Characteristic polynomials and subspaces of matrices

Jean-Guillaume Dumas and John Sheekey

Université de Grenoble
Laboratoire Jean Kuntzmann
Applied Mathematics and Computer Science Department

University College Dublin
Claude Shannon Institute
Discrete Mathematics, Coding, Cryptography and Information Security

Subspaces of matrices

- $\mathbb{F}_q$ is a finite field
- $\mathbb{V}$ is a vector space of finite dimension *n* over $\mathbb{F}_q$
- Consider subspaces $\mathcal{M}$ of $\mathrm{End}_{\mathbb{F}_q}(\mathbb{V})$ in which each non-zero element has a prescribed rank
- $\mathcal{M}$ is a (n^2, d) code
  - Search for large dimensions d
  - *[Gow et. al]* investigate maximum dimension subspaces

Type	Any			GF(2)	GF(2)	Skew-symmetric			
Dimension	$m \times n$	$m \times n$	$m \times n$	4×5	5×5	$n \times n$	$n \times n$	$n \times n$	$3n \times 3n$
Constraint	rank $\geq r$	rank $= r$	rank $= r$	rank $= 3$	rank $= 4$	rank $\geq 2r$	rank $= 2r$	rank $= 2$	rank $= 2n$
Max dim.	$\leq n(m-r+1)$	$\geq n$	$\leq m+n-r$	6	6	$\geq n(n-2r+1)/2$	$\leq 2n-2r-1$	$\leq n-1$	$\geq 3n$

Finite Semi-fields

- A finite non-associative ring $\mathbb{D}$ where nonzero elements are closed under multiplication is called a **presemifield**
- If $\mathbb{D}$ has an identity element it is called **semifield**

- *[L.E. Dickson, 1906]*
- *[A.A. Albert, 50s]*
- *[D.E. Knuth, 1965]* for projective semifield planes

- *[Kantor 2006, Dempwollf 2008, Rúa et al. 2009]* ...
  - Representation as subspaces of invertible matrices
  - Classification of semifields of order $81=3^4$, $64=2^6$, ...

Equivalence testing

- Space equivalence
  - Classification, exhaustive search, etc.
  - $\mathcal{M}$ eq. $\mathcal{S}$ iff $\mathcal{S} = U^{-1}\,\mathcal{M}\,V$, for some invertible U and V

- Some ideas to reduce the search when $Id \in \mathcal{M}$

1. Reduction to similarity
   - $Id = U^{-1} A V \Rightarrow V = A^{-1} U$
   - ie. $S_i = U^{-1} M_j V$ if and only if $S_i = U^{-1} (M_j A^{-1}) U$

   $\Rightarrow \mathcal{M}$ eq. $\mathcal{S}$ iff $\exists A \in \mathcal{M}$ such that $\mathcal{M}\,A^{-1}$ sim. $\mathcal{S}$

2. Less admissible characteristic polynomials
   - Chevaley-Warning theory ⇨ prescribed coefficients
   - $A+x\cdot Id \in \mathcal{M}$ ⇨ no eigen value in the field for $A \neq \lambda\, Id$

   ⇨No linear factor in the characteristic polynomial

Computing the characteristic polynomial and matrix normal forms

- Motivations
  - Subspaces of matrices, semifields
  - But also: Isomorphism of graphs, certified eigenvalues …
- Computations
  - Similarity via matrix normal forms
  - Frobenius normal form and characteristic polynomial
  - Krylov iterations
  - Reductions to matrix multiplication
- Finite semi-fields of order $243=3^5$?

Companion matrix

If $P(X) = X^n + \sum_{i=0}^{n-1} p_i X^i$ then

$$Companion(P(X)) = \begin{bmatrix} 0 & 0 & \dots & 0 & -p_0 \\ 1 & 0 & \dots & 0 & -p_1 \\ & \ddots & \ddots & & \vdots \\ & & & 1 & 0 & -p_{n-2} \\ & & & & 1 & -p_{n-1} \end{bmatrix}$$

Charpoly (Companion(P)) = P

Minpoly (Companion(P)) = P

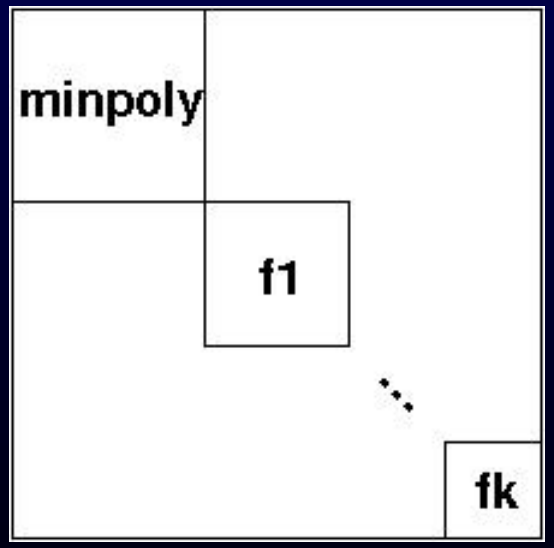
Frobenius normal form

- Similarity

⇨ Tested via a change of basis to a normal form

  - Gauß-Jordan normal form
  - Frobenius (rational canonical form)
    - Block diagonal companion matrix

    - $f_k \mid f_{k-1} \mid \ldots \mid f_1 \mid f_0 = \text{Minpoly}$
    - $f_k \cdot f_{k-1} \cdot \ldots \cdot f_1 \cdot f_0 = \text{Charpoly}$
    - $\text{Minpoly} \mid \text{Charpoly} \mid \text{Minpoly}^n$



Space equivalence via similarity

$|\mathbf{GL(n, q)}| \approx \mathbf{q^{n^2}}$

- Equivalent
  - $\exists$ U, V $\in$ GL(n,q)
    $\forall S_i \in \mathcal{S}, \forall M_j \in \mathcal{M}$
    $S_i = U^{-1} \cdot M_j \cdot V$

Complexity bound reduced
from $\mathbf{O\left(n^3 q^{2n^2}\right)}$ to $\mathbf{O\left(n^6 + n^3 q^{2n}\right)}$

$|\mathbf{Centralizer}| \approx \mathbf{q^n}$

- Equivalent
  - $\exists A \in \mathcal{M}$, Frob$\{\mathcal{S}\}$ = Frob$\{\mathcal{M}\ A^{-1}\}$
  - S_* ~~> Frob$_*$
    $\exists A, M_a \in \mathcal{M}$, $M_a A^{-1}$ ~~> Frob$_*$

$\Rightarrow S_* = K_a^{-1} \cdot (M_a A^{-1}) \cdot K_a$

- $\exists Y_* \in$ Centralizer(S_*)

$U = Y_* \cdot K_a^{-1}$ and $V = A^{-1} \cdot K_a \cdot Y_*^{-1}$

$\forall S_i \in \mathcal{S}, \forall M_j \in \mathcal{M}$
$S_i = U^{-1} M_j V$

Algebraic complexity model

- Counting arithmetic operations
- E.g. Matrix multiplication
  - Classic $2n^3 - n^2$
  - *[Strassen 1969]* $7n^{2.807} + o(n^{2.807})$
  - *[Winograd 1971]* $6n^{2.807} + o(n^{2.807})$
  - ...
  - *[Coppersmith Winograd 1990]* $O(n^{2.376})$

  $O(n^{\omega})$, where ω denotes an admissible exponent
- Reductions to matrix multiplication
  - ⇨ Better complexity
  - ✚ Better efficiency in practice
    - Block versions optimize memory hierarchy usage

Examples of Matrix multiplication reductions

- Triangular system solving with n×n matrix right hand side
  - TRSM(n) = n^3 or $1/(2^{\omega-1}-2) \cdot MM(n)$
- TRMM(n) = n^3 or $1/(2^{\omega-1}-2) \cdot MM(n)$
- Inverse of well-behaved matrices *[Strassen 1969]*
  - INVERSE(n) = $2n^3$ or $3 \cdot 2^{\omega} /(2^{\omega}-4) /(2^{\omega}-2) \cdot MM(n)$
  - INVT(n) = $1/3\ n^3$ or $4/(2^{\omega}-4) /(2^{\omega}-2) \cdot MM(n)$
- LQUP of any matrix *[Ibarra-Moran-Hui 1982]*
  - LQUP(n) = $2/3\ n^3$ or $2^{\omega} /(2^{\omega}-4) /(2^{\omega}-2) \cdot MM(n)$
  - Rank
  - Determinant
- Charpoly, Frobenius form ?

Characteristic polynomial Computations, pre-Strassen

- *[Leverrier 1840]*
  - trace of powers of A, and Newton's formula
  - improved/rediscovered by Souriau, Faddeev, Frame and Csanky
  - $O(n^4)$ operations using matrix multiplication
  - Still suited for parallel computations
- *[Danilevskii 1937]*
  - elementary row/column operations
  - $O(n^3)$
- *[Hessenberg 1942]*
  - transformation to quasi-upper triangular and determinant expansion formula
  - $O(n^3)$

Characteristic polynomial Computations, post-Strassen

- *[Preparata & Sarwate 1978]*
  - Update Csanky with fast matrix multiplication
  - $O(n^{\omega+1})$
- *[Keller-Gehrig 1985]*
  - Using a Krylov basis
  - $O(n^{\omega} \log n)$
- *[Keller-Gehrig 1985]*
  - Danilevskii block operations
  - $O(n^{\omega})$ BUT only valid with well-behaved matrices

Krylov iteration

- Degree d Krylov matrix of one vector v
  - $K = [\, v \mid Av \mid A^2v \mid \ldots \mid A^{d-1}v \,]$
- Krylov property: d maximal, K full rank

  ⇨ $A \cdot K = [\, Av \mid A^2v \mid \ldots \mid A^dv \,] = K \cdot C = K \cdot$

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & * \\ 1 & 0 & \ldots & 0 & * \\ & \ddots & \ddots & & \vdots \\ & & & 1 & 0 & * \\ & & & & 1 & * \end{bmatrix}$$

  ⇨ C is the companion matrix of the minimal polynomial of A,v

  ☺ If v is chosen randomly and the field is sufficiently large $\text{minpoly}_{A,v} = \text{minpoly}_A$ with high probability

  👍 As minpoly_A is annihilating the sequence of projections we always have $\text{minpoly}_{A,v} \mid \text{minpoly}_A$

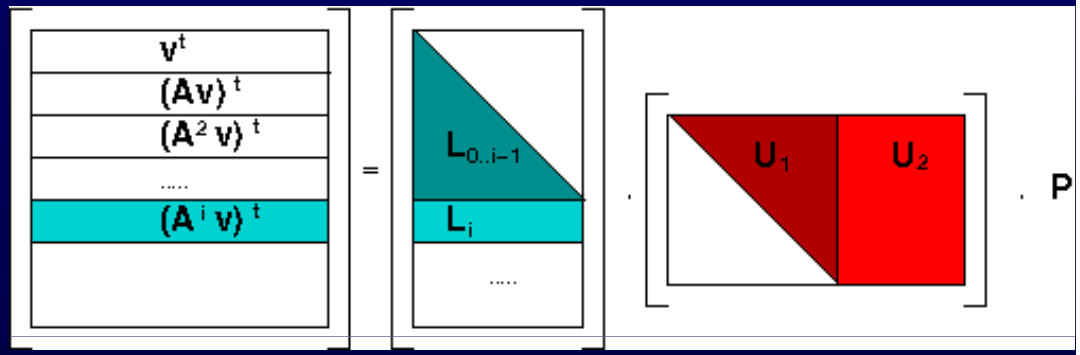
  - e.g. suppose K square, inv., minpoly=charpoly ⇨ $C = K^{-1} A K$

Minpoly $\leq$ Krylov+LUP+TRSM

[D., Pernet, Wan 2005]

1. QLUP factorisation of the Krylov matrix

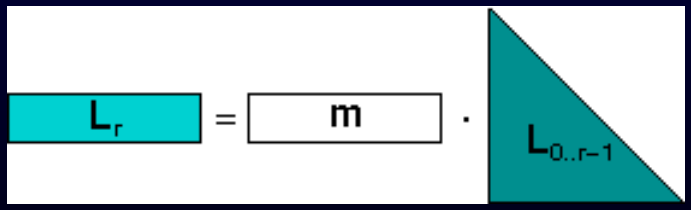
$\Rightarrow \left(A^i v\right)^t = L_i \cdot U \cdot P$



2. Cayley-Hamilton $\quad \text{Minpoly}_A(A) = 0 \Leftrightarrow A^r = -\sum m_i A^i$

1. + 2. $\quad (A^r v)^t = L_r \cdot U \cdot P = -\sum m_i \left(A^i v\right)^t = -\sum m_i (L_i \cdot U \cdot P)$

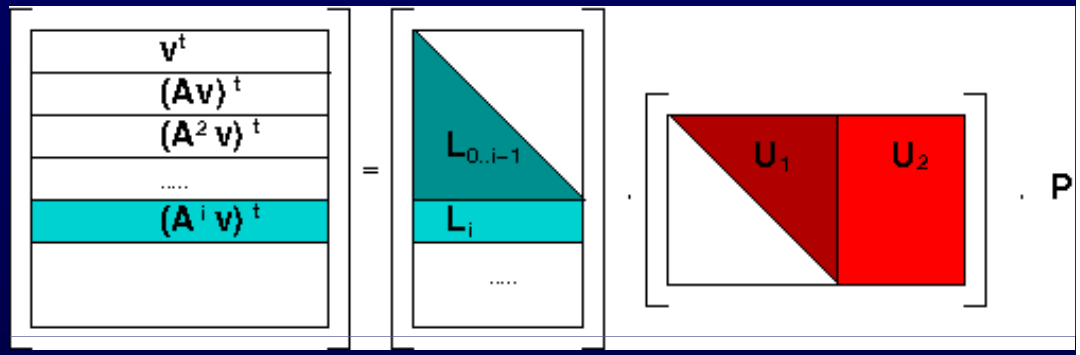
$$L_r U P P^{-1} \begin{bmatrix} U_1^{-1} \\ 0 \end{bmatrix} = L_r \cdot Id =$$

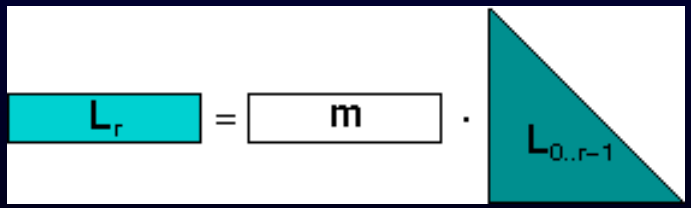


$\Rightarrow$ $\text{Minpoly}_{A,v}$ solves $\quad L_r = m \cdot L_{0..r-1}$

Minpoly $\leq$ Krylov+LUP+TRSM

[D., Pernet, Wan 2005]





LUKrylov algorithm: two problems

1. Krylov space is iterative: $2n^3$
   - 👍 *[Keller-Gehrig 85]*
     - $A, A^2, A^4, A^8, \ldots, A^{2\wedge\log(n)}$ in only log(n) matrix multiplication
     - $A^2 \cdot [v, Av] = [A^2v, A^3v]$
     - $A^4 \cdot [v, Av, A^2v, A^3v] = [A^4v, A^5v, A^6v, A^7v]$ …
     - … full Krylov iteration in $O(n^\omega \log(n))$
     - 😐 in practice log(n) matrix multiplication

2. Charpoly = Minpoly + Charpoly(Schur complement)
   - – Charpoly = $O(\sum n^2 \cdot k_i + k_i^2 n)$ or $O(\sum n^{\omega-1}k_i\log(k_i)+k_i^{\omega-1}n\log(n))$
   - ☺ With $\sum k_i = n$ and $\sum k_i^2 \leq n^2$ the latter gives $O(n^3)$

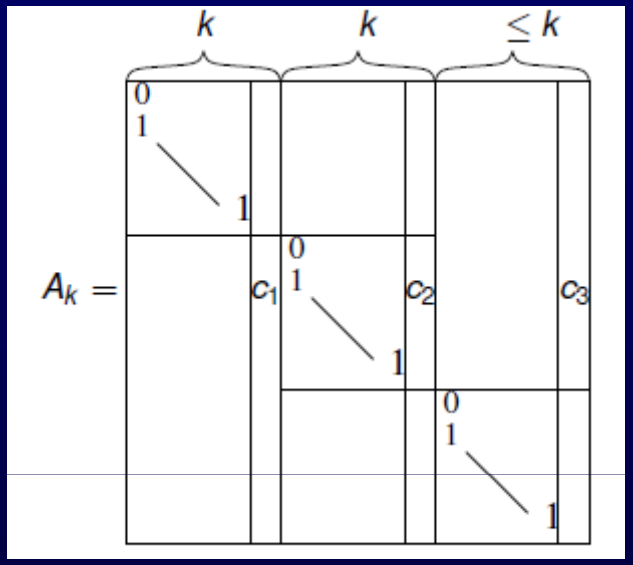
   - ✚ Frobenius form can be recovered along the way
   - ☹ But not $O(n^\omega \log(n))$ even with fast matrix multiplication and Keller-Gehrig's trick …

Simultaneously compute the blocks

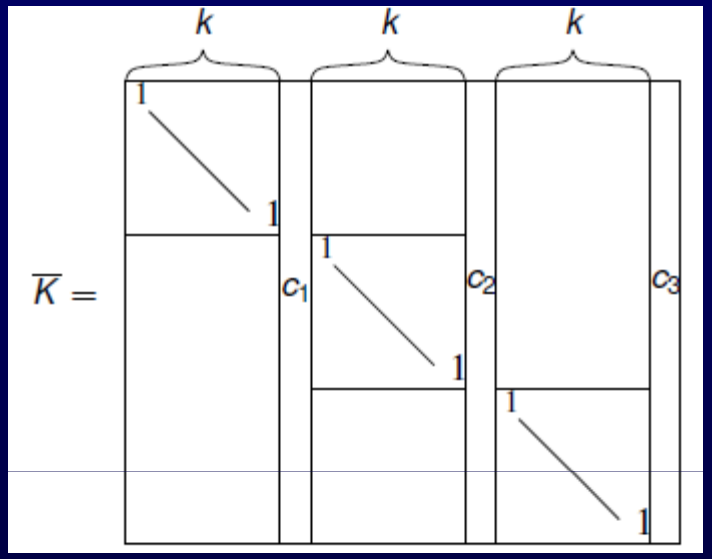
- Krylov matrix of several vectors v_i
  - $K = [\ v_1 \mid \dots \mid A^{k1-1}v_1 \mid v_2 \mid \dots \mid A^{k2-1}v_2 \mid \dots \mid v_l \mid \dots \mid A^{kl-1}v_l\ ]$
  - *[Eberly 2000]* Finds several blocks in the Frobenius form plus the change of basis, but also in either $O(n^3)$ or $O(n^\omega \log(n))$
- *[Pernet-Storjohann 2007]*
  - Start with $A_0 = A = [\ Ae_1 \mid Ae_2 \mid \dots \mid Ae_n\ ]$
  - Expand it to $K' = [e_1 \mid Ae_1 \mid e_2 \mid Ae_2 \mid \dots \mid e_n \mid Ae_n\ ]$
  - Find K_1, the first n independent columns
  - $A_0 \cdot K_1 = K_1 \cdot A_1 = K_1 \cdot [e_2 \mid * \mid e_4 \mid * \mid \dots\ ]$
  - …

  Iterate while reordering the columns to get increasingly large ordered identity parts
  - End by Frobenius = $A_d = K_d^{-1} \dots K_1^{-1}\ A\ K_1 \dots K_d$

From k-shifted form to (k+1)-shifted form

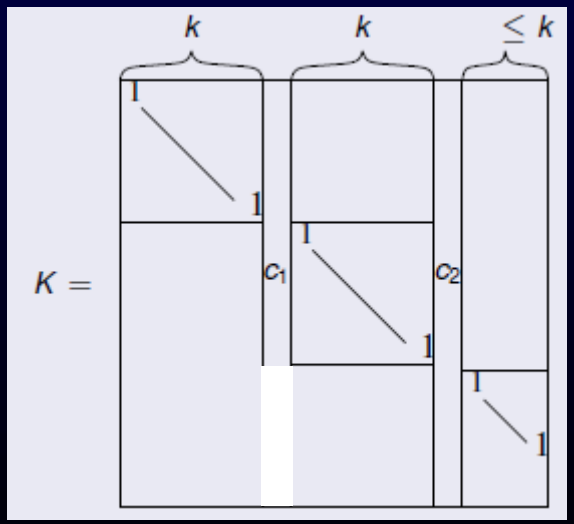


⇨ build $n \times (n+k)$



⇨ select first n independent columns

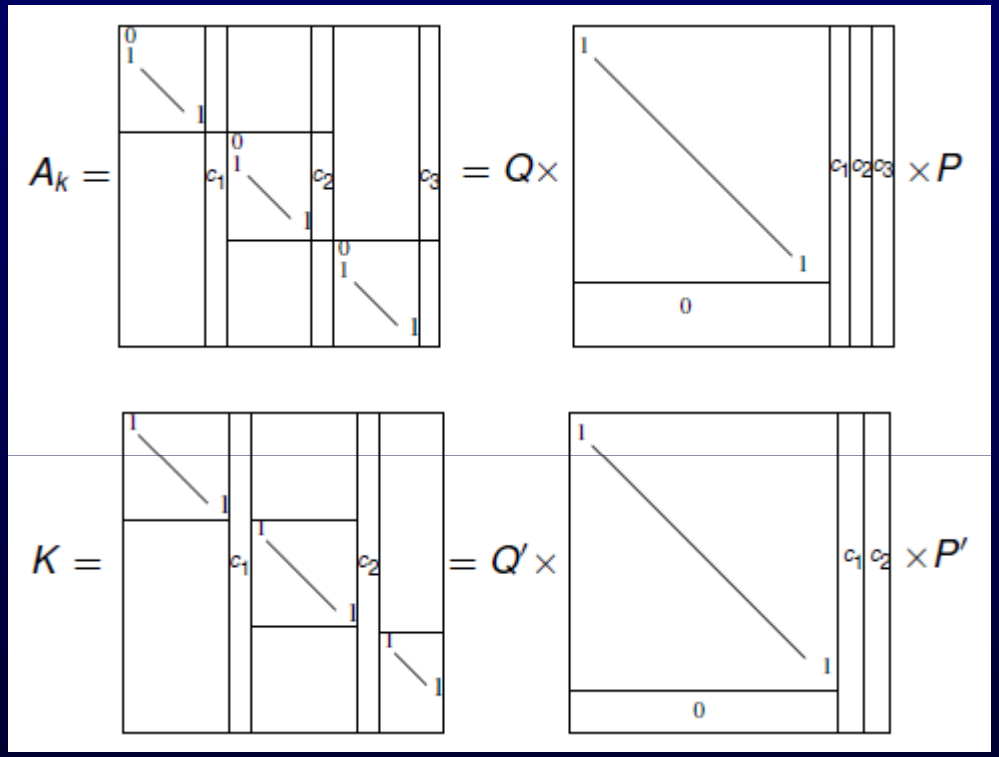
⇨ LQUP



- If $\#\mathbb{F}_q > n^2$, w. h. p.,

⇨ $A_{k+1} = K^{-1} A_k K$ is in (k+1)-shifted normal form

Using fast rectangular matrix multiplication



Id	A	×	Id	C	=	Id	C + AD
O	B		O	D		O	BD

- $n \times (n/k)$ by $(n/k) \times (n/k)$

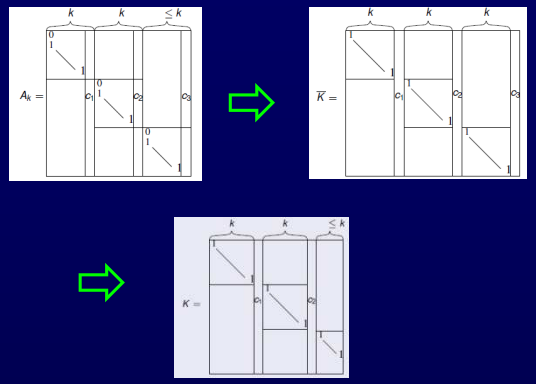
👍 Multiply k blocks of size (n/k)

⇨ $O(k (n/k)^{\omega})$

Overall complexity

[Pernet-Storjohann 2007]

- Rank profile n×(n/k)
  - derived from LQUP
  - $O(n(n/k)^{\omega-1}) = O(k (n/k)^{\omega})$



- Similarity transformation n×(n/k)
  - Parenthesizing
  - $O(k (n/k)^{\omega})$

$$K^{-1}AK = Q'^{T}\left(\begin{bmatrix} I & * \\ 0 & * \end{bmatrix}\left(P'^{T}Q\left(\begin{bmatrix} I & * \\ 0 & * \end{bmatrix}\left(PQ'\begin{bmatrix} I & * \\ 0 & * \end{bmatrix}\right)\right)\right)\right)P'$$

- Overall complexity bound
  - summing for each iteration

$$\sum_{k=1}^{n} k\left(\frac{n}{k}\right)^{\omega} = n^{\omega}\sum_{k=1}^{n}\left(\frac{1}{k}\right)^{\omega-1} = O(n^{\omega})$$

Blocking for Efficiency

Athlon 2200, 1.8 GHz, 2Gb

n	magma-2.11	LU-Krylov	Pernet-Storjohann
100	0.010s	0.005s	0.006s
300	0.830s	0.294s	0.105s
500	3.810s	1.316s	0.387s
800	15.64s	4.663s	1.387s
1000	29.96s	10.21s	2.755s
1500	102.1s	33.36s	7.696s
2000	238.0s	79.13s	17.91s
3000	802.0s	258.4s	61.09s
5000	3793s	1177s	273.4s
7500	MT	4209s	991.4s
10 000	MT	8847s	2080s

- Dominant factors of complexity bounds
  - LUKrylov $\approx 2n^3+2/3n^\omega \approx 4.33\ n^3$
  - Pernet-Storjohann $\approx ((6+2/3)\zeta(\omega-1)-6)n^\omega$ close to $4.96\ n^\omega$

About Probabilistic methods

- Monte-Carlo (always fast, probably correct)
- Less than 1/q to be wrong
  - Examples: $\text{minpoly}_{A,v}=\text{minpoly}_A$
  - Solution: divisibilty ensures lcm will converge with $1/q^k$
- Las Vegas (always correct, probably fast)
  - Examples: charpoly from LUKrylov algorithm
  - Divisibility ensures that poly is correct if degree is n
  - Solution: start again when check detects failure
- Frobenius
  - Preconditioning requires $\#\mathbb{F}_q > n^2$
  - Solution: select vectors from an extension field

Perspectives

- Frobenius $\leq$ Matrix Multiplication *[Pernet-Storjohann 2007]*
- Change of basis with an extra log(log(n)) factor
  - Application to semifields classification on small matrices …
- Sparse matrices ?
  - Rank, Det, Solve, Minpoly in $O(n^2)$ *[Wiedemann 86]*
  - Charpoly
    - Best algorithm $O(n^{2.5}\log^2(n)\log(\log(n))$ *[Villard 2000]*
    - Heuristic $O(n^{2.5})$ *[D.-Pernet-Saunders 2009]*
- Arbitrary precision Integer matrices ?
  - Coefficients growth ⇨ naïve methods exponential in n … still
  - Determinant $O(n^{\omega}\log^a(n))$ *[Storjohann 2005]*
  - Charpoly $O(n^{2.7}\log^a(n))$ *[Kaltofen-Villard 2004]*

Semi fields of order $243=3^5$ …

- Specialized Packed matrix routines
  - *[D. 2008, Boothby-Bradshaw 2010]*
  - Among 3^{20} = 3 486 784 401 matrices, 38 267 664 are invertible with 1 prescribed column and restricted charpoly
  - ☺ 2856s on 8 processors
- 26 × 38 267 664 = 994 959 264 Frobenius forms
  - ☺ Degree 5 with no linear factors ⇨ minpoly=charpoly
  - ⇨ Simple Krylov iteration with 1 vector, w.h.p yields Frobenius
  - 😐 Estimation 22 CPU days, Memory is the bottleneck …
- 994 959 264 × the number of equivalent classes $\langle I,F,A_i \rangle$: Comparisons
  - ⇨ Full equivalence testing: 243 (|centralizer|) tests for each comparison
  - ? Some pre-filtering might still be necessary …
- Then append the remaining 2 admissible matrices one at a time
  - ☺ Generation of adequate matrices: estimation 3 hours
  - ? Compute the inequivalent classes $\langle I,F,A_3,A_4,A_5 \rangle$ …