{k=1}^{20} \left(x - \frac{k}{20} \right).$$

- We choose 20 random points on $(0, 1)$
- Evaluate W_1 at each of these points and compute eigenvalues of GCM.
- Roots found with maximum error of 7.1×10^{-12}
- C.f. 10^{-7} with Bernstein basis.



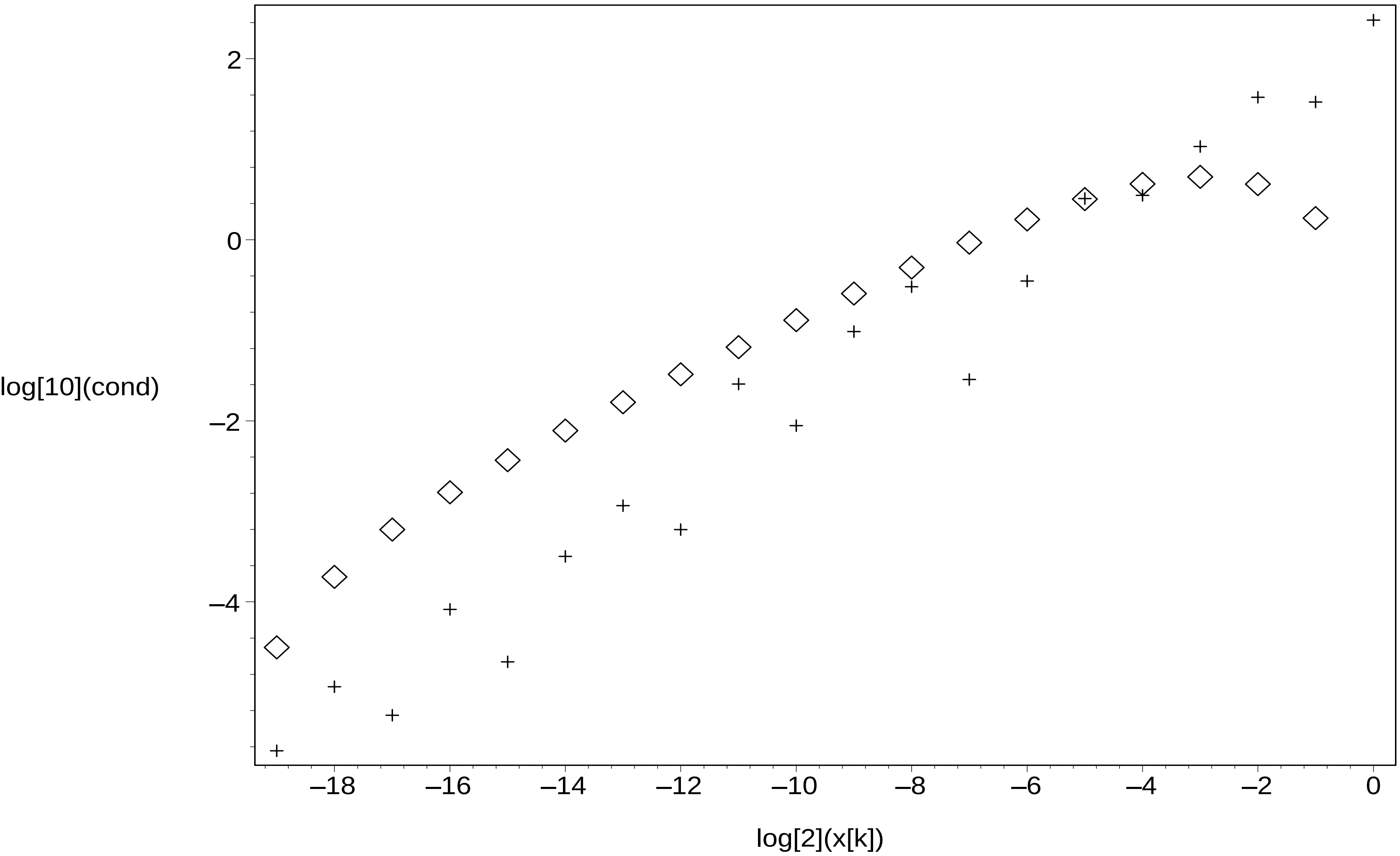
Root condition of the Wilkinson polynomial $\prod_{i=1}^{20}(x - \frac{i}{20})$ expressed in the Bernstein basis (diamonds) and a random Lagrange basis (crosses).

Example 2: The 2nd Wilkinson Polynomial

- We use the second Wilkinson polynomial example of FG

$$W_2 = \prod_{k=0}^{19} \left(x - 2^{-k}\right)$$

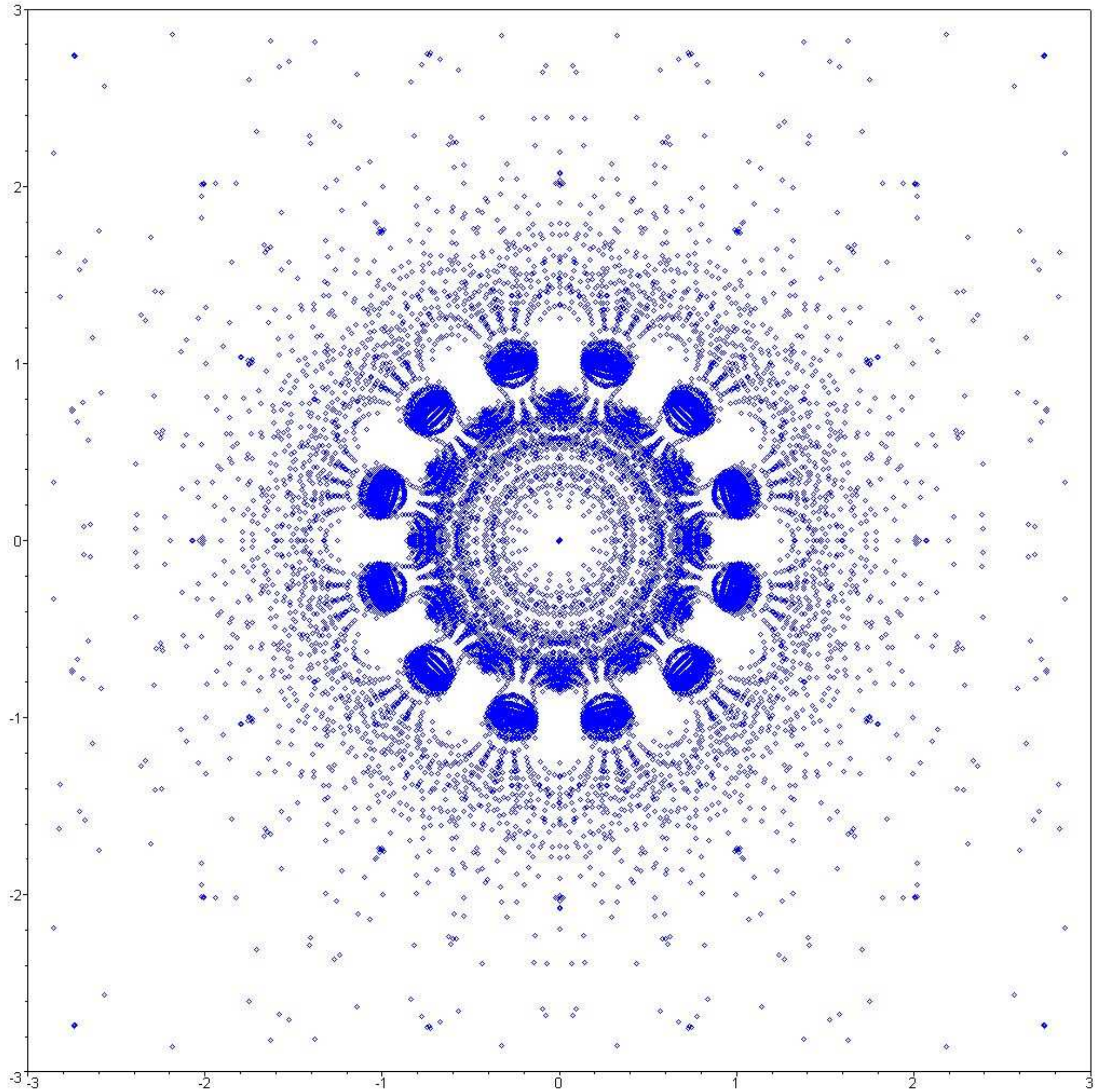
- FG find Bernstein basis only slightly better than power basis.
- Random Lagrange is worse.
- Lagrange with points chosen randomly in $[2^{-k-1}, 2^{-k}]$ does well. (This choice is reasonable if we suspect roots near origin.)
- Smallest root accurate to 11 places.



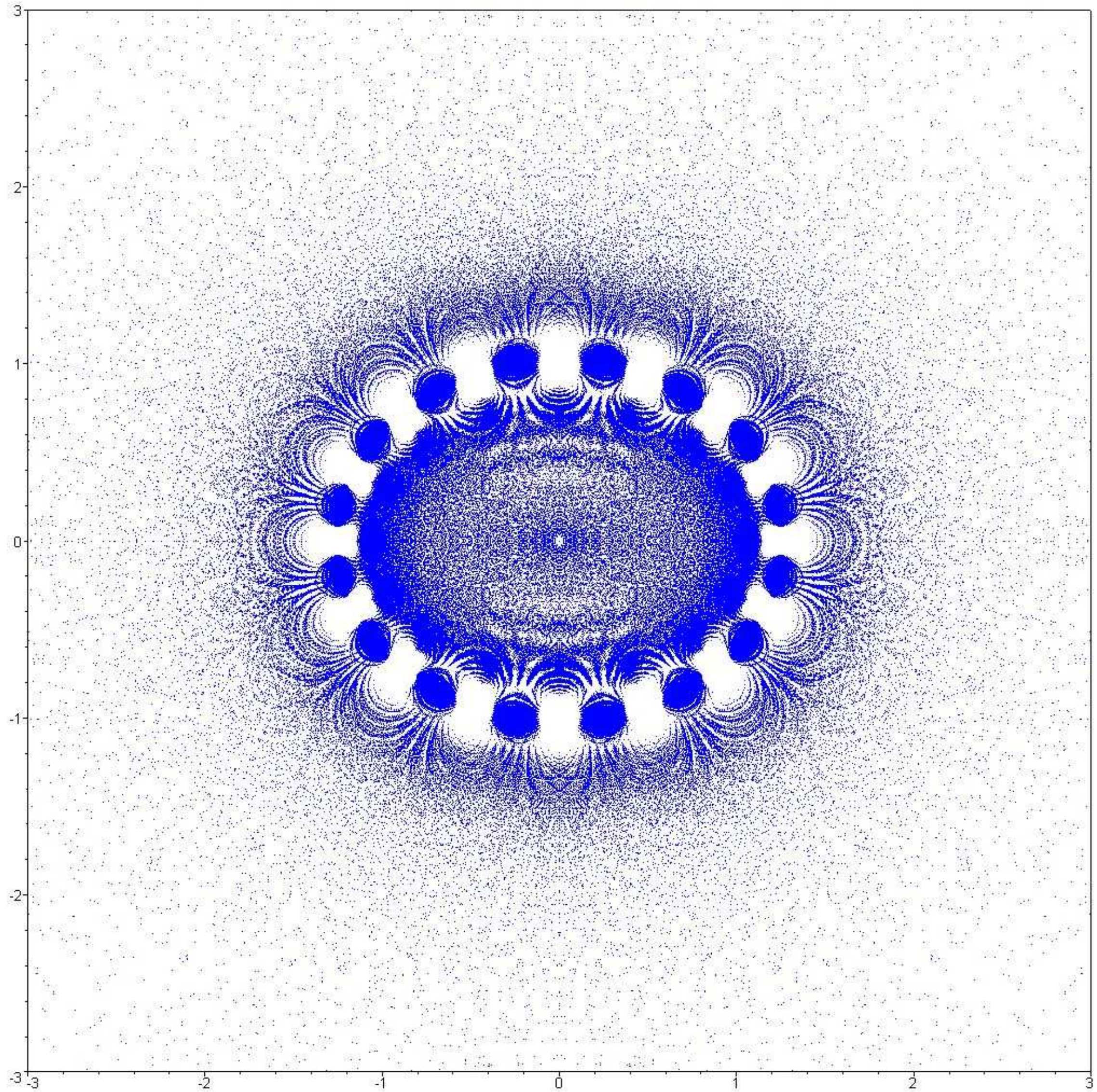
Root condition of the second Wilkinson polynomial $\prod_{i=0}^{19}(x-2^{-i})$ expressed in the Bernstein basis (diamonds) and a nonuniform random Lagrange basis (crosses).

Example 4: Structures in Number Theory

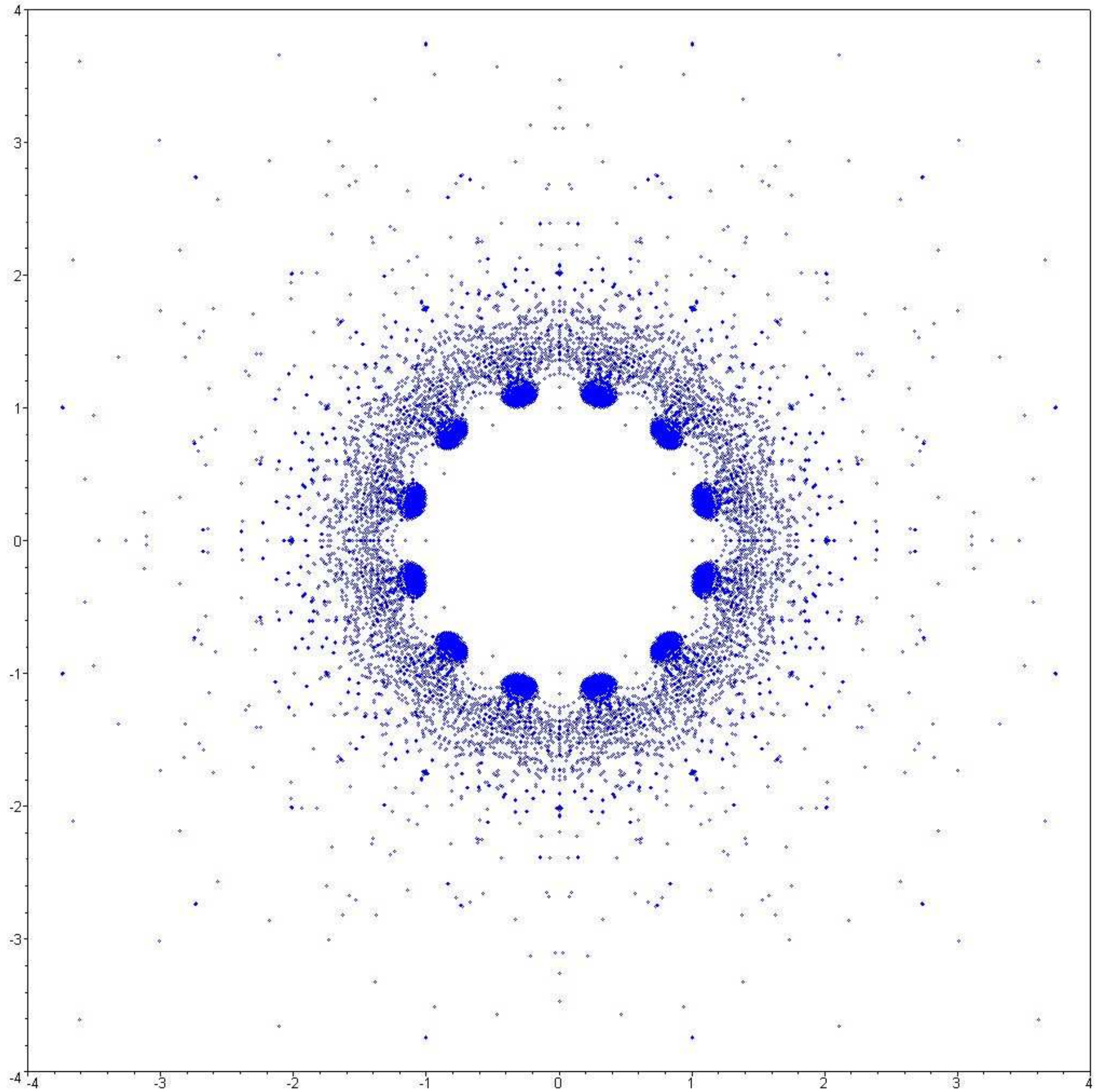
- We consider certain sets of polynomials taking on values from $\{-1, 0, +1\}$ at equally spaced points.



All roots of polynomials taking on values ± 1 at the 12^{th} roots of unity.



All roots of polynomials taking on values ± 1 at 16 equally spaced points on ellipse (> 400000 points)



All roots of polynomials taking on values $\{0, 1\}$ at 12^{th} roots of unity.

- First problem is all polynomials with values ± 1 at 12^{th} roots of unity.
- $N \times 2^N = 49152$ roots.
- Sampled check found roots were computed accurately to all but last digit.

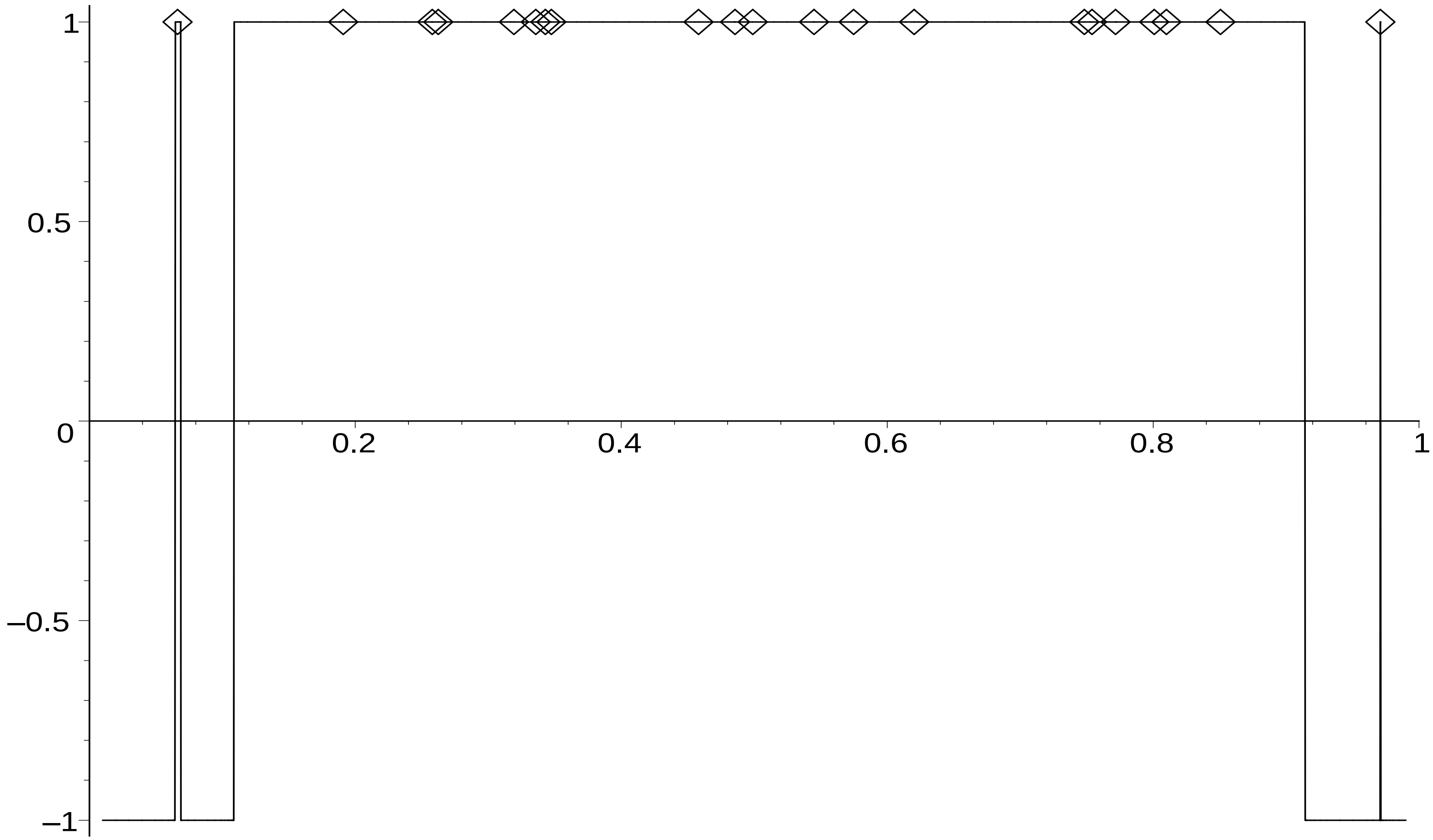
Proposition 1

Fix a set of interpolation points $[x_0, x_1, ..., x_n]$.

If any basis B *can be expressed as a non-negative combination of the Lagrange basis on this set of points, then there exists a set* T, *depending on* f *and containing the interpolation points, in which* $C_L(f,T) \leq C_B(f,T)$.

If, further, the inequality is strict on an interpolation point, that is $C_L(f,T) < C_B(f,T)$, *then the set* T *has a non-empty interior.*

The size of T relative to Ω is of immediate practical interest.



Sign of $C_B(W_1, t) - C_L(W_1, t)$ for Example 1.

The Lagrange basis is better at the positive values.

Proposition 2

If we choose n of our $n+1$ interpolation points to be the roots, then $C_L(f,r)=0$. That is, if we are lucky enough to interpolate at all the roots, the conditioning is perfect.

Conclusions

- Broadening the range of bases considered gives us choices that are better in certain regards.
- Compared to other popular choices, Lagrange bases are well-conditioned for root-finding.
- Remark: We can *over-sample* and get even better behaviour!