Padé and Hermite-Padé approximants

Alin Bostan

Specfun, Inria

MPRI C-2-22
October 26, 2020

Part 2 from the last homework

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

(a) Show that the parity of all coefficients of P can be determined in $O(\mathsf{M}(N))$ bit ops.

(b) Show that P satisfies a linear differential equation of order 1 with polynomial coefficients.

(c) Determine a linear recurrence of order 2 satisfied by the sequence $(p_i)_i$.

(d) Give an algorithm that computes p_N in $O(\mathsf{M}_{\mathbb{Z}}(N \log N) \log N)$ bit ops.

Solution (a)

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

(a) Show that the parity of all coefficients of P can be determined in $O(\mathsf{M}(N))$ bit ops.

▷ n is even if and only if $n = 0$ in $\mathbb{K} := \mathbb{Z}/2\mathbb{Z}$.

▷ It is sufficient to compute $P_N(x) := (1 + x + x^2)^N$ in $\mathbb{K}[x]$

▷ DAC algorithm based on $P_N(x) = P_{\lfloor N/2 \rfloor}(x)^2 \cdot (1 + x + x^2)^{N \bmod 2}$ in $\mathbb{K}[x]$

▷ Cost recurrence: $\mathsf{C}(N) = \mathsf{C}(N/2) + \mathsf{M}(N/2) + O(N)$

▷ Conclusion: $\mathsf{C}(N) = \mathsf{M}(N) + O(N)$

Solution (b), (c)

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

(b) Show that P satisfies a linear differential equation of order 1 with polynomial coefficients.

(c) Determine a linear recurrence of order 2 satisfied by the sequence $(p_i)_i$.

▷ Logarithmic derivative: $\dfrac{P'(x)}{P(x)} = \dfrac{N\,(2\,x+1)}{x^2+x+1}$

▷ $P = \sum\limits_{i \geq 0} p_i x^i, \quad P' = \sum\limits_{i \geq 0} (i+1) p_{i+1} x^i, \quad [x^i](x^2+x+1)P' - N(2x+1)P = 0$

▷ $(i-1)p_{i-1} + ip_i + (i+1)p_{i+1} = 2Np_{i-1} + Np_i,$ for all $i \geq 0$

▷ The recurrence satisfied by the sequence (p_i) is

$$p_{i+1} = \frac{1}{i+1}\Big((N-i)p_i + (2N-i+1)p_{i-1}\Big) \quad \text{for} \quad i \geq 0.$$

Solution (d)

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

(d) Give an algorithm that computes p_N in $O(\mathsf{M}_{\mathbb{Z}}(N \log N) \log N)$ bit ops.

▷ The recurrence rewrites in matrix form: $F_i = \frac{1}{i+1} A_i F_{i-1}$, where

$$A_i = \begin{pmatrix} N - i & 2N - i + 1 \\ i + 1 & 0 \end{pmatrix} \quad \text{and} \quad F_i = \begin{pmatrix} p_{i+1} \\ p_i \end{pmatrix}.$$

▷ By unrolling it, we obtain the equality: $F_i = \dfrac{1}{(i+1)!} A(i) \cdots A(1) \begin{pmatrix} N \\ 1 \end{pmatrix}$.

▷ To compute p_N we determine F_N by binary splitting on the integer $f = (N+1)!$ and on the matrix $B = A(N) \cdots A(1)$, followed by the matrix-vector product $v = B \times \begin{pmatrix} N & 1 \end{pmatrix}^T$ and an (exact) division $\frac{1}{f} v$.

▷ The integer f, and the elements of B and v, have $O(N \log(N))$ bits.

▷ Cost: $O(\mathsf{M}_{\mathbb{Z}}(N \log(N)) \log(N))$ bit ops.

Bonus

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

Questions:

1. What is the total bitsize of P?
2. What is the bit complexity for computing P:
   - using the algorithm in (a)?
   - using the linear recurrence in (c)?

Bonus

Let $P = \sum_{i=0}^{2N} p_i x^i \in \mathbb{Z}[x]$ be the polynomial $P(x) = (1 + x + x^2)^N$.

Questions:

1. What is the total bitsize of P? $O(N^2)$
2. What is the bit complexity for computing P:
   - using the algorithm in (a)? $\tilde{O}(N^2)$
   - using the linear recurrence in (c)? $\tilde{O}(N^2)$

▷ Similar to

$$\sum_{k=0}^{N} \log \binom{N}{k} = (N+1)\log N! - 2\sum_{k=0}^{N} \log k! \sim \frac{N(N+1)}{2}$$

by Stirling's approximation $\log k! = k \log k - k + \frac{1}{2}\log k + \log \sqrt{2\pi} + o(1)$.

▷ Without Stirling, by using AGM inequality only:

$$p_0 \cdots p_{2N} \leq \left(\frac{p_0 + \cdots + p_{2N}}{2N+1}\right)^{2N+1} = \left(\frac{3^N}{2N+1}\right)^{2N+1} \implies \sum_i \log(p_i) = O(N^2).$$

Resultants and Euclid’s algorithm

Back to the Homework: computation with algebraic numbers

Let $A = \prod_i (x - \alpha_i)$ and $B = \prod_j (x - \beta_j)$ be polynomials of $\mathbb{K}[x]$. Then

$$A \oplus B := \prod_{i,j} (t - (\alpha_i + \beta_j)) \text{ is equal to } \mathsf{Res}_x(A(x), B(t - x)).$$

Proof: By Poisson's formula, $\mathsf{Res}\,(P, Q) \; = \; \mathsf{lc}(P)^{\deg Q} \cdot \prod_{P(\gamma)=0} Q(\gamma)$.

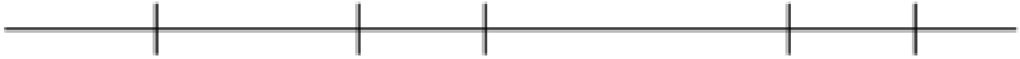
Thus, $\mathsf{Res}_x(A(x), B(t - x)) = \prod_i B(t - \alpha_i) = \prod_{i,j} (t - \alpha_i - \beta_j) = A \oplus B$.

▷ Algorithm and Complexity? If $\deg A, \deg B \leq d$, then $\deg(A \oplus B) \leq d^2$. Evaluation-interpolation on $d^2 + 1$ points $t_i \in \mathbb{K}$, and $d^2 + 1$ computations of $\mathsf{Res}_x(A(x), B(t_i - x))$, each in $O(d^\theta)$. Total: $O(d^{\theta+2})$ ops.

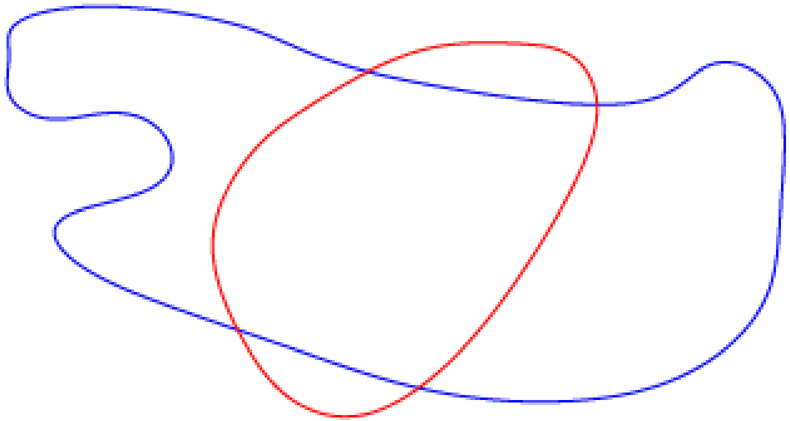
▷ Upcoming: each $\mathsf{Res}_x(A(x), B(t_i - x))$ in $O(d^2)$ [so, total in $O(d^4)$], and even in $\tilde{O}(d)$ [so, total in $\tilde{O}(d^3)$]. Not today: Total in $\tilde{O}(d^2)$ = quasi-optimal.

Systems of two equations and two unknowns

Geometrically, roots of a polynomial $f \in \mathbb{Q}[x]$ correspond to points on a line.



Roots of polynomials $A \in \mathbb{Q}[x, y]$ correspond to plane curves $A = 0$.

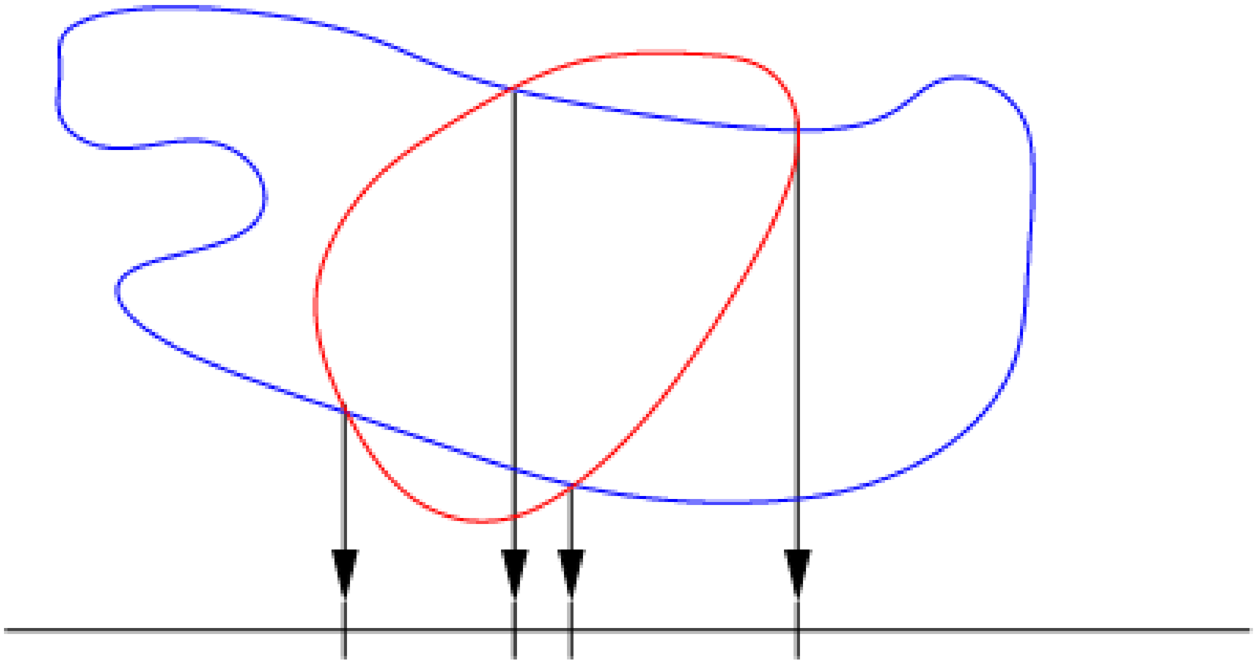


Let now A and B be in $\mathbb{Q}[x, y]$. Then:

- either the curves $A = 0$ and $B = 0$ have a common component,
- or they intersect in a finite number of points.

Application: Resultants compute projections

Theorem. Let $A = a_m y^m + \cdots$ and $B = b_n y^n + \cdots$ be polynomials in $\mathbb{Q}[x][y]$. The roots of $\mathsf{Res}_y(A, B) \in \mathbb{Q}[x]$ are either the abscissas of points in the intersection $A = B = 0$, or common roots of a_m and b_n.



Proof. Elimination property: $\mathsf{Res}_y(A, B) = UA + VB$, for $U, V \in \mathbb{Q}[x, y]$.

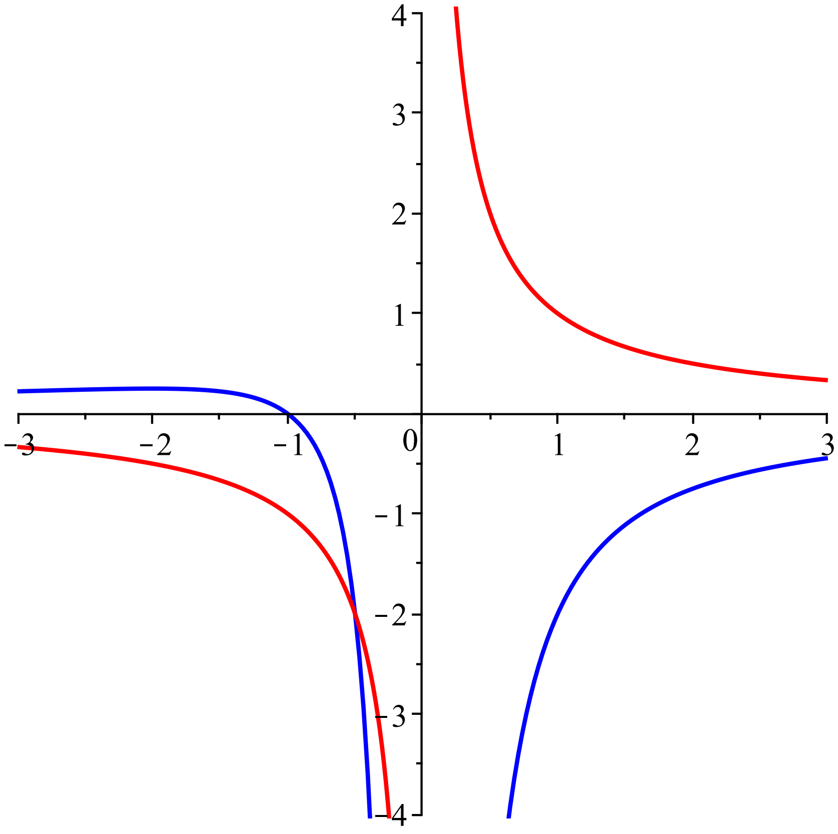
Thus $A(\alpha, \beta) = B(\alpha, \beta) = 0$ implies $\mathsf{Res}_y(A, B)(\alpha) = 0$

Conversely: if $\mathsf{Res}_y(A, B)(\alpha) = 0$ and $a_m(\alpha) \neq 0$, then $\mathsf{Res}_y(A(\alpha, y), B(\alpha, y))$ is equal to $\mathsf{Res}_y(A, B)(\alpha) = 0$, thus $\exists \beta \in \overline{\mathbb{Q}}$ with $A(\alpha, \beta) = B(\alpha, \beta) = 0$.

Application: Resultants compute projections

Graphically, the *degenerated* roots of the second case ("common roots of a_m and b_n") correspond to the presence of vertical asymptotes

▷ Example: taking $A = x^2y + x + 1$ and $B = xy - 1$, one has $\mathsf{Res}_y(A, B) = -x(2x + 1)$ (asymptote in $x = 0$, *true* solution in $x = -\frac{1}{2}$).



Application: implicitization of parametric curves

Task: Given a rational parametrization of a curve

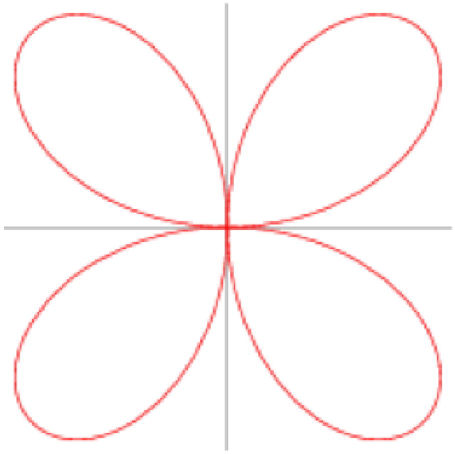
$$x = A(t), \quad y = B(t), \qquad A, B \in \mathbb{K}(t),$$

compute a non-trivial polynomial in x and y vanishing on the curve.

Recipe: take the resultant in t of numerators of $x - A(t)$ and $y - B(t)$.

Example: for the four-leaved clover (a.k.a. quadrifolium) given by

$$x = \frac{4t(1-t^2)^2}{(1+t^2)^3}, \quad y = \frac{8t^2(1-t^2)}{(1+t^2)^3},$$



$$\mathrm{Res}_t((1+t^2)^3x - 4t(1-t^2)^2, (1+t^2)^3y - 8t^2(1-t^2)) = 2^{24}\left((x^2+y^2)^3 - 4x^2y^2\right).$$

Euclidean-type algorithm for the resultant

- If $A = QB + R$, and $R \neq 0$, then (by Poisson's formula)

$$\mathsf{Res}\,(A, B) = (-1)^{\deg A \deg B}\, \mathsf{lc}(B)^{\deg A - \deg R}\, \mathsf{Res}\,(B, R).$$

- If B is constant, then $\quad \mathsf{Res}\,(A, B) = B^{\deg A}$.

If $(R_0, \dots, R_{\ell-1}, R_\ell = \gcd(A, B), 0)$ is the remainder sequence produced by the Euclidean algorithm for $R_0 = A$ and $R_1 = B$, then

- either $\deg R_\ell \neq 0$, in which case $\mathsf{Res}\,(A, B) = 0$,
- else: $\mathsf{Res}\,(A, B) = R_\ell^{\deg R_{\ell-1}} \displaystyle\prod_{i=0}^{\ell-2} (-1)^{\deg R_i \deg R_{i+1}}\, \mathsf{lc}(R_{i+1})^{\deg R_i - \deg R_{i+2}}$.

▷ This leads to a $O(n^2)$ algorithm for $\mathsf{Res}\,(A, B)$, where $\deg(A), \deg(B) \leq n$.

▷ Divide-and-conquer $O(\mathsf{M}(n) \log n)$ algorithms exist but require extra-work.

Euclidean algorithm

Euclid(A, B)

```
Input  $A$ and $B$ in $\mathbb{K}[x]$.
Output  A gcd $G$ of $A$ and $B$.
  1. $R_0 := A$; $R_1 := B$; $i := 1$.
  2. While $R_i$ is non-zero, do:
         $R_{i+1} := R_{i-1} \bmod R_i$
         $i := i + 1$.
  3. Return $R_{i-1}$.
```

▷ Termination: $\deg(B) > \deg(R_2) > \deg(R_3) > \cdots$

▷ Correctness: $\gcd(A, B) = \gcd(B, A \bmod B)$

▷ Quadratic complexity: $O\big(\deg(A)\deg(B)\big)$ operations in $\mathbb{K}$

Euclidean-type algorithm for the resultant

Res(A, B)

Input A and B in $\mathbb{K}[x]$.

Output The resultant $\mathrm{Res}(A, B)$.

```
1. $R_0 := A$; $R_1 := B$; $r_1 := 1$; $i := 1$.
2. While $\deg(R_i) > 0$, do:
       $R_{i+1} := R_{i-1} \bmod R_i$
       $r_i := (-1)^{\deg(R_{i-1})\deg(R_i)} \cdot \mathsf{lc}(R_i)^{\deg(R_{i-1})-\deg(R_{i+1})} \cdot r_{i-1}$
       $i := i + 1$.
3. If $R_i \neq 0$, then return $r_{i-1} \cdot R_i^{\deg(R_{i-1})}$. Else return 0.
```

▷ Termination: $\deg(B) > \deg(R_2) > \deg(R_3) > \cdots$

▷ Correctness: In step 3, R_i is a constant. If $R_i = 0$, then $R_{i-1} = \gcd(A, B)$ has degree > 0, thus Res $(A, B) = 0$. If $R_i \neq 0$, then $R_i = \gcd(A, B)$.

▷ Quadratic complexity: $O\big(\deg(A)\deg(B)\big)$ operations in $\mathbb{K}$

Padé approximants

Rational reconstruction

Let $\mathbb{K}$ be a field, $A \in \mathbb{K}[x]$ of degree $n > 0$ and $B \in \mathbb{K}[x]$ of degree $< n$. For $k \in \{1, \dots, n\}$, the k-th rational reconstruction of B modulo A is the problem of finding a pair $(R, V) \in \mathbb{K}[x]^2$ satisfying:

$$(\mathsf{RR}_k) \quad \gcd(V, A) = 1, \quad \frac{R}{V} \equiv B \bmod A, \quad \deg(R) < k, \quad \deg(V) \leq n - k.$$

▷ Particular cases: Padé approximation for $A = x^n$, and Cauchy (rational) interpolation for $A = \prod_i (x - u_i)$ where $u_1, \dots, u_n \in \mathbb{K}$ mutually distinct.

▷ Remarks:

- If $k = n$, then $(R, V) = (B, 1)$ is a solution of (RR_k)
- If $k < n$, (RR_k) may have no solution! E.g., $A = x^3, B = x^2 + 1, k = 2$
- If (RR_k) admits a solution (R, V), then R/V is unique:

  if $(R_1, V_1) \in \mathbb{K}[x]^2$ another solution, then $\underbrace{A}_{\deg n}$ divides $\underbrace{R_1 V - V_1 R}_{\deg < n}$

A simpler problem

Let $\mathbb{K}$ be a field, $A \in \mathbb{K}[x]$ of degre $n > 0$ and $B \in \mathbb{K}[x]$ of degre $< n$.
For $k \in \{1, \ldots, n\}$, the k-th simplified rational reconstruction of B modulo A is the problem of finding a pair $(R, V) \in \mathbb{K}[x]^2$ satisfying:

$$(\mathsf{SRR}_k) \qquad R \equiv VB \bmod A, \quad \deg(R) < k, \quad \deg(V) \leq n - k.$$

▷ Remarks:

- (SRR_k) always admits a non-trivial solution!
  Indeed, (SRR_k) is equivalent to a homogeneous linear algebra problem with n equations and $k + (n - k + 1) = n + 1$ unknowns $\qquad O(n^\theta)$

- If (SRR_k) admits a solution (R, V), then R/V is unique:
  if $(R_1, V_1) \in \mathbb{K}[x]^2$ another solution, then $\underbrace{A}_{\deg n}$ divides $\underbrace{R_1V - V_1R}_{\deg < n}$

Teasers

▷ One can solve (SRR_k) using the Extended Euclidean Algorithm $O(n^2)$

▷ One can deduce a decision/computation procedure for (RR_k) $O(n^2)$

▷ Intuition of the link with the gcd:

$$R \equiv VB \bmod A \quad \text{iff} \quad \exists U, \text{ s.t. } UA + VB = R$$

▷ The case $k = 1$:

- If $\gcd(A, B) = 1$ then $R = UA + VB$ for $R = 1$ and $\deg V < \deg A = n$
- If $\gcd(A, B) \neq 1$ then $R = VB \bmod A$ for $R = 0$ and $V = \operatorname{lcm}(A, B)/B$

Extended Euclidean Algorithm

EEA(A, B)

```
Input  $A$ and $B$ in $\mathbb{K}[x]$.

Output  A gcd $G$ of $A$ and $B$, and cofactors $U$ and $V$.

  1. $R_0 := A$; $U_0 := 1$; $V_0 := 0$; $R_1 := B$; $U_1 := 0$; $V_1 := 1$; $i := 1$.
  2. While $R_i$ is non-zero, do:
     (a) $(Q_i, R_{i+1}) := \mathrm{QuotRem}(R_{i-1}, R_i)$          $\# R_{i-1} = Q_i R_i + R_{i+1}$
     (b) $U_{i+1} := U_{i-1} - Q_i U_i$; $V_{i+1} := V_{i-1} - Q_i V_i$
     (c) $i := i + 1$
  3. Return $\big(R_{i-1}, U_{i-1}, V_{i-1}\big)$.
```

▷ Correctness: $R_i = U_i A + V_i B$ (by induction):

$$R_{i+1} = R_{i-1} - Q_i R_i = U_{i-1} A + V_{i-1} B - Q_i (U_i A + V_i B) = U_{i+1} A + V_{i+1} B$$

▷ Quadratic complexity: $O\big(\deg(A)\deg(B)\big)$ operations in $\mathbb{K}$

Extended Euclidean Algorithm: properties

▷ Matrix reformulation:

$$\begin{pmatrix} U_i & V_i \\ U_{i+1} & V_{i+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -Q_i \end{pmatrix} \times \begin{pmatrix} U_{i-1} & V_{i-1} \\ U_i & V_i \end{pmatrix}$$

▷ Consequences:

(1) $U_i V_{i+1} - V_i U_{i+1} = (-1)^i$

(2) $\gcd(U_i, V_i) = 1$

(3) $\gcd(R_i, V_i) = \gcd(A, V_i)$

▷ Degrees of cofactors: If $n_0 = n = \deg(A) > n_1 = \deg(B) > n_2 > \ldots > n_\ell$ are the degrees of R_i in $\mathsf{EEA}(A, B)$, then $\deg(V_i) = n - n_{i-1}$ for all $i \geq 1$

Proof by complete induction: since $\deg(V_{k-1}) = n - n_{k-2} < n - n_{k-1} = \deg(V_k)$,
$\deg(V_{k+1}) = \deg(Q_k V_k) = \deg(Q_k) + \deg(V_k) = (n_{k-1} - n_k) + (n - n_{k-1}) = n - n_k$

Main result for Rational Reconstruction

Theorem: Let $A \in \mathbb{K}[x]$ have degree $n > 0$ and let $B \in \mathbb{K}[x]$ have degree $< n$. Let $k \in \{1, 2, \ldots, n\}$ and let R_j, U_j, V_j be the jth row in $\mathsf{EEA}(A, B)$, where j is minimal s.t. $\deg(R_j) < k$. Then:

(1) (R_j, V_j) is a nontrivial solution of (SRR_k).
If moreover $\gcd(R_j, V_j) = 1$, then (R_j, V_j) also solves (RR_k).

(2) If (RR_k) admits a solution, then $\gcd(R_j, V_j) = 1$.
Precisely, if $R/V \in \mathbb{K}(x)$ is an irreducible form of the solution of (RR_k), then $\exists \alpha \in \mathbb{K} \setminus \{0\}$ s.t. $R = \alpha R_j$ and $V = \alpha V_j$.

In summary: (RR_k) admits a solution iff $\gcd(R_j, V_j) = 1$ iff $\gcd(A, V_j) = 1$.

Proof of (1)

Let us prove that (R_j, V_j) is a nontrivial solution of (SRR_k):

- $R_j = U_j A + V_j B \equiv V_j B \mod A$
- $\deg R_j < k$ (by assumption)
- $\deg V_j = n - \deg R_{j-1} \leq n - k$ (by minimality of j s.t. $\deg(R_j) < k$)

If moreover $1 = \gcd(R_j, V_j) = \gcd(A, V_j)$, then (R_j, V_j) also solves (RR_k).

Proof of (2)

Assume (R, V) is a solution of (RR_k). Then:

(a) $R = UA + VB$ for a certain $U \in \mathbb{K}[x]$

(b) $U_j V = U V_j$: otherwise, the system $\begin{pmatrix} U_j & V_j \\ U & V \end{pmatrix} \times \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} R_j \\ R \end{pmatrix}$ is of Cramer-type, and thus $A = \begin{vmatrix} R_j & V_j \\ R & V \end{vmatrix} \Big/ \begin{vmatrix} U_j & V_j \\ U & V \end{vmatrix}$, satisfies:

$$\begin{aligned}\deg(A) \leq \deg(R_j V - R V_j) &\leq \max\{k-1+\deg(V), \deg(R)+n-\deg(R_{j-1})\} \\ &\leq \max\{k-1+(n-k), (k-1)+n-k)\} = n-1.\end{aligned}$$

(c) V_j divides $U_j V$, and thus V. Write $V = \alpha V_j$, for $\alpha \in \mathbb{K}[x]$. Then, $U = \alpha U_j$ and $R = UA + VB = \alpha(U_j A + V_j B) = \alpha R_j$.

(d) $\alpha \in \mathbb{K} \setminus \{0\}$, since $\gcd(R, V) = 1$

Algorithm for Rational Reconstruction

$\mathsf{RR}(A, B, k)$

Input A, B in $\mathbb{K}[x]$ with $\deg B < \deg A = n$ and $k \in \{1, \ldots, n\}$

Output A solution (R, V) of (RR_k), or FAIL if no solution

```
1. $R_0 := A$; $V_0 := 0$; $R_1 := B$; $V_1 := 1$; $i := 1$.
2. While $\deg R_i \geq k$, do:
   (a) $(Q_i, R_{i+1}) := \mathrm{QuotRem}(R_{i-1}, R_i)$        $\# R_{i-1} = Q_i R_i + R_{i+1}$
   (b) $V_{i+1} := V_{i-1} - Q_i V_i$
   (c) $i := i + 1$
3. If $\gcd(A, V_i) = 1$ then return $(R_i, V_i)$; else return FAIL.
```

- Correctness: Previous result.
- Quadratic complexity: $O\big(\deg(A)\deg(B)\big) = O(n^2)$ operations in $\mathbb{K}$
- There exist quasi-linear time algorithms $O(\mathsf{M}(n) \log n)$

Algorithm for Padé approximation

Pade(B, n, k): for computing an approximant of type $(k-1, n-k)$

```
Input  $B$ in $\mathbb{K}[x]$ with $\deg B < n$ and $k \in \{1, \ldots, n\}$
Output  $(R, V)$ s.t. $R/V = B \bmod x^n$, $\deg R < k, \deg V \leq n-k$, or FAIL
  1. $R_0 := x^n$; $V_0 := 0$; $R_1 := B$; $V_1 := 1$; $i := 1$.
  2. While $\deg R_i \geq k$, do:
     (a) $(Q_i, R_{i+1}) := \mathrm{QuotRem}(R_{i-1}, R_i)$          $\# R_{i-1} = Q_i R_i + R_{i+1}$
     (b) $V_{i+1} := V_{i-1} - Q_i V_i$
     (c) $i := i + 1$
  3. If $V_i(0) \neq 0$ then return $(R_i, V_i)$; else return FAIL.
```

- ▷ Quadratic complexity: $O(n^2)$ operations in $\mathbb{K}$
- ▷ There exist quasi-linear time algorithms $O(\mathsf{M}(n) \log n)$

First exercise for next time (2/11/2020)

Show, using the previous algorithm, that there is no Padé approximant of type $(1,1)$ for $1+x^2$, i.e. no pair (R,V) of polynomials of degree at most 1 such that $V(0) \neq 0$ and $\dfrac{R}{V} = 1 + x^2 \bmod x^3$.

Algorithm for Cauchy interpolation

Cauchy$(\mathbf{u}, \mathbf{v}, k)$

```
**Input** $u_1, \ldots, u_n \in \mathbb{K}$ mutually distinct, $v_1, \ldots, v_n \in \mathbb{K}$, and $k \in \{1, \ldots, n\}$
**Output** $\frac{R}{V}$ s.t. $\frac{R}{V}(u_i) = v_i$ for all $i$, $\deg R < k, \deg V \leq n - k$, or FAIL
  1. $A := \prod_i (x - u_i)$ and $B$ s.t. $B(u_i) = v_i$ for all $i$, $\deg B < n$.
  2. $R_0 := A$; $V_0 := 0$; $R_1 := B$; $V_1 := 1$; $i := 1$.
  3. While $\deg R_i \geq k$, do:
     (a) $(Q_i, R_{i+1}) := \mathrm{QuotRem}(R_{i-1}, R_i)$          $\# R_{i-1} = Q_i R_i + R_{i+1}$
     (b) $V_{i+1} := V_{i-1} - Q_i V_i$
     (c) $i := i + 1$
  4. If $V_i(u_j) \neq 0$ for all $j$, then return $(R_i, V_i)$; else return FAIL.
```

- ▷ Quadratic complexity: $O(n^2)$ operations in $\mathbb{K}$
- ▷ There exist quasi-linear time algorithms $O(\mathsf{M}(n) \log n)$

Guessing: what's the next term of the sequence?

- $1, \quad 1, \quad 1, \quad 1, \quad 1$
- $1, \quad 1, \quad 2, \quad 3, \quad 5$
- $1, \quad 1, \quad 2, \quad 5, \quad 14$
- $1, \quad 2, \quad 9, \quad 54, \quad 378$
- $1, \quad 2, \quad 16, \quad 192, \quad 2816$
- $1, \quad 3, \quad 30, \quad 420, \quad 6930$

Guessing: what's the next term of the sequence?

• 1, 1, 1, 1, 1	1
• 1, 1, 2, 3, 5	8
• 1, 1, 2, 5, 14	42
• 1, 2, 9, 54, 378	2916
• 1, 2, 16, 192, 2816	46592
• 1, 3, 30, 420, 6930	126126

Guessing: what's the next term of the sequence?

- $1, \quad 1, \quad 1, \quad 1, \quad 1$ $\qquad 1/(1-t)$
- $1, \quad 1, \quad 2, \quad 3, \quad 5$ $\qquad 1/(1-t-t^2)$
- $1, \quad 1, \quad 2, \quad 5, \quad 14$ $\qquad (1-\sqrt{1-4t})/(2t)$
- $1, \quad 2, \quad 9, \quad 54, \quad 378$ $\qquad 27\,t^2T^2 + (1-18\,t)\,T + 16\,t - 1$
- $1, \quad 2, \quad 16, \quad 192, \quad 2816$ $\qquad 64\,t^2\,T^3 + 16\,t\,T^2 + (1-72\,t)\,T + 54\,t - 1$
- $1, \quad 3, \quad 30, \quad 420, \quad 6930$ $\qquad \left(27\,t^2 - t\right)y''(t) + (54\,t - 2)\,y'(t) + 6\,y(t)$

▷ Automated guessing: algorithmic computation of these equations

Berlekamp-Massey algorithm

—guessing linear recurrences with constant coefficients—

Linear recurrences with constant coefficients

Def. $(a_n)_{n\geq 0}$ is a linearly recurrent sequence with constant coefficients (l.r.s.c.c., or C-recursive) if $\exists f_0, \dots, f_d \in \mathbb{K}$ not all zero, s.t.

$$f_d a_{n+d} + \cdots + f_0 a_n = 0, \quad \text{for all} \quad n \geq 0.$$

▷ $f(x) = f_d x^d + \cdots + f_0$ is called a characteristic polynomial of $(a_n)_{n\geq 0}$.

▷ The minimal polynomial of $(a_n)_{n\geq 0}$, denoted $\mathsf{MinPol}(a_n)$ is the monic, minimal degree, characteristic polynomial of $(a_n)_{n\geq 0}$.

▷ Computing $\mathsf{MinPol}(a_n)$ is equivalent to solving the Padé approximation pb:

$$\frac{R}{V} \equiv A \bmod x^{2N}, \quad x \nmid V, \ \deg(R) < N, \ \deg(V) \leq N \ \text{ and } \ \gcd(R, V) = 1,$$

where $\deg \mathsf{MinPol}(a_n) \leq N$ and $A = a_0 + a_1 x + a_2 x^2 + \cdots + a_{2N-1} x^{2N-1}$.

Recall: duality lemma

Duality lemma (link between l.r.s.c.c. and rational functions)
Let $A(x) = \sum_{n\geq 0} a_n x^n \in \mathbb{K}[[x]]$ be the generating function of $(a_n)_{n\geq 0}$.
The following assertions are equivalent:

(i) (a_n) is a l.r.s.c.c., having P as characteristic polynomial of degree d.

(ii) $A(x)$ is rational, of the form $A = Q/\mathsf{rev}_d(P)$ for some $Q \in \mathbb{K}[x]_{<d}$, where $\mathsf{rev}_d(P) = P(\frac{1}{x})x^d$.

Moreover, if P is the minimal polynomial of $(a_n)_{n\geq 0}$, then

$$d = \max\{1 + \deg(Q), \deg(\mathsf{rev}_d(P))\} \quad \text{and} \quad \gcd(Q, \mathsf{rev}_d(P)) = 1.$$

▷ E.g., the generating function of the Fibonacci sequence $(F_n)_{n\geq 0}$ given by $F_0 = F_1 = 1, F_{n+2} = F_{n+1} + F_n$ is $1/(1 - x - x^2)$. Here $P = x^2 - x - 1$, $Q = 1$.

Berlekamp-Massey algorithm

Input A bound $N \in \mathbb{N}$ on the degree of the minimal polynomial of $(a_n)_{n\geq 0}$ and the first $2N$ terms $a_0, \dots, a_{2N-1} \in \mathbb{K}$.

Output the minimal polynomial of $(a_n)_{n\geq 0}$.

(1) $A = a_0 + a_1 x + \cdots + a_{2N-1} x^{2N-1}$.

(2) Compute the solution $(R, V) \in \mathbb{K}[x]^2$ of $\mathsf{Pade}(A, 2N, N)$ s.t. $V(0) = 1$.

(3) $d = \max\{1 + \deg(R), \deg(V)\}$. Return $\mathsf{rev}_d(V) = V(1/x)x^d$.

- ▷ Quadratic complexity: $O(N^2)$ operations in $\mathbb{K}$
- ▷ There exist quasi-linear time algorithms $O(\mathsf{M}(N)\log N)$

Berlekamp-Massey algorithm, a variant

```
Input  A bound $N \in \mathbb{N}$ on the degree of the minimal polynomial of $(a_n)_{n \geq 0}$
       and the first $2N$ terms $a_0, \dots, a_{2N-1} \in \mathbb{K}$.

Output  the minimal polynomial of $(a_n)_{n \geq 0}$.

 1. $R_0 := x^{2N}$; $V_0 := 0$; $R_1 := a_{2N-1} + \cdots + a_0 x^{2N-1}$; $V_1 := 1$; $i := 1$.
 2. While $\deg R_i \geq N$, do:
    (a) $(Q_i, R_{i+1}) := \mathrm{QuotRem}(R_{i-1}, R_i)$                $\# R_{i-1} = Q_i R_i + R_{i+1}$
    (b) $V_{i+1} := V_{i-1} - Q_i V_i$
    (c) $i := i + 1$
 3. Return $V_i / \mathrm{lc}(V_i)$.
```

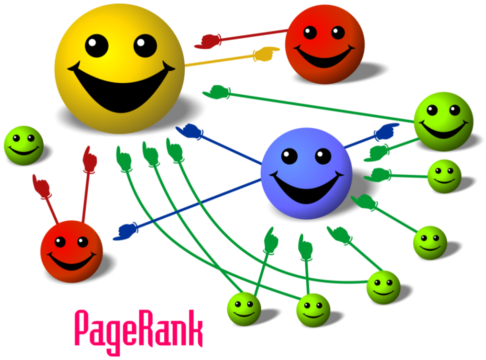
- ▷ Quadratic complexity: $O(N^2)$ operations in $\mathbb{K}$
- ▷ There exist quasi-linear time algorithms $O(\mathsf{M}(N) \log N)$

Wiedemann's algorithm

—solving sparse matrices—

Sparse matrices

Def. A matrix $M \in \mathcal{M}_n(\mathbb{K})$ is called sparse if it has $s \ll n^2$ nonzero entries.



▷ Typically: $s \approx n$

▷ Cost of matrix-vector product Mv, for $v \in \mathbb{K}^n$, is $c(M) = O(s)$ ops. in $\mathbb{K}$.

▷ Applications: web ranking, integer factorization, discrete logarithm

Algorithms for sparse matrices

Def. A matrix $M \in \mathcal{M}_n(\mathbb{K})$ is called sparse if it has $s \ll n^2$ nonzero entries.

▷ Algorithmic questions for sparse matrices

- solving linear systems $My = b$ ($b \in \mathbb{K}^n$)
- computing determinant $\det(M)$, rank $\mathrm{rk}(M)$, etc.
- computing characteristic polynomial $\chi_M(x)$, minimal polynomial $\mu_M(x)$

▷ Answers: probabilistic algorithms of complexity $O(ns + n^2)$, or $\tilde{O}(ns + n^2)$

▷ Roughly speaking, "$\omega_{\mathsf{sparse}} = 2$"

Wiedemann's algorithm for solving sparse linear systems

▷ solving $My = b$ reduces to computing an annihilating polynomial $\mu_{\mathbf{v}}$:

if $\mu_{\mathbf{v}} = \sum_{j=0}^{d} m_j x^j$ is the minimal polynomial of the vector sequence $\mathbf{v} = (M^i b)_{i \geq 0}$, i.e. the minimal degree polynomial s.t. $\sum_{j=0}^{d} m_j M^j b = 0$, then

(1) $\mu_{\mathbf{v}}$ divides μ_M (since μ_M annihilates $\mathbf{v}$)

(2) thus $\mu_{\mathbf{v}}$ also divides χ_M, and in particular $d \leq n$

(3) if M is invertible, then $m_0 \neq 0$ (since $\chi_M(0) = \det(M) \neq 0$)

(4) sol. $y = M^{-1} b = -\left(\frac{m_1}{m_0} b + \cdots + \frac{m_d}{m_0} M^{d-1} b\right)$ in $O(d(c(M) + n)) = O(ns + n^2)$

▷ computing $\mu_{\mathbf{v}}$ reduces (probabilistically) to computing $\mathsf{MinPol}((u^T \cdot M^i \cdot b)_i)$ for a randomly chosen vector $u \in \mathbb{K}^n$:

(1) $(a_i)_i = (u^T \cdot M^i \cdot b)_i$ is a l.r.s.c.c.

(2) its minimal polynomial $\mu_a = \mathsf{MinPol}((a_i)_i)$ has degree at most n

(3) μ_a divides $\mu_{\mathbf{v}}$; in addition, they coincide with good probability

(4) μ_a can be computed from $a_0, \ldots, a_{2n-1}$ in $O(n^2)$ by Berlekamp-Massey algo

(5) $a_0, \ldots, a_{2n-1}$ can be computed in $O(n(c(M) + n)) = O(ns + n^2)$

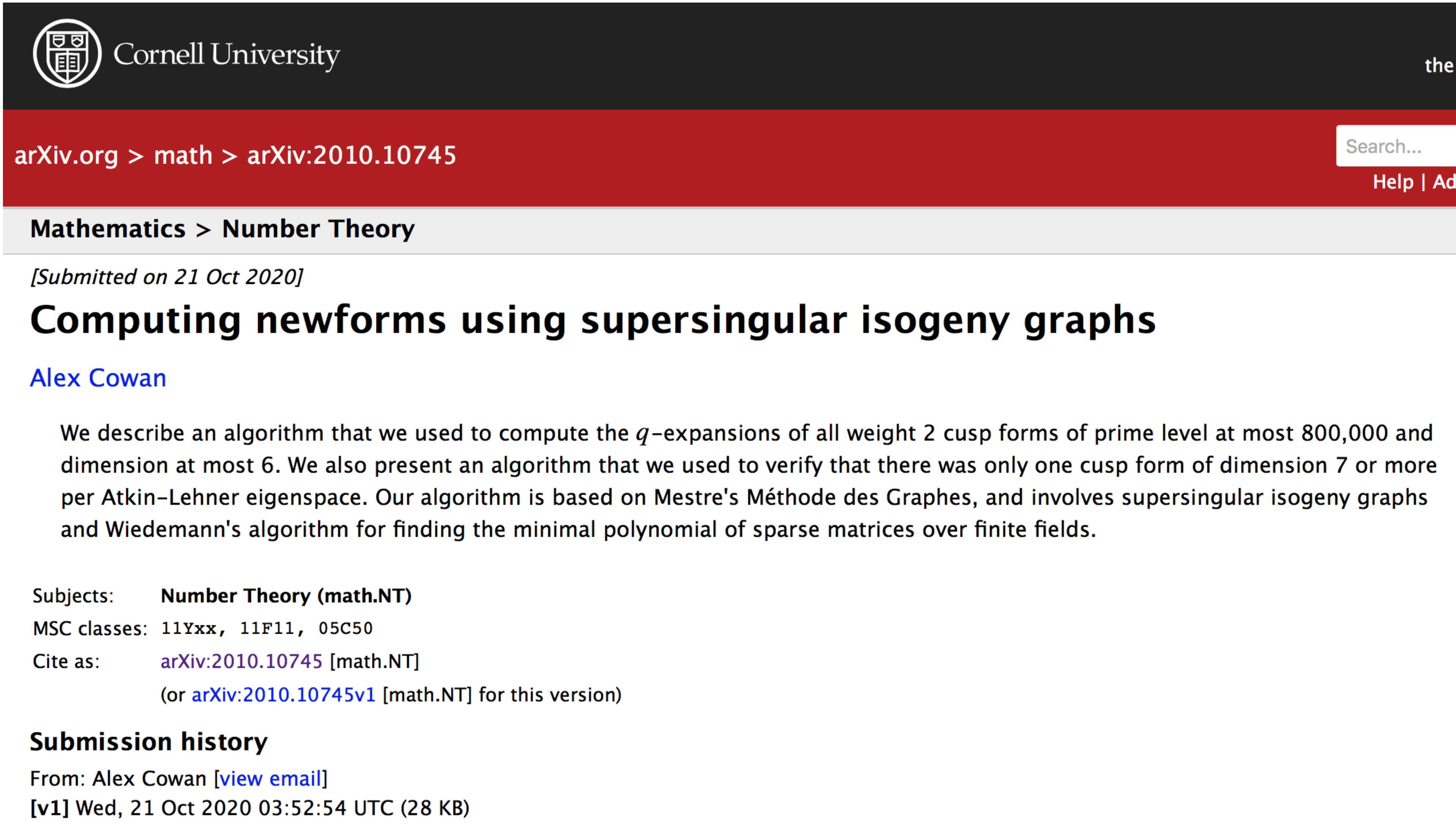
Wiedemann's algorithm for solving sparse linear systems

```
Input  A sparse invertible matrix $M \in \mathcal{M}_n(\mathbb{K})$ and $b \in \mathbb{K}^n$.
Output  The solution $y \in \mathbb{K}^n$ of the linear system $My = b$.

(1) Compute $\mu_{\mathbf{v}} = \mathsf{MinPol}((M^i b)_{i \geq 0})$, as below.
(2) Compute $h = -\frac{\mu_{\mathbf{v}} - \mu_{\mathbf{v}}(0)}{\mu_{\mathbf{v}}(0)x} \in \mathbb{K}[x]$ and return $y = h(M)b$.
```

```
Input  A sparse matrix $M \in \mathcal{M}_n(\mathbb{K})$ and $b \in \mathbb{K}^n$.
Output  The minimal polynomial of the vector sequence $\mathbf{v} = (M^i b)_{i \geq 0}$.

(1) If $b = 0$, then return 1.
(2) Choose $U \subseteq \mathbb{K}$ a finite subset of cardinal $\geq 2n$.
(3) Choose $u \in U^n$ uniformly at random; compute $a_i = u^T v_i$ for $0 \leq i < 2n$.
(4) Compute the minimal polynomial $\mu_a$ of the l.r.s.c.c. $(a_i)_i$.
(5) If $\mu_a(M)b = 0$ in $\mathbb{K}^n$, return $\mu_a$; else go back to (3).
```

▷ Randomized algorithm; expected complexity $O(nc(M) + n^2)$ if $|\mathbb{K}| \geq 2n$.

▷ Probability$(\mu_a = \mu_{\mathbf{v}}) \geq 1 - \frac{n}{|U|}$, thus 2 expected iterations if $|\mathbb{K}| \geq 2n$.


Cornell University
the
arXiv.org > math > arXiv:2010.10745
Search...
Help | Ad
Mathematics > Number Theory

[Submitted on 21 Oct 2020]

Computing newforms using supersingular isogeny graphs

Alex Cowan

We describe an algorithm that we used to compute the q-expansions of all weight 2 cusp forms of prime level at most 800,000 and dimension at most 6. We also present an algorithm that we used to verify that there was only one cusp form of dimension 7 or more per Atkin-Lehner eigenspace. Our algorithm is based on Mestre's Méthode des Graphes, and involves supersingular isogeny graphs and Wiedemann's algorithm for finding the minimal polynomial of sparse matrices over finite fields.

Subjects:	**Number Theory (math.NT)**
MSC classes:	11Yxx, 11F11, 05C50
Cite as:	arXiv:2010.10745 [math.NT]
	(or arXiv:2010.10745v1 [math.NT] for this version)

Submission history

From: Alex Cowan [view email]
[v1] Wed, 21 Oct 2020 03:52:54 UTC (28 KB)

Probability of success in Wiedemann's algorithm

Let $\mathcal{P}$ be $\text{Probability}\big(\text{MinPol}((u^T M^i b)_{i\geq 0}) = f\big)$, where $f = \text{MinPol}\,\big((M^i b)_{i\geq 0}\big)$

(1) The map $\psi : \mathbb{K}^n \to \mathbb{A} = \mathbb{K}[x]/(f)$ defined by $\psi(u) := \sum_{i=0}^{n-1} (u^T \cdot M^i \cdot b) x^i$ is $\mathbb{K}$-linear, surjective and such that

$$f \text{ is the minpoly of } (u^T \cdot M^i \cdot b)_{i\geq 0} \iff \psi(u) \text{ is invertible in } \mathbb{A}.$$

(2) $R(y_1, \ldots, y_n) := \text{Res}_x(y_1\psi(e_1) + \cdots + y_n\psi(e_n), f(x)) \in \mathbb{K}[y_1, \ldots, y_n] \setminus \{0\}$ has total degree at most $d := \deg(f)$ and for all $u = (u_1, \ldots, u_n)^T \in \mathbb{K}^n$,

$$\psi(u) \text{ is invertible in } \mathbb{A} \iff R(u_1, \ldots, u_n) \neq 0.$$

(3) R admits at most $\deg(R) \cdot |U|^{n-1} \leq d \cdot |U|^{n-1}$ roots U^n.

(4) Probability that an element in U^n is a root of R is at most $d/|U|$.

(5) $\mathcal{P} \geq 1 - \dfrac{d}{|U|}$.

A second exercise for next time (2/11/2020)

Let $\mathbb{K} = \mathbb{F}_5$ be the finite field with 5 elements, let $M \in \mathcal{M}_3(\mathbb{K})$ and $b \in \mathbb{K}^3$ be

$$M = \begin{pmatrix} 1 & 4 & 4 \\ 4 & 0 & 3 \\ 1 & 2 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$$

We want to find $y \in \mathbb{K}^3$ such that $My = b$, by using Wiedemann's algorithm.

1. For the choice $u = (1,0,0)^T$, show that the algorithm computes the sequence $(3,0,4,2,3,0,\ldots)$, then its minimal polynomial $x^2 + 2x + 2$, and that it eventually rejects this choice of u.

2. Apply the algorithm for the choice $u = (1,2,0)^T$, and deduce that the minimal polynomial of $(M^i b)_{i \geq 0}$ equals $x^3 + 3x + 1$.

3. Determine the solution y by using this minimal polynomial.

Hermite-Padé approximants

Definition of Hermite-Padé approximants

Definition: Given a column vector $\mathbf{F} = (f_1, \ldots, f_n)^T \in \mathbb{K}[[x]]^n$ and an n-tuple $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{N}^n$, a Hermite-Padé approximant of type $\mathbf{d}$ for $\mathbf{F}$ is a row vector $\mathbf{P} = (P_1, \ldots, P_n) \in \mathbb{K}[x]^n$, $(\mathbf{P} \neq 0)$, such that:

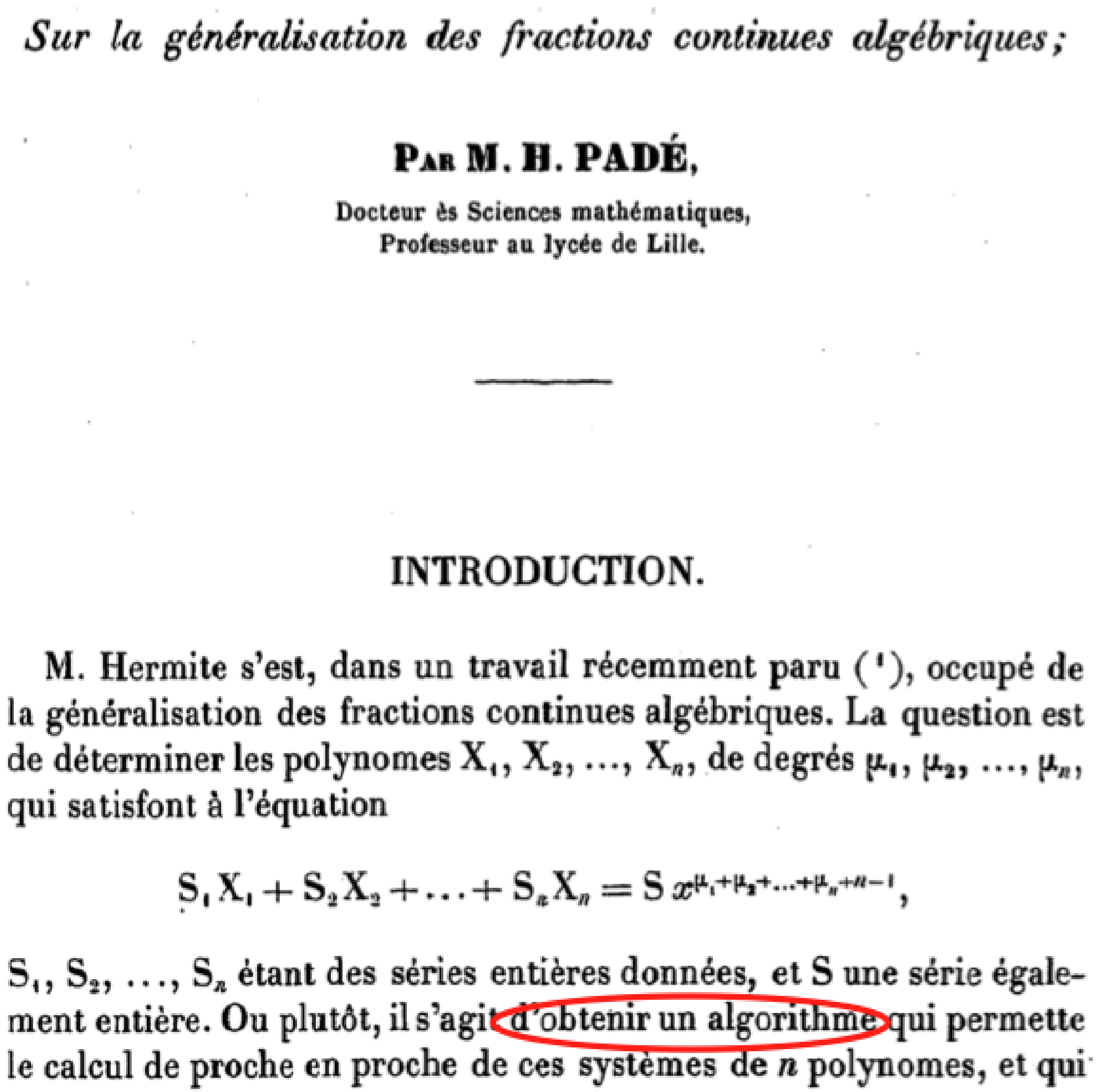
(1) $\mathbf{P} \cdot \mathbf{F} = P_1 f_1 + \cdots + P_n f_n = O(x^\sigma)$ with $\sigma = \sum_i (d_i + 1) - 1$,

(2) $\deg(P_i) \leq d_i$ for all i.

σ is called the order of the approximant $\mathbf{P}$.

▷ Very useful concept in number theory (irrationality/transcendence):

- [Hermite, 1873]: e is transcendent.
- [Lindemann, 1882]: π is transcendent; so does e^α for any $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$.
- [Apéry, 1978; Beukers, 1981]: $\zeta(3) = \sum_{n \geq 1} \frac{1}{n^3}$ is irrational.
- [Rivoal, 2000]: there exist infinite values of k such that $\zeta(2k+1) \notin \mathbb{Q}$.



Sur la généralisation des fractions continues algébriques;

Par M. H. PADÉ,

Docteur ès Sciences mathématiques,
Professeur au lycée de Lille.

INTRODUCTION.

M. Hermite s'est, dans un travail récemment paru (1), occupé de la généralisation des fractions continues algébriques. La question est de déterminer les polynomes $X_1, X_2, \dots, X_n$, de degrés $\mu_1, \mu_2, \dots, \mu_n$, qui satisfont à l'équation

$$S_1 X_1 + S_2 X_2 + \dots + S_n X_n = S\, x^{\mu_1+\mu_2+\dots+\mu_n+n-1},$$

$S_1, S_2, \dots, S_n$ étant des séries entières données, et S une série également entière. Ou plutôt, il s'agit d'obtenir un algorithme qui permette le calcul de proche en proche de ces systèmes de n polynomes, et qui

[Padé, 1894]