H–bases and approximate varieties

Tomas Sauer
Lehrstuhl für Numerische Mathematik
Justus–Liebig–Universität Gießen
Tomas.Sauer@math.uni-giessen.de

23 February 2006

> There's al–gebra. That's like sums with letters. For . . . for people whose brains aren't clever enough for numbers, see?
>
> — Terry Pratchett, *Jingo*

The original problem

- **Given:** finite set of points $\Xi \subset \mathbb{R}^d$.
- **Wanted:** "Dimension reduction" – if possible.
- **Idea:** Find algebraic variety $\mathscr{V}$ such that $\Xi \subset \mathscr{V}$.
- **Side conditions:**
  - Many points: $\#\Xi$ is large.
  - Noisy data: Ξ is only approximation.
- **Applications:**
  - Learning Theory.
  - Brain Science (Microstate Analysis).

What is the problem?

- Determination of variety $\mathscr{V}(\Xi)$ – determination of basis for the ideal

$$\mathscr{I}_\Xi = \{f \in \mathbb{R}[x] \ : \ f(\Xi) = 0\}, \qquad x = (x_1, \dots, x_d).$$

- Computation of Gröbner basis for $\mathscr{I}_\Xi$ from Ξ:

Buchberger–Möller algorithm, 1982

- However:
  - Numerical application of Gröbner bases can be dangerous.
  - What about inexact points?
  - Do we want a zero dimensional ideal at all?

A very simple example

- Points on a line:

$$\Xi = \{k\eta \ : \ k = 0, \dots, N\}, \qquad \eta \in \mathbb{R}^2_+.$$

- Gröbner basis (with $x \prec y$)

$$\left\{ x\,(x - \eta_1) \cdots (x - N\eta_1), y - \frac{\eta_2}{\eta_1} x \right\}.$$

- Roles: "Variety" – "Finiteness of Ξ".
- Distinction: **total degree**.
- "Gröbner problem": $\eta_1 \to 0$. "Representation singularity".

The goals

(i) Determine a “low degree” part of the ideal.

(ii) Use associated variety.

(iii) Find a numerically stable procedure.

(iv) Continuity with respect to Ξ.

- “Solution” for the simple example:

$$\left\{ \frac{\eta_2}{\|\eta\|_2} x - \frac{\eta_1}{\|\eta\|_2} y \right\}, \qquad \eta \in \mathbb{R}^2 \setminus \{0\}.$$

- H–basis, normalized, changes continuously with η.

H–bases

Definition. $H \subset \mathbb{R}[x]$ *is called* H–basis *for an ideal* $\mathscr{I}$ *if*

$$f \in \mathscr{I} \quad \Leftrightarrow \quad f = \sum_{h \in H} f_h \, h, \quad \deg f \geq \deg f_h + \deg h.$$

- Alternatively:

$$\Lambda(\mathscr{I}) = \langle \Lambda(H) \rangle, \qquad \Lambda = \text{leading term}.$$

Homogeneous ideals

- Of course:

> Any Gröbner basis with respect to a graded term order is also an H–basis.

- So – what's different? No term orders!

Homogeneous reduction

- Set

$$\Pi_k^0 := \left\{ f(x) = \sum_{|\alpha|=k} f_\alpha x^\alpha \; : \; f_\alpha \in \mathbb{R} \right\}, \qquad \Pi_k = \bigoplus_{j=0}^{k} \Pi_j^0.$$

Homogeneous of degree k Of degree at most k

- For $F \subset \mathbb{R}[x]$ and $k \in \mathbb{N}_0$ consider

$$V_k(F) := \left\{ \sum_{f \in F} g_f \Lambda(f) \; : \; g_f \in \Pi_{k-\deg f}^0, \; f \in F \right\} \subseteq \Pi_k^0 \qquad W_k(F) := \Pi_k^0 \ominus V_k(F)$$

with respect to *inner product* $(\cdot,\cdot) : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$.

Homogeneous reduction – The algorithm

- **Given:** $F \subset \mathbb{R}[x]$, $g \in \mathbb{R}[x]$.
- Iteration for $k = \deg g, \dots, 0$
  - ▷ Decompose
$$\Lambda(g) = \underbrace{\sum_{f \in F} g_f \, \Lambda(f)}_{\in V_k(F)} \oplus r_k, \qquad r_k \in W_k(F).$$
  - ▷ Continue with
$$\Pi_{k-1} \ni g \leftarrow g - \sum_{f \in F} g_f \, f - r_k$$
- **Result:** *remainder* $g_{\underset{F}{\to}} := r := \sum_{k=0}^{\deg g} r_k$.

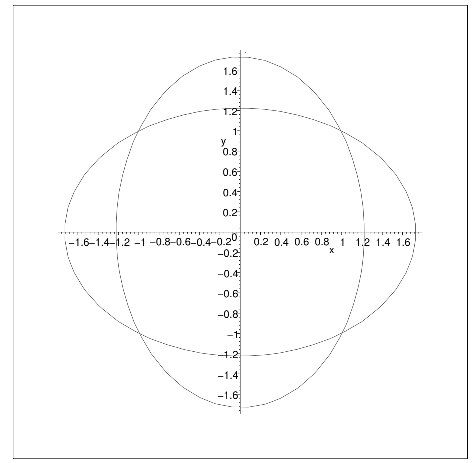
Homogeneous reduction – Properties

Theorem. *If* H *is an H–basis then* $f \underset{H}{\to}$ *depends only on* $\langle H \rangle$*.*

- Properties of the algorithm:
  - Only orthogonal projections – stable and continuous.
  - Can be used for H–basis computation – reduction of syzygies.
  - Variant of Buchberger's algorithm.
  - **But:** Basis for module of syzygies is more complicated.
- H–basis construction available for zero–dimensional ideals [Möller&Sauer '00].
- Homogeneous H–bases are **not** unique, not even reduced ones.
- *Normal form* $\nu_{\langle H \rangle}(f) := f \underset{H}{\to}$.

H–bases and system solving

- Intersection of two ellipses [Möller&Sauer '99]:



$$
\begin{aligned}
0 &= f_1(x,y) = \frac{1}{3}x^2 + \frac{2}{3}y^2 - 1 \\
0 &= f_2(x,y) = \frac{2}{3}x^2 + \frac{1}{3}y^2 - 1
\end{aligned}
$$

- Perturbation:

$$
\widehat{f}_2(x,y) = f_2\left(R_\varphi \begin{bmatrix} x \\ y \end{bmatrix}\right), \qquad R_\varphi = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}
$$

- Solve by eigenvalue method, see [Auzinger&Stetter, Möller&Stetter].

With Gröbner bases . . .

- Quotient spaces $\mathscr{P} = \mathbb{R}[x]/\langle F \rangle$

$$\mathscr{P}^{\varphi}_{\mathrm{lex}} = \begin{cases} \{1, x, x^2, x^3\} & \varphi \neq 0, \\ \{1, x, y, xy\} & \varphi = 0, \end{cases} \qquad \mathscr{P}^{\varphi}_{\mathrm{glex}} = \begin{cases} \{1, x, y, x^2\} & \varphi \neq 0, \\ \{1, x, y, xy\} & \varphi = 0. \end{cases}$$

- Multiplication matrices:

$$M_x^0 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad M_y^0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

. . . we are in trouble

- For $\varphi \neq 0$ we have

$$M^{\varphi}_{x,\mathrm{lex}} = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad M^{\varphi}_{y,\mathrm{lex}} = \begin{bmatrix} 0 & \star & 0 & \star \\ \star & 0 & \star & 0 \\ 0 & \star & 0 & \star \\ \star & 0 & \star & 0 \end{bmatrix}$$

and

$$M^{\varphi}_{x,\mathrm{glex}} = \begin{bmatrix} 0 & 0 & \star & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \star & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad M^{\varphi}_{y,\mathrm{glex}} = \begin{bmatrix} 0 & \star & \frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 1 & \star & 0 & 1 \\ 0 & 0 & -\frac{1}{2} & 0 \end{bmatrix}$$

- **Bad news:** $\varphi \to 0$ does **not** reproduce $M^{0}_{x/y}$ and $\lim_{\varphi \to 0} \star = \infty$.

- For φ small: **Complex** zeros in `Maple`.

On the contrary

- Proper homogeneous H–basis ("least interpolation") gives

$$\mathscr{P}_H^\varphi = \left\{ 1, x, y, \frac{8\sin\varphi\ x^2 - \cos\varphi\ xy - 4\sin\varphi\ y^2}{\left(40\sin^2\varphi + \cos^2\varphi\right)^{1/2}} \right\}.$$

Continuous change of spaces.

- Matrix

$$M_{x,H}^\varphi = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = M_x^0.$$

More varying coefficients.

The interpolation problem

- Another approach to zero–dimensional ideals:

  Dual functionals – interpolation

- *Lagrange interpolation problem*:

  To $\Xi \subset \mathbb{R}^d$ and $f \in \mathbb{R}[x]$ find $Lf \in \mathbb{R}[x]$ such that $f(\Xi) = Lf(\Xi)$.

- *Interpolation space* $\mathscr{P} \subset \mathbb{R}[x]$:

  For $f \in \mathbb{R}[x]$ there exists a **unique** $L_{\mathscr{P}} f \in \mathscr{P}$ such that $f(\Xi) = L_{\mathscr{P}} f(\Xi)$.

- *Degree reducing* interpolation space:

$$\deg L_{\mathscr{P}} f \le \deg f, \qquad f \in \mathbb{R}[x].$$

Degree reducing interpolation

- Introduced by de Boor&Ron, "least interpolant"

$$\mathscr{P} := \bigcap_{f \in \mathscr{I}_\Xi} \Lambda(f)(D).$$

Joint kernel of differential operators, *Box spline* theory.

- "Natural interpolation space": normal forms $\mathscr{P} := \nu_{\mathscr{I}_\Xi}(\mathbb{R}[x])$.
- Difference in $\mathscr{I}_\Xi \Rightarrow$ Remainder interpolates.
- "Least" is normal form for

$$(f, g) = (f(D)g)(0) = \sum_{\alpha \in \mathbb{N}_0^d} \alpha!\, f_\alpha\, g_\alpha.$$

Newton bases

- Recall univariate case: $\Xi = \{\xi_0, \dots, \xi_n\}$.

- Degree reducing interpolation space: Π_n and

$$L_n f(x) = \sum_{k=0}^{n} [\xi_0, \dots, \xi_k] f \cdot (x - \xi_0) \cdots (x - \xi_{k-1}).$$

Newton form of the interpolation polynomial, divided differences.

- *Newton basis* $p_k(x) := (x - \xi_0) \cdots (x - \xi_{k-1})$ satisfies

$$p_k(\xi_j) = 0, \qquad p_k(\xi_k) \neq 0, \qquad 0 \leq j < k \leq n.$$

Newton bases II

- *Degree of a space:* $\deg \mathscr{P} := \max_{p \in \mathscr{P}} \deg p$.
- $N = \{N_k \ : \ k = 0, \dots, \deg \mathscr{P}\}$ is a *Newton basis* for $\mathscr{P}$ with respect to Ξ if
  - ▷ $N_k \subset \Pi_k$, $k = 0, \dots, \deg \mathscr{P}$.
  - ▷ there exists a decomposition

$$\Xi = \bigcup_{k=0}^{\deg \mathscr{P}} \Xi_k \qquad \text{such that} \qquad N_k\left(\Xi_j\right) = \delta_{jk} I, \quad 0 \le j \le k \le \deg \mathscr{P}.$$

  - ▷ we have that

$$\Pi_k^0 = \operatorname{span} \Lambda\left(N_k\right) \oplus \left(\Lambda(\mathscr{I}_\Xi) \cap \Pi_k^0\right).$$

Direct sum, not orthogonality!

Newton bases III

- Newton basis in $\nu_{\mathscr{I}_\Xi}(\mathbb{K}[x])$ is characterized entirely by

$$N_k(\Xi_j) = \delta_{jk} I, \qquad 0 \le j \le k \le \deg \mathscr{P}.$$

- Recall: Orthogonality of leading terms.

Theorem. *$\mathscr{P} \subset \mathbb{R}[x]$ is a degree reducing interpolation space if and only if $\mathscr{P}$ has a Newton basis.*

- **Algorithm:** Construct $\mathscr{P}$ by determining a Newton basis.
- **By–product:** H–basis for $\mathscr{I}_\Xi$.
- Use this to compute approximate variety.

The algorithm – principal remarks

- Computation of Newton and ideal basis by
  - Gauss elimination.
  - Gram–Schmidt for partial orthogonalization.
  - Gauss elimination by segments.
- Essentially the same. And like BM!
- Automatically determines decomposition of Ξ. In no way unique!
- "Pivoting strategies" influence numerical outcome.

Fundamental Theorem of Numerical Analysis: Details matter!

Approximating BM

☐ How to approximate BM well?

Pragmatic approach! BM and approximators.

The algorithm

- Initialize

$$k \leftarrow 0, \qquad \Xi' \leftarrow \Xi, \qquad H \leftarrow \emptyset.$$

- `while` $\Xi' \neq \emptyset$
  - ▷ Choose basis P of $W_k(H)$.
  - ▷ Set $P' \leftarrow P - L_{\mathscr{P}_{k-1}} P$; interpolation with respect to $\Xi_0, \dots, \Xi_{k-1}$.
  - ▷ Determine Ξ_k and decompose:

$$P' = F \cup Q, \qquad \text{s.th.} \qquad Q(\Xi) = 0, \quad \det F(\Xi_k) \neq 0.$$

  - ▷ Set $N_k \leftarrow F(\Xi_k)^{-1} F$.
  - ▷ Update

$$k \leftarrow k+1, \qquad \Xi' \leftarrow \Xi' \setminus \Xi_k, \qquad H \leftarrow H \cup Q.$$

The algorithm II – how to decompose

- Consider *Vandermonde matrix* $P'(\Xi')$.
- Rearrange P' and Ξ' such that

$$P'(\Xi') = \begin{bmatrix} P'_+(\Xi'_+) & P'_-(\Xi'_+) \\ P'_-(\Xi'_+) & P'_-(\Xi'_-) \end{bmatrix}, \qquad \operatorname{rank} P'(\Xi') = \operatorname{rank} P'_+(\Xi'_+) = \#\Xi'_+.$$

- Set $F = P'_+$, $\Xi_k = \Xi'_+$ and

$$Q = P'_- - P'_-(\Xi'_+)\, P'_+(\Xi'_+)^{-1}\, P'_+.$$

Gauss elimination in block matrix form! Dangerous!

The algorithm III – properties

Theorem. *The algorithm terminates after finitely many steps and yields*

(i) *a decomposition* $\Xi = \Xi_0 \cup \cdots \cup \Xi_n$.

(ii) *a Newton basis* $N_0, \dots, N_n$ *for a degree reducing interpolation space* $\mathscr{P}$.

(iii) *an H–basis* H *(after another iteration).*

- Important part: Determine rank of a matrix. **Numerically difficult**.
- Up to measure zero rank is always maximal!
- Possible inversion of **near singular** matrices. Numerical suicide.

Numerical aspects

- Distinction $f(\Xi) = 0$ for $f \in \mathscr{I}_\Xi$ is too hard.
- Use $|f(\Xi)| \le \varepsilon$ instead $\Rightarrow \mathscr{I}_\Xi(\varepsilon)$. No ideal!
- $\varepsilon =$ tolerance. Greater than "unit roundoff".
- Matter of *normalization*: $\frac{\varepsilon}{\max |f(\Xi)|} f \in \mathscr{I}_\Xi(\varepsilon)$.
- Normalization:

$$1 = \|f\|_2^2 = (f, f) = \sum_{\alpha \in \mathbb{N}_0^d} f_\alpha^2.$$

- Note

$$C_{\deg f}^{-1} \|f\|_2 \le \max_{x \in \Omega} |f(x)| \le C_{\deg f} \|f\|_2, \qquad \Omega \subset \mathbb{R}^d \text{ compact.}$$

Back to decomposition

- Vandermonde matrix

$$F(\Xi) = [\, f(\xi) \ : \ f \in F, \xi \in \Xi \,] = F \cdot \Xi^T.$$

- What are reasonable operations on F and Ξ?
  - F is basis of a vector space as well as

$$\left[g_f := \sum_{f' \in F} a_{f,f'} \, f' \ : \ f \in F \right] = A\, F, \qquad A = \left[a_{f,f'} \ : \ f, f' \in F \right].$$

  - Ξ can be permuted $\Rightarrow \Xi\, P$, P Permutation.

- "Invariance":

$$F(\Xi) \simeq (A\,F) \cdot (P\,\Xi)^T = A\,F(\Xi)\,P^T \simeq A\,F(\Xi)\,P.$$

A citation

> *The great stability of unitary transformations in numerical analysis springs from the fact that both the ℓ_2–norm and the Frobenius norm are unitarily invariant. This means in practice that even when rounding errors are made, no substantial growth takes place in the norms of the successive transformed matrices.*
>
> — J. H. Wilkinson, 1965

- Try to use as much orthogonality as possible. In particular for A.
- Fits well with the homogeneous reduction process.
- Apply stable orthogonal transformations.

Back to decomposition II

Credo. *Orthogonality is good!*

- In the matrix decomposition $F(\Xi) = A\, F(\Xi)\, P$ choose $A = Q$ *orthogonally* such that
$$Q\, F(\Xi)\, P = R = \begin{bmatrix} * & \dots & * \\ & \ddots & \vdots \\ & & * \end{bmatrix}.$$
- Method: Householder transformations!
- Action of P –*row pivoting* with respect to $|\cdot|_2$.
- Get biggest columns in front.

Properties of the QR–decomposition

Proposition. *Let* $R = Q\,F(\Xi)\,P$ *be obtained by* QR*–factorization with row pivoting*

(i) *the diagonal elements of* R *satisfy*

$$|r_{11}| \geq |r_{22}| \geq \cdots \geq |r_{nn}| .$$

(ii) *if* F *is orthonormal, i.e.* $(f, f') = \delta_{f,f'}$, $f, f' \in F$, *then*

$$\|(QF)_f\|_2 = 1, \qquad f \in F.$$

- Property (i) is a *stopping criterion*: stop if $|r_{kk}| \leq \varepsilon$.
- Property (ii) is *preservation of normalization*.

More on the stopping condition

- Further property of R:

$$|r_{jj}| \geq |r_{jk}|, \qquad k = j+1, \dots, N.$$

Diagonal elements dominate!

- With threshold ε and $|r_{kk}| > \varepsilon \geq |r_{k+1,k+1}|$

$$|R| \leq \left[\begin{array}{ccc|ccccc} |r_{11}| & \dots & |r_{11}| & |r_{11}| & \dots & |r_{11}| & \dots & |r_{11}| \\ & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ & & |r_{kk}| & |r_{kk}| & \dots & |r_{kk}| & \dots & |r_{kk}| \\ \hline & & & \varepsilon & \dots & \varepsilon & \dots & \varepsilon \\ & & & & \ddots & \vdots & \ddots & \vdots \\ & & & & & \varepsilon & \dots & \varepsilon \end{array}\right] := \left[\begin{array}{c|c} R_+ & A \\ \hline 0 & E \end{array}\right].$$

Approximate ideal functions and Newton basis

Consequence. *The polynomials* $q_j := (Q\,F)_j$, $j = k+1, \dots, n$ *satisfy*

$$\|q_j\|_2 = 1, \qquad |q_j(\Xi)| \le \varepsilon.$$

- $q_{k+1}, \dots, q_n$ are *ideal* part of F.
- These polynomials **really** vanish approximately at Ξ. Normalization!
- The "selected points" are $\Xi' = \left(P^T \Xi\right)_j$, $j = 1, \dots, k$.
- Newton basis:

$$N := \left[(Q\,F)_j \left(\left(P^T \Xi\right)_l \right) \ : \ j, l = 1, \dots, k \right]^{-1} \begin{bmatrix} (Q\,F)_0 \\ \vdots \\ (Q\,F)_k \end{bmatrix}.$$

More numerical considerations

- Algorithm works well if $|r_{kk}| \gg \varepsilon$.
- Otherwise: "almost numerically singular points".
- Condition number of inversion:

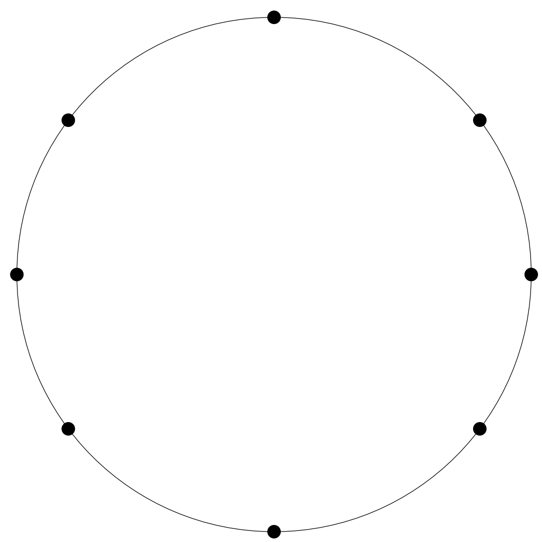
$$\kappa(R_+) \le k\, 2^k \frac{|r_{11}|}{|r_{kk}|}.$$

Very pessimistic estimate for $\|R_+\|^{-1}$ — but sharp.

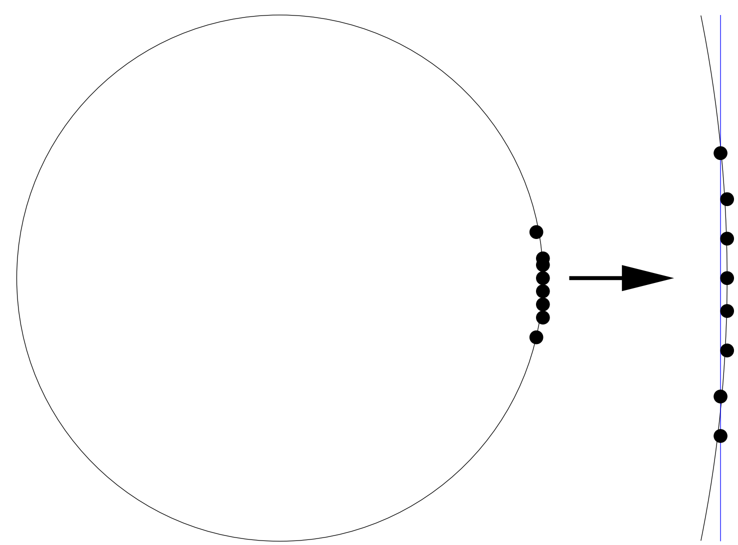
- Better if k is small $\Leftarrow$ many low degree polynomials in ideal.
- Growth rate of k depends on $\dim V_{\deg F}(\mathscr{I}_\Xi)$; related to Hilbert function.

More numerical considerations II

- What does ill–conditioned mean?
- Example with 8 points.



Well distributed, far from all lines.



Poor distribution — Small perturbation of line.

- Practically **any** variety $\mathscr{V}$ becomes “linear” for dense points.

A little bit of perturbation analysis

- Suppose that $\Xi \subset [-1,1]^d$ and $\widehat{\Xi} := \Xi + \Theta$, $\|\Theta\|_\infty < \varepsilon$, and that F is orthonormal. Then (componentwise)

$$\left| F\left(\widehat{\Xi}\right) - F(\Xi) \right| \le 2d \, (\deg F)^2 \sqrt{\binom{\deg F + d}{d}} \, \varepsilon =: C\,\varepsilon.$$

Markov inequality + simple estimates; "complexity quantities".

- Factors of *Householder matrix* $I - 2y\,y^T$ for column $\widehat{f}_j := \left[\widehat{f}_{jj}, \dots, \widehat{f}_{jn}\right]^T$ of $\widehat{F}^{(j)}(\Xi)$

$$y = \frac{\widehat{f}_j + \epsilon \,|\widehat{r}_{jj}|\; e_j}{2\,|\widehat{r}_{jj}| \left(|\widehat{r}_{jj}| + \left|\widehat{f}_{jj}\right|\right)}, \qquad \widehat{r}_{jj} = \pm \left\|\widehat{f}_j\right\|_2, \qquad \epsilon \in \{-1, 1\}.$$

Continuous as long as problem is well–conditioned. Error amplification: $\sim |r_{jj}|^{-2}$.

The modified algorithm

- Initialize

$$k \leftarrow 0, \qquad \Xi' \leftarrow \Xi, \qquad H \leftarrow \emptyset.$$

- `while` $\Xi' \neq \emptyset$
  - ▷ Choose basis P of $W_k(H)$.
  - ▷ Set $P' \leftarrow P - L_{\mathscr{P}_{k-1}} P$; [Back–substitution by means of Newton basis is stable]
  - ▷ Orthonormalize P': $(p, p') = \delta_{p,p'}$, $p, p' \in P'$.
  - ▷ Determine Ξ_k and decompose:

$$P' = F \cup Q, \qquad |Q(\Xi)| \leq \varepsilon, \qquad \operatorname{diag} F(\Xi_k) > \varepsilon.$$

  - ▷ Set $N_k \leftarrow F(\Xi_k)^{-1} F$.
  - ▷ Update

$$k \leftarrow k+1, \qquad \Xi' \leftarrow \Xi' \setminus \Xi_k, \qquad H \leftarrow H \cup Q.$$

And the approximate variety

- Algorithm stops for some $k = \min\{j : \mathscr{P} \subset \Pi_j\}$.
- For $0 \le l \le k+1$ set $H_l := \{h \in H : \deg h \le l\}$.

Theorem. *Each* H_l *is an H–basis for* $\langle H_l \rangle$, $l = 1, \dots, k$.

- H–basis for computations modulo ideal – dimension etc.
- Since $H_l \subseteq H_{l+1}$ the varieties $\mathscr{V}_l$ are of decreasing dimension with

$$\mathbb{R}^d = \mathscr{V}_0 \supseteq \mathscr{V}_1 \supseteq \cdots \supseteq \mathscr{V}_{k+1} = \Xi.$$

- Which one to pick? Tradeoff between simplicity and tightness!

Back to varieties

Back to the example

- For the simple case $\Xi = \{k\eta \; : \; k = 0, \dots, N\} \dots$
- Interpolation space ($N = 2M + \epsilon$)

$$\mathscr{P} = \mathrm{span} \left\{ \eta_1 x + \eta_2 y, \eta_1 x^2 + \eta_2 xy + \eta_1 y^2, \dots, \eta_1 x^N + \cdots + \eta_{1+\epsilon} y^N \right\}.$$

- Ideal bases:

$$\begin{aligned} H_0 &= \{0\} \\ H_1 = \cdots = H_N &= \left\{ \frac{\eta_2}{\|\eta\|_2} x - \frac{\eta_1}{\|\eta\|_2} y \right\} \\ \cdots = H_{N+2} = H_{N+1} &= \left\{ \frac{\eta_2}{\|\eta\|_2} x - \frac{\eta_1}{\|\eta\|_2} y, \frac{\eta_1 x^{N+1} + \cdots + \eta_\epsilon y^{N+1} + \cdots}{\|\dots\|_2} \right\} \end{aligned}$$

Conclusion

- Orthogonal projections and transformations are stable operations.
- Homogeneous H–bases can be used in an "orthogonal fashion".
- **Even better:** Use inner product $(f, g) = (f(D)g)(0)$ – RKHS!
- Determination of "approximate ideals" and "approximate varieties".
- **But:**
  - ▷ Still no efficient implementations.
  - ▷ Cannot compensate ill–conditioning.
- Nevertheless – it's worth a try.

Thank you for your attention!