Computation of the Matrix Exponential by Generalized Polar Decomposition

Bari, 3 Dicembre 2003

Antonella Zanna
University of Bergen, Norway

email: anto@ii.uib.no
http://www.ii.uib.no/~anto

Example of Lie-group methods

- Order-four Gauss–Legendre:

$$
\begin{aligned}
F_1 &= \tfrac{1}{4}A_1 + (\tfrac{1}{4} - \tfrac{\sqrt{3}}{6})A_2 + (\tfrac{5}{144} - \tfrac{\sqrt{3}}{48})[A_1, A_2],\\
A_1 &= hA\big(t_n + c_1 h,\ \mathrm{expm}(F_1)Y_n\big),\\
F_2 &= (\tfrac{1}{4} + \tfrac{\sqrt{3}}{6})A_1 + \tfrac{1}{4}A_2 - (\tfrac{5}{144} + \tfrac{\sqrt{3}}{48})[A_1, A_2],\\
A_2 &= hA\big(t_n + c_2 h,\ \mathrm{expm}(F_2)Y_n\big),\\
\Theta &= \tfrac{1}{2}(A_1 + A_2) - \tfrac{\sqrt{3}}{12}[A_1, A_2],\\
Y_{n+1} &= \mathrm{expm}(\Theta)Y_n,
\end{aligned}
$$

where $c_i = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$, $i = 1, 2$.

- Order-four Gauss–Lobatto:

$$
\begin{aligned}
F_1 &= \boldsymbol{O},\\
A_1 &= hA(t_n, Y_n),\\
F_2 &= \tfrac{5}{24}A_1 + \tfrac{1}{3}A_2 - \tfrac{1}{24}A_3\\
&\quad - (\tfrac{11}{480}[A_1, A_2] + \tfrac{5}{1152}[A_1, A_3] + \tfrac{1}{144}[A_2, A_3]),\\
A_2 &= hA\big(t_n + \tfrac{1}{2}h,\ \mathrm{expm}(F_2)Y_n\big),\\
F_3 &= \tfrac{1}{6}A_1 + \tfrac{2}{3}A_2 + \tfrac{1}{6}A_3 - (\tfrac{1}{15}[A_1, A_2] + \tfrac{1}{60}[A_1, A_3] + \tfrac{1}{15}[A_2, A_3]),\\
A_3 &= hA\big(t_n + h,\ \mathrm{expm}(F_3)Y_n\big),\\
\Theta &= \tfrac{1}{6}A_1 + \tfrac{2}{3}A_2 + \tfrac{1}{6}A_3 - (\tfrac{1}{15}[A_1, A_2] + \tfrac{1}{60}[A_1, A_3] + \tfrac{1}{15}[A_2, A_3]),\\
Y_{n+1} &= \mathrm{expm}(\Theta)Y_n.
\end{aligned}
$$

Relaxing the collocation conditions, it is possible to obtain explicit methods.

- An explicit order-three scheme:

$$
\begin{aligned}
A_1 &= hA(t_n, Y_n),\\
A_2 &= hA\big(t_n + \tfrac{1}{2}h,\ \mathrm{expm}(A_1)Y_n\big),\\
A_3 &= hA\big(t_n + h,\ \mathrm{expm}(-A_1 + 2A_2)Y_n\big),\\
\Theta &= \tfrac{1}{6}A_1 + \tfrac{2}{3}A_2 + \tfrac{1}{6}A_3 - [A_1 - A_3, \tfrac{1}{15}A_2 + \tfrac{1}{60}A_3],\\
Y_{n+1} &= \mathrm{expm}(\Theta)Y_n.
\end{aligned}
$$

Motivations

- Integration methods using exponentials in GI need **fast** algorithms that approximate the matrix exponential

  why?

  - The numerical methods require repeated computations of exponentials/tangent maps

- Exact computation is not an issue – but it is crucial that the exponential approximation is in the Lie group.

  why?

  - The order is "decided" from the underlying ODE method
  - Approximation from the **Lie algebra** to the **Lie group** is needed for intrinsic methods

Class of methods from "19 dubious ways. . . "

1. **Series expansions:**

   Taylor: $\exp(A) = I + A + \frac{A^2}{2!} + \cdots$

   Padé: $\exp(A) \approx [D_{pq}(A)]^{-1} N_{pq}(A)$

2. **ODE methods:** Too expensive
3. **Polynomial methods:**

   characteristic polynomial

   $$c(z) = \det(zI - A) = \prod_{i=0}^{r} (z - \lambda_i)^{\alpha_i}$$

   $\alpha_i :=$ alg. mult. of λ_i.

   minimal polynomial

   $$p(z) = \prod_{i=1}^{r} (z - \lambda_i)^{j_i}$$

   $j_i := \dim(J_i)$, largest Jordan block for λ_i.

   $$f(z) = q(z)d(z) + r(z),$$

   $d(z) = c(z), p(z)$ and $\deg(r) < \deg(d)$.

4. **Matrix decompositions:** $\exp(A) = S\exp(B)S^{-1}$
   Requires Schur decompositions, for repeated computations of the same exponential.
5. **Splitting methods:**

   $$\exp(A) = \lim_{m\to\infty} \left(\exp\left(\frac{B}{m}\right)\exp\left(\frac{C}{m}\right)\right)^m,$$

   $A = B + C$, $m = 2^j$.

6. **Krylov methods:**

   $$\exp(A)\mathbf{v} = \beta V_m \exp(H_m)\mathbf{e}_1,$$

   where $\beta = \|\mathbf{v}\|$, $\mathbf{v}_1 = \mathbf{v}/\beta$,
   V_m is a basis of $\mathcal{K}_m = \{\mathbf{v}_1, A\mathbf{v}_1, \ldots, A^{m-1}\mathbf{v}_1\}$
   and H_m upper Hessenberg.

Back
Close

Computation of the matrix exponential by GPDs

- These methods can be thought at **splitting methods**

$$\exp(A) \approx \exp(B)\exp(C), \qquad B + C = A$$

- An introduction to the theory of GPD
- Application to the computation of $\exp$: A first approach with full matrices
- Faster methods based on reduction to banded form
- A domain-decomposition approach for large problems
- Error analysis
- Some concluding remarks

Generalized polar decompositions

Ingredients:

- A Lie group $(G, \cdot\,)$
- An involutive automorphism $\sigma : G \to G$, (σ invertible, 1-to-1)

$$\sigma(xy) = \sigma(x)\sigma(y), \qquad \sigma^2(x) = x, \qquad \forall x, y \in G$$

What we can do with them:

- We can factorize

  $$z = xy \qquad \text{GPD}$$

  where

  $$\sigma(x) = x^{-1}$$

  and

  $$\sigma(y) = y$$

where x, y are appropriate group elements which are determined by z and σ.

In particular, if z is sufficiently close to the group identity e, $z = \exp(tZ)$, it is true that

$$\exp(tZ) = \exp(X(t))\exp(Y(t))$$

and for t sufficiently small, the functions $X(t)$ and $Y(t)$ are uniquely determined.

Properties of the decomposition: at the group level

$$z = \exp(tZ) = \exp(X(t))\exp(Y(t)) = xy \qquad \text{GPD of } z$$

with $\sigma(x) = x^{-1}$, $\sigma(y) = y$.

Consider the sets

$$\begin{aligned} G^\sigma &= \{z \in G : \sigma(z) = z\} && \text{fixed points of } \sigma \\ G_\sigma &= \{z \in G : \sigma(z) = z^{-1}\} && \text{anti-fixed points of } \sigma \end{aligned}$$

- G^σ has the structure of a group:

$$z_1, z_2 \in G^\sigma \qquad \Rightarrow \qquad z_1 z_2 \in G^\sigma, \quad z_1^{-1} \in G^\sigma$$

- G_σ has the structure of a *symmetric space*,

$$z_1, z_2 \in G_\sigma \qquad \Rightarrow \qquad z_1 \star z_2 = z_1 z_2^{-1} z_1 \in G_\sigma.$$

At the algebra level...

Assume $z = \exp(tZ)$, where $Z \in \mathfrak{g}$, the Lie-algebra of G. The group automorphism σ induces a Lie-algebra map

$$\mathrm{d}\sigma(Z) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \sigma(\exp(tZ)), \qquad Z \in \mathfrak{g}$$

which is also an involutive automorphism since

$$\mathrm{d}\sigma([A,B]) = [\,\mathrm{d}\sigma(A),\, \mathrm{d}\sigma(B)], \quad A, B \in \mathfrak{g}, \qquad \mathrm{d}\sigma^2 = \mathrm{id},$$

We denote

$$\begin{aligned} \mathfrak{k} &= \{Z \in \mathfrak{g} : \mathrm{d}\sigma(Z) = Z\} && \text{subalgebra of } \mathfrak{g} \\ \mathfrak{p} &= \{Z \in \mathfrak{g} : \mathrm{d}\sigma(Z) = -Z\} && \text{Lie triple system.} \end{aligned}$$

The subspaces $\mathfrak{p}$ and $\mathfrak{k}$ obey important inclusion properties:

$$\begin{aligned} [\mathfrak{k}, \mathfrak{k}] &\subseteq \mathfrak{k}, && (+1) \times (+1) = (+1) \\ [\mathfrak{k}, \mathfrak{p}] &\subseteq \mathfrak{p}, && (+1) \times (-1) = (-1) \\ [\mathfrak{p}, \mathfrak{p}] &\subseteq \mathfrak{k}, && (-1) \times (-1) = (+1). \end{aligned}$$

It is true that

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k},$$

in other words, every $Z \in \mathfrak{g}$ can be uniquely written as

$$Z = P + K, \qquad \mathrm{d}\sigma(P) = -P, \quad \mathrm{d}\sigma(K) = K,$$

where

$$P = \frac{1}{2}(Z - \mathrm{d}\sigma(Z)), \qquad K = \frac{1}{2}(Z + \mathrm{d}\sigma(Z)).$$

In summary:

Group level	Algebra level
$G^\sigma := \{z \in G : \sigma(z) = z\}$, subgrp	$\mathfrak{k} = \{Z \in \mathfrak{g} : \mathrm{d}\sigma(Z) = Z\}$, subalg.
$G_\sigma := \{z \in G : \sigma(z) = z^{-1}\}$, symm. sp.	$\mathfrak{p} = \{Z \in \mathfrak{g} : \mathrm{d}\sigma(Z) = -Z\}$, LTS
$G = G_\sigma \cdot G^\sigma$	$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$
$z = xy$	$Z = P + K$
$x = \exp(X(t)) \in G_\sigma$	$X(t) = \sum_{i=1}^\infty X_i t^i \in \mathfrak{p}, \quad X_i \in \mathfrak{p}$
$y = \exp(Y(t)) \in G^\sigma$	$Y(t) = \sum_{i=1}^\infty Y_i t^i \in \mathfrak{k}, \quad Y_i \in \mathfrak{k}$

The decomposition

$$Z = P + K$$

completely determines the functions $X(t)$ and $Y(t)$:

$$\begin{aligned} X &= Pt - \tfrac{1}{2}[P,K]t^2 - \tfrac{1}{6}[K,[P,K]]t^3 \\ &\quad + \left(\tfrac{1}{24}[P,[P,[P,K]]] - \tfrac{1}{24}[K,[K,[P,K]]]\right)t^4 \\ &\quad + \left(\tfrac{7}{360}[K,[P,[P,[P,K]]]] - \tfrac{1}{120}[K,[K,[K,[P,K]]]] - \tfrac{1}{180}[[P,K],[P,[P,K]]]\right)t^5 \\ &\quad + \mathcal{O}(t^6), \end{aligned}$$

$$\begin{aligned} Y &= Kt - \tfrac{1}{12}[P,[P,K]]t^3 + \Big(\tfrac{1}{120}[P,[P,[P,[P,K]]]] \\ &\qquad + \tfrac{1}{720}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]]\Big)t^5 + \mathcal{O}(t^7). \end{aligned}$$

In general, all the terms in the expansion of $X(t)$ and $Y(t)$ are obtained by explicit recurrence relations in terms of P and K.

Examples

- Classical polar decomposition:

$$\begin{aligned} \sigma(z) &= z^{-\top}, \quad z \in G \\ \mathrm{d}\sigma(Z) &= -Z^{\top}, \quad Z \in \mathfrak{g} \end{aligned}$$

The splitting:

$$\begin{aligned} P &= \frac{1}{2}(Z - \mathrm{d}\sigma(Z)) = \frac{1}{2}(Z + Z^{\top}), \qquad \text{symm. matrix} \\ K &= \frac{1}{2}(Z + \mathrm{d}\sigma(Z)) = \frac{1}{2}(Z - Z^{\top}), \qquad \text{skew-symm. matrix} \\ Z &= P + K \\ \exp(Z) &\approx \exp(P)\exp(K) \qquad \text{to first order} \end{aligned}$$

we recover one of the splitting methods described in "19 dubious ways. . . ".

- In this case, $\exp(tZ) = \exp(X(t))\exp(Y(t))$ is the continuous analytic version of the classical polar decomposition of a matrix as the product of a **symmetric PD** matrix ($\exp(X(t))$) and an **orthogonal** matrix ($\exp(Y(t))$).

- Low-rank decompositions: Let $S = \mathrm{diag}(1, 1, 1 \ldots, -1)$ and

$$\begin{aligned} \sigma(z) &= SzS^{-1}, \quad z \in G \\ \mathrm{d}\sigma(Z) &= SZS^{-1}, \quad Z \in \mathfrak{g} \end{aligned}$$

The splitting:

$$P = \frac{1}{2}(Z - \mathrm{d}\sigma(Z)) = \left[\begin{array}{cccc|c} 0 & \cdots & 0 & & z_{1,n} \\ \vdots & 0 & \cdots & & \vdots \\ 0 & \vdots & \ddots & & z_{n-1,1} \\ \hline z_{n,1} & \cdots & z_{n,n-1} & & 0 \end{array}\right], \quad \text{rank-2 matrix}$$

$$K = \frac{1}{2}(Z + \mathrm{d}\sigma(Z)) = \left[\begin{array}{ccc|c} z_{1,1} & \cdots & z_{1,n-1} & 0 \\ \vdots & \cdots & \vdots & \\ z_{n-1,1} & \vdots & z_{n-1,n-1} & 0 \\ \hline 0 & \cdots & 0 & z_{n,n} \end{array}\right], \quad \text{block diagonal matrix}$$

$$Z = P + K$$

$$\exp(Z) \approx \exp(P)\exp(K) \quad \text{to first order}$$

Approximation of the matrix exponential

Recall that by GPD:

$$\exp(tZ) = \exp(X(t))\exp(Y(t)),$$

where

$$\begin{aligned} X &= Pt - \tfrac{1}{2}[P,K]t^2 - \tfrac{1}{6}[K,[P,K]]t^3 \\ &\quad + \left(\tfrac{1}{24}[P,[P,[P,K]]] - \tfrac{1}{24}[K,[K,[P,K]]]\right)t^4 \\ &\quad + \left(\tfrac{7}{360}[K,[P,[P,[P,K]]]] - \tfrac{1}{120}[K,[K,[K,[P,K]]]] - \tfrac{1}{180}[[P,K],[P,[P,K]]]\right)t^5 \\ &\quad + \mathcal{O}(t^6), \end{aligned}$$

$$\begin{aligned} Y &= Kt - \tfrac{1}{12}[P,[P,K]]t^3 + \left(\tfrac{1}{120}[P,[P,[P,[P,K]]]]\right. \\ &\qquad \left. + \tfrac{1}{720}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]]\right)t^5 + \mathcal{O}(t^7). \end{aligned}$$

and

$$Z = P + K, \qquad P = \frac{1}{2}(Z - \mathrm{d}\sigma(Z)), \qquad K = \frac{1}{2}(Z + \mathrm{d}\sigma(Z)).$$

Back
Close

A splitting method that approximates $\exp(tZ)$:

- Choose an appropriate σ.
- Split $Z = P + K$.
- Truncate the expansion

$$X(t) = Pt + \frac{1}{2}t^2[P,K] + \cdots, \quad Y(t) = Kt - \frac{1}{12}t^3[P,[P,K]] + \cdots$$

  to desired order.

- Compute the exponential of $X(t) \in \mathfrak{p}$
- Set $Z_1 = Y(t)$
- Repeat

In general, we iterate the procedure on the reduced space until we get a space of low dimension. At the end,

$$\exp(tZ) \approx \exp(X^{[1]})\exp(X^{[2]})\cdots\exp(X^{[m]})\exp(Y^{[m]}).$$

What are good choices of σ?

- The splitted factors should be easy to compute.
- Commutators should have a low complexity.
- Exponential/commutators of splitted parts should be easy to compute (either approximately or preferably exactly).

A good choice is splitting in matrices of low rank, for instance, borderded matrices: take

$$S = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & 0 \\ \hline 0 & \cdots & 0 & & -1 \end{array}\right], \qquad \mathrm{d}\sigma(Z) = SZS,$$

then

$$P = \left[\begin{array}{ccc|c} 0 & \cdots & 0 & z_{1,n} \\ \vdots & 0 & \cdots & \vdots \\ 0 & \vdots & \ddots & z_{n-1,1} \\ \hline z_{n,1} & \cdots & z_{n,n-1} & 0 \end{array}\right], \qquad K = \left[\begin{array}{ccc|c} z_{1,1} & \cdots & z_{1,n-1} & 0 \\ \vdots & \cdots & \vdots & \\ z_{n-1,1} & \vdots & z_{n-1,n-1} & 0 \\ \hline 0 & \cdots & 0 & z_{n,n} \end{array}\right].$$

Such automorphisms work for $\mathrm{GL}(n), \mathrm{SL}(n), \mathrm{SO}(n)$.
Note that the commutators appearing in the expansion can be computed in $\mathcal{O}(n^2)$ computations (n^3 if the procedure is iterated for matrices of decreasing dimension)

An Euler–Rodrigues like formula for bordered matrices

The exponential of bordered matrices can be computed exactly by means of a formula similar to the Euler–Rodrigues formula for computing the exponential of a 3×3 skew-symmetric matrices.

Assume that $A \in \mathfrak{p}$ is of the form

$$A = \left[\begin{array}{c|c} O & \mathbf{a} \\ \hline \mathbf{b}^T & 0 \end{array}\right], \qquad \mathbf{a}, \mathbf{b} \in \mathbb{R}^n. \tag{3.1}$$

Then,

$$\exp(A) = \begin{cases} I + \frac{\sinh\theta}{\theta} A + \frac{1}{2}\left(\frac{\sinh(\theta/2)}{\theta/2}\right)^2 A^2, & \text{if } \mathbf{a}^T\mathbf{b} > 0, \theta = \sqrt{\mathbf{a}^T\mathbf{b}}, \\ I + A + \frac{1}{2}A^2, & \text{if } \mathbf{a}^T\mathbf{b} = 0, \\ I + \frac{\sin\theta}{\theta} A + \frac{1}{2}\left(\frac{\sin(\theta/2)}{\theta/2}\right)^2 A^2, & \text{if } \mathbf{a}^T\mathbf{b} < 0, \theta = \sqrt{-\mathbf{a}^T\mathbf{b}} \end{cases}$$

where

$$A^2 = \left[\begin{array}{c|c} \mathbf{a}\mathbf{b}^T & \mathbf{0} \\ \hline \mathbf{0}^T & \theta^2 \end{array}\right].$$

Minimal polynomial:

$$p(z) = \lambda(\lambda^2 - \mathbf{a}^\top\mathbf{b}), \qquad \deg(p) = 3$$

Characteristic polynomial

$$c(z) = \lambda^{n-2}(\lambda^2 - \mathbf{a}^\top\mathbf{b}), \qquad \deg(c) = n.$$

Back
Close

Note that $\exp(A)$ never needs being computed explicitely but always applied to a vector/matrix $\mathbf{v}$.

$$\begin{bmatrix} \mathbf{w}_k \\ w \end{bmatrix} = \exp(A)\mathbf{v} = \exp(A) \begin{bmatrix} \mathbf{v}_k \\ v \end{bmatrix} = \begin{bmatrix} \mathbf{v}_k + \zeta_1 \mathbf{a} \\ \zeta_2 \end{bmatrix},$$

where

$$\begin{aligned} \zeta_1 &= [\eta_1 v + \eta_2(\mathbf{b}^\top \mathbf{v}_k)], \\ \zeta_2 &= (1+\theta)v + (\mathbf{b}^\top \mathbf{v}_k). \end{aligned}$$

Cost of the computation (including both addition and multiplication) of the 'Euler–Rodrigues' exponential. The (k,k) column corresponds to the case when $\mathbf{a}, \mathbf{b}$ are full, the (k,p) corresponds to the case when $\mathbf{a}$ is full while only the last p components of $\mathbf{b}$ are nonzero and finally the (p,p) column corresponds to both $\mathbf{a}$ and $\mathbf{b}$ having only the last p components nonzero.

Cost of $\exp(A)$	(k,k)	(k,p)	(p,p)
$\mathbf{a}^\top \mathbf{b}$	$2k$	$2p$	$2p$
$\mathbf{b}^\top \mathbf{v}_k$	$2k$	$2p$	$2p$
$\zeta_1 \mathbf{a}$	k	k	p
$\mathbf{w}_k$	k	k	p
total, stage k	$6k$	$2k+4p$	$6p$
total, summing $1 \le k \le n$ (vector)	$3n^2$	n^2+4pn	$6pn$
matrix (n vectors)	$2n^3$	n^3+2pn^2	$4pn^2$

On the computation of commutators

Our first observation is that the involutions S are usually chosen so that $P = \Pi_{\mathfrak{p}}(Z)$ has low rank, hence only just a few nonzero eigenvalues.

- Use the theory of minimal polynomial, the least degree monic polynomial such that

$$p(\mathrm{ad}_A) = 0.$$

Lemma 3.1 *Consider the bordered matrix A in (3.1) with $\mathbf{a}\mathbf{b}^\top \neq O$. The minimal polynomial of ad_A is*

$$\begin{aligned} p(\lambda) &= \lambda(\lambda - 2\theta)(\lambda + 2\theta)(\lambda - \theta)(\lambda + \theta) \qquad (3.2) \\ &= \lambda^5 - 5\mathbf{b}^\top\mathbf{a}\lambda^3 + 4(\mathbf{b}^\top\mathbf{a})^2\lambda, \end{aligned}$$

where $\theta = \sqrt{\mathbf{b}^\top\mathbf{a}}$. If $\mathbf{a}\mathbf{b}^\top = O$, and $\mathbf{a}$ and $\mathbf{b}$ are not both zero, then the minimal polynomial is

$$p(\lambda) = \lambda^3. \qquad (3.3)$$

Proof. If A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ with algebraic multiplicities $r_1, r_2, \ldots, r_m$ respectively, the minimal polynomial of A has the form

$$q(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{g_i},$$

where g_i is the order of the largest Jordan block of A corresponding to the eigenvalue λ_i.

Let us assume first that $\mathbf{b}^\top \mathbf{a} \neq 0$. Assume that $\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ v_2 \end{bmatrix}$ is eigenvector of A corresponding to the eigenvalue λ. Imposing $A\mathbf{v} = \lambda \mathbf{v}$, we have

$$\begin{aligned} \mathbf{a} v_2 &= \lambda \mathbf{v}_1 \\ \mathbf{b}^\top \mathbf{v}_1 &= \lambda v_2 \end{aligned}$$

and deduce immediately that the eigenvalues of A are $\lambda = \pm\theta = \pm\sqrt{\mathbf{b}^\top \mathbf{a}}$ and $\lambda = 0$ with algebraic multiplicities one, one, and $n-2$ respectively. It is easily verified that these are also their geometric multiplicities: for $\lambda = \pm\theta$, eigenvectors are of the form $[\mathbf{a}, \pm 1]^\top$; for the zero eigenvalues, eigenvectors are of the form $[\mathbf{v}_1, 0]^\top$, $\mathbf{0} \neq \mathbf{v}_1 \in R^{n-1}$, satisfying $\mathbf{b}^\top \mathbf{v}_1 = 0$, furthermore, it is possible to find $n-2$ of those that are linearly independent.
Since the eigenvalues and eigenvectors of ad_A are the form $\lambda_i - \lambda_j$ and $\mathbf{y}_i^\top \mathbf{x}_j$ respectively, the λ_is being eigenvalues of A with left and right eigenvector $\mathbf{y}_i$ and $\mathbf{x}_i$ respectively, we deduce that ad_A has eigenvalues

$$\lambda = \pm 2\theta, \qquad \lambda = \pm\theta$$

with algebraic/geometric multiplicities one each, and

$$\lambda = 0$$

with algebraic and geometric multiplicity $n^2 - 4$. This implies that all Jordan blocks have size one, from which it follows directly that the minimal polynomial of ad_A is of the form (3.2).
Next, if $\theta = 0$ but $\mathbf{a}\mathbf{b}^\top \neq O$, namely $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$, the eigenvalues of A, that we write as $[\mathbf{v}_1, v_2]^\top$, must obey the conditions

$$\begin{aligned} \mathbf{a} v_2 &= \mathbf{0} \\ \mathbf{b}^\top \mathbf{v}_1 &= 0. \end{aligned}$$

Since $\mathbf{a} \neq 0$, it must necessarily be $v_2 = 0$. Therefore eigenvalues must be of the form $[\mathbf{v}_1, 0]$. Recall that $\mathbf{v}_1$ has $n-1$ entries ($n-1$ free parameters) while the second equation $\mathbf{b}^\top \mathbf{v}_1 = 0$ gives only a linear constraint: This mean that we can find only $n-2$ linearly independent eigenvalues and two further linearly independent generalized eigenvalues. In terms of Jordan blocks, this means that A has a Jordan block of the form

$$J(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

hence λ^3 is the minimal polynomial of A and, as a consequence, $A^3 = O$. Passing to the adjoint operator ad_A, recall that, for an arbitrary matrix C,

$$\mathrm{ad}_A^k C = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} A^i C A^{k-i}, \qquad k = 1, 2, \ldots. \tag{3.4}$$

Clearly, $\mathrm{ad}_A^5 C = O$ since in all terms there appears a power A^i with $i \geq 3$. For lower order powers, there are always terms of the type $A^i C A^{k-i}$ where $i, k-i \leq 2$. This means that it is always possible to find a matrix C for which at least one of terms does not vanish. Hence the minimal polynomial of ad_A is

$$p(\lambda) = \lambda^5.$$

Finally, in the case when either $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$, by direct computation,

$$A^2 = O,$$

hence the minimal polynomial of A is λ^2. Insofar as ad_A is concerned, the first power to vanish in (3.4) is ad_A^3, and no lower power vanishes for arbitrary matrices C. Hence the minimal polynomial is

$$p(\lambda) = \lambda^3.$$

This completes the proof of the lemma. □

Theorem 3.2 *Assume that the matrix A is of the form (3.1). Then, for every $k = 1, 2, \ldots$, commutators by A can be computed as*

$$\mathrm{ad}_A^k = [C_1 + (-1)^k C_2] 2^k \theta^k + [C_3 + (-1)^k C_4] \theta^k, \qquad k = 1, 2, \ldots \tag{3.5}$$

when $\theta = \sqrt{\mathbf{b}^\top \mathbf{a}} \neq 0$, and

$$\begin{aligned} C_1 - C_2 &= \frac{1}{6}\left(-\frac{\mathrm{ad}_A}{\theta} + \frac{\mathrm{ad}_A^3}{\theta^3}\right) \\ C_3 - C_4 &= \frac{1}{3}\left(\frac{4\mathrm{ad}_A}{\theta} - \frac{\mathrm{ad}_P^3}{\theta^3}\right) \\ C_1 + C_2 &= \frac{1}{12}\left(-\frac{\mathrm{ad}_A^2}{\theta^2} + \frac{\mathrm{ad}_A^4}{\theta^4}\right) \\ C_3 + C_4 &= \frac{1}{3}\left(\frac{4\mathrm{ad}_A^2}{\theta^2} - \frac{\mathrm{ad}_A^4}{\theta^4}\right). \end{aligned} \tag{3.6}$$

If $\theta = 0$ but $\mathbf{a}\mathbf{b}^\top \neq O$, then

$$\mathrm{ad}_A^k = O, \qquad k = 5, 6, 7, \ldots.$$

If $\theta = 0$ and either $\mathbf{a}$ or $\mathbf{b}$ is a zero vector, then

$$\mathrm{ad}_A^k = O, \qquad k = 3, 4, 5, \ldots.$$

Proof. It follows from the minimal polynomial (3.2). □

Complexity of the algorithms for full matrices:

Order 2	$\mathfrak{sl}(n), \mathfrak{so}(p,q)$ vector	$\mathfrak{sl}(n), \mathfrak{so}(p,q)$ matrix	$\mathfrak{so}(n)$ vector	$\mathfrak{so}(n)$ matrix
splitting	$1\frac{1}{3}n^3$	$1\frac{1}{3}n^3$	$\frac{2}{3}n^3$	$\frac{2}{3}n^3$
assembly exp	$3n^2$	$2n^3$	$3n^2$	$2n^3$
total	$1\frac{1}{3}n^3$	$3\frac{1}{3}n^3$	$\frac{2}{3}n^3$	$2\frac{2}{3}n^3$

Order 3(4)	$\mathfrak{sl}(n), \mathfrak{so}(p,q)$ vector	$\mathfrak{sl}(n), \mathfrak{so}(p,q)$ matrix	$\mathfrak{so}(n)$ vector	$\mathfrak{so}(n)$ matrix
splitting	$5(7)n^3$	$5(7)n^3$	$2\frac{1}{2}(4)n^3$	$2\frac{1}{2}(4)n^3$
assembly exp	$3n^2$	$2n^3$	$3n^2$	$2n^3$
total	$5(7)n^3$	$7(9)n^3$	$2\frac{1}{2}(4)n^3$	$4\frac{1}{2}(6)n^3$

These algorithms have a complexity that is comparable with other classical algorithms, like for instance diagonal Padé approximants.

Feasible algorithms are up to order 4, because for higher order the complexity becomes larger (although still $\mathcal{O}(n^3)$).

Faster algorithms

Combine GI and classical Linear Algebra techniques

The main difference with the approach presented before is that the matrix Z is preprocessed and reduced to a 'sparse' form **stable under commutation**, which is

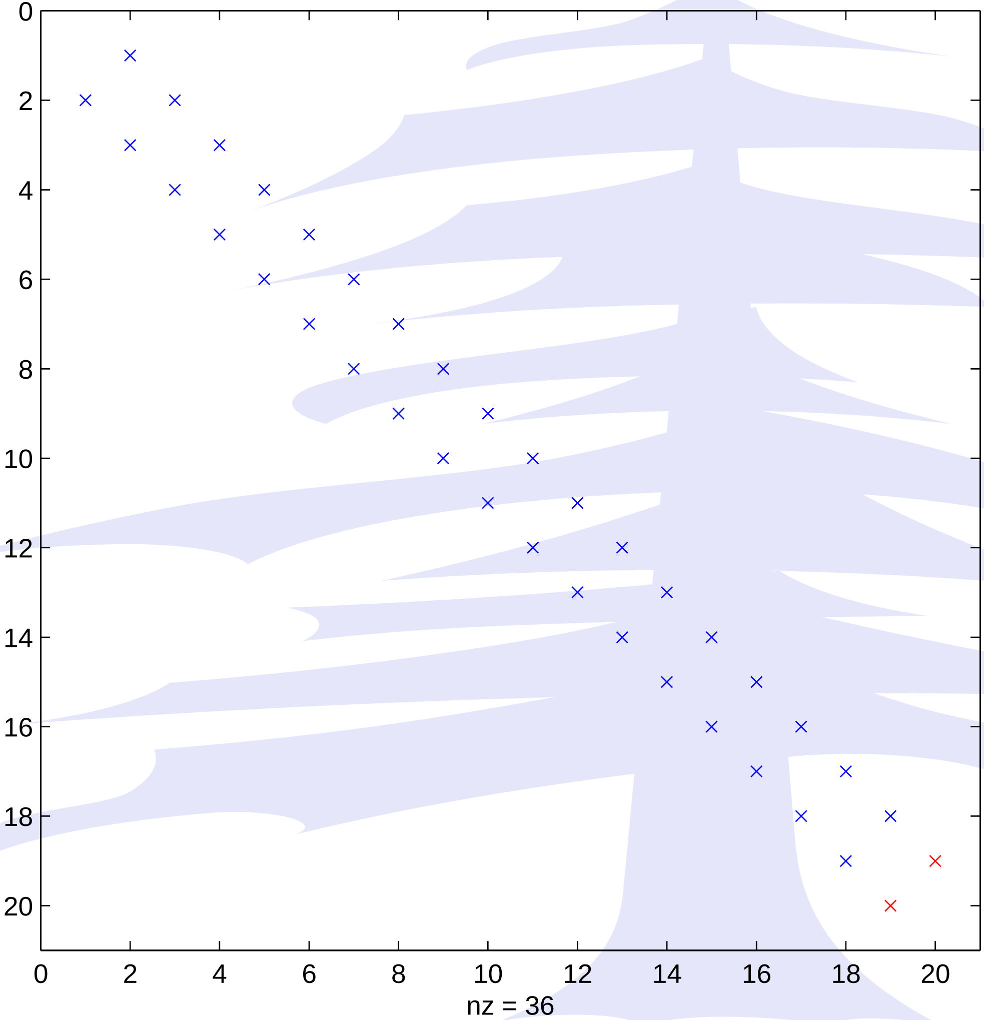
- tridiagonal (for symmetric and skew-symmetric matrices),
- upper Hessenberg (for matrices in $\mathfrak{sl}(n)$),
- butterfly form (for symplectic matrices).

$$Z = VBV^{-1}$$

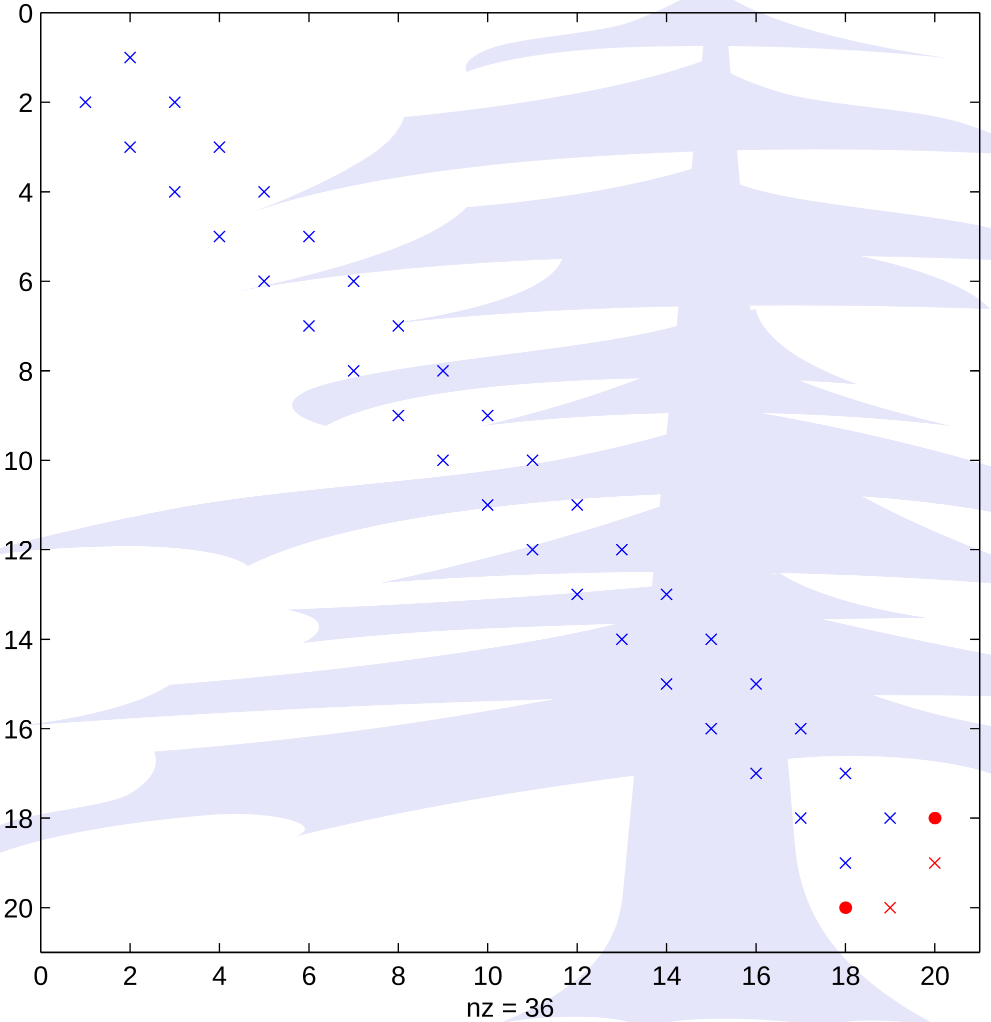
Again, we split rows and columns and start computing commutators.
In the following example, we consider a skew-symmetric tridiagonal matrix.

- 'red' for the $\mathfrak{p}$-part, 'blue' for the $\mathfrak{k}$-part
- updated elements are denoted with dots instead of crosses

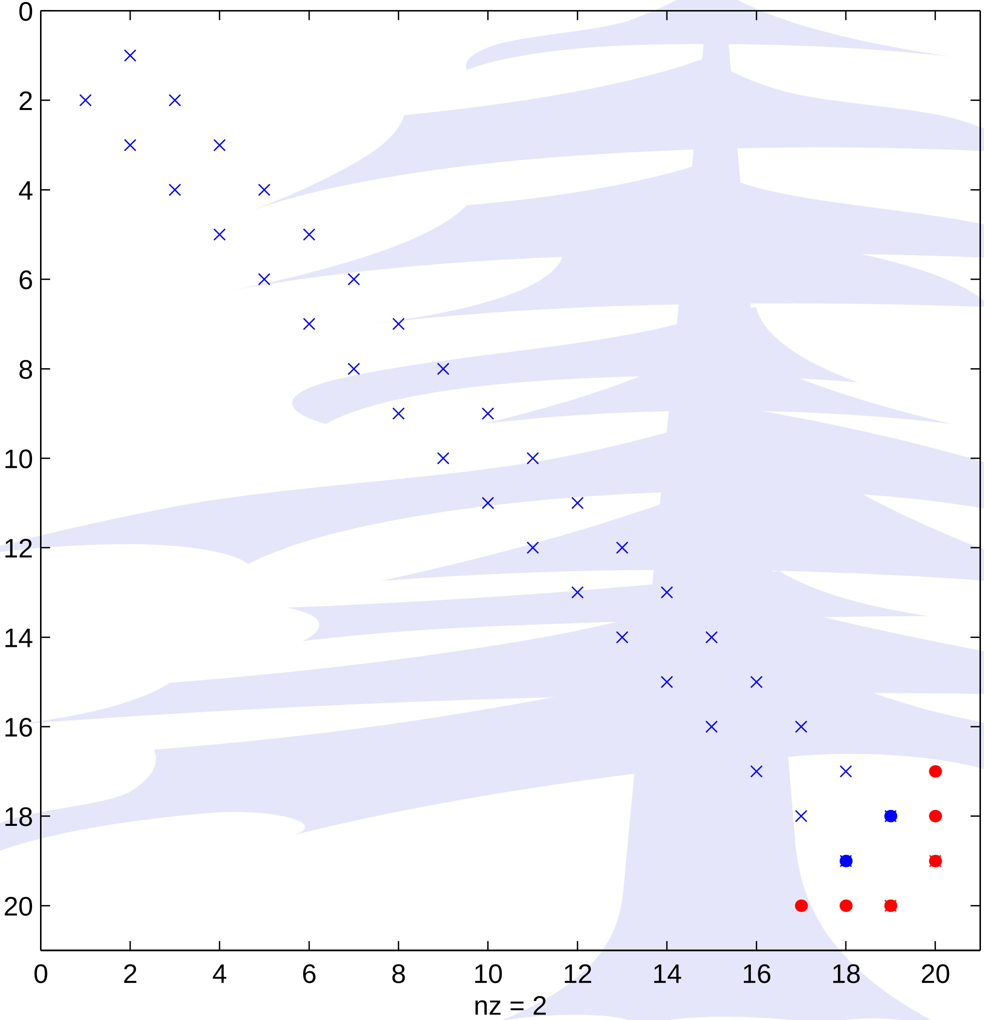
Order 1:



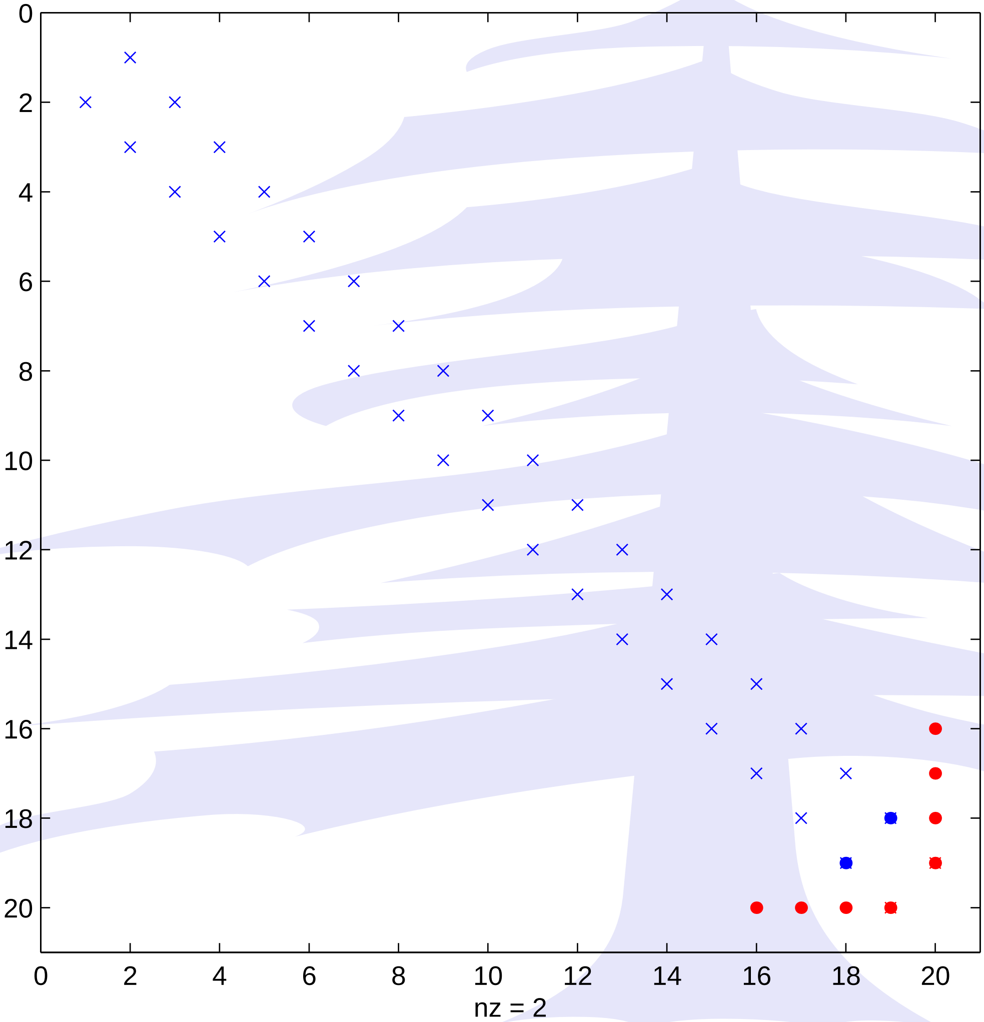
Order 2:



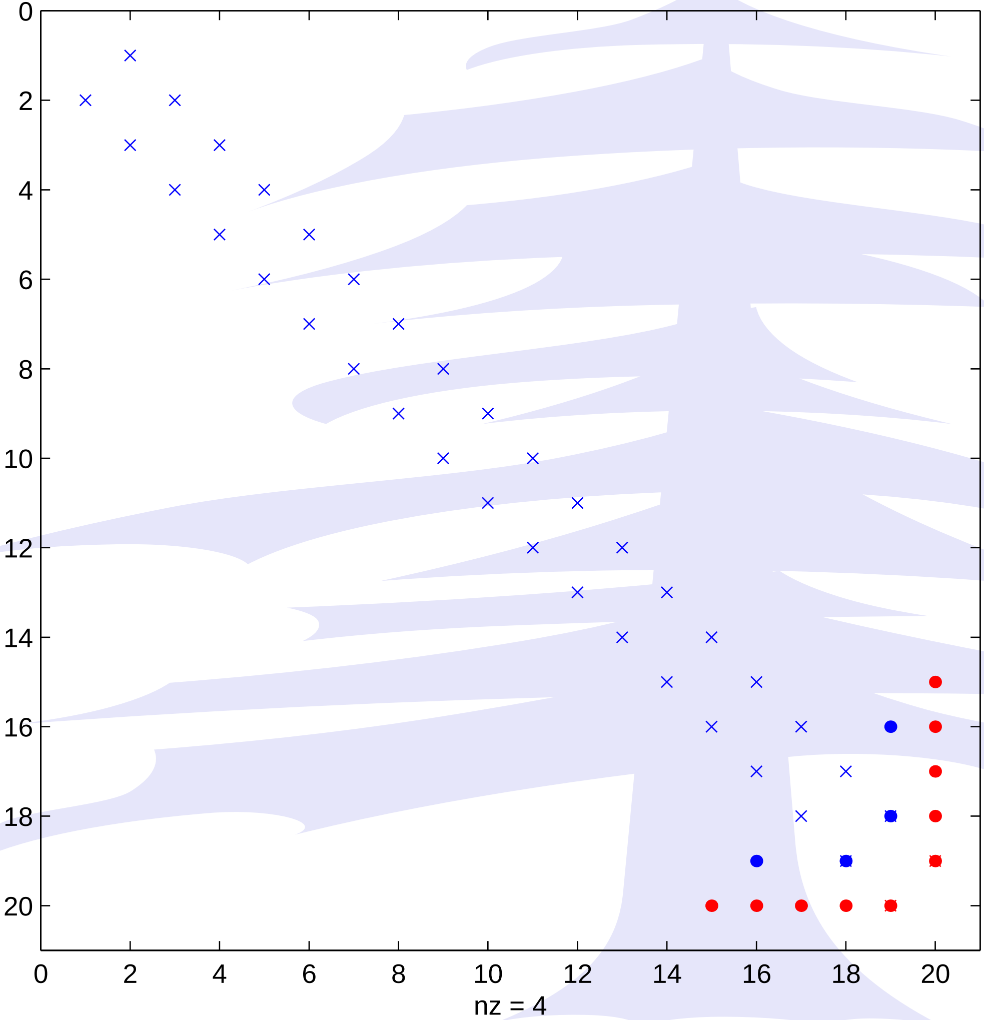
Order 3:



Order 4:



Order 5:

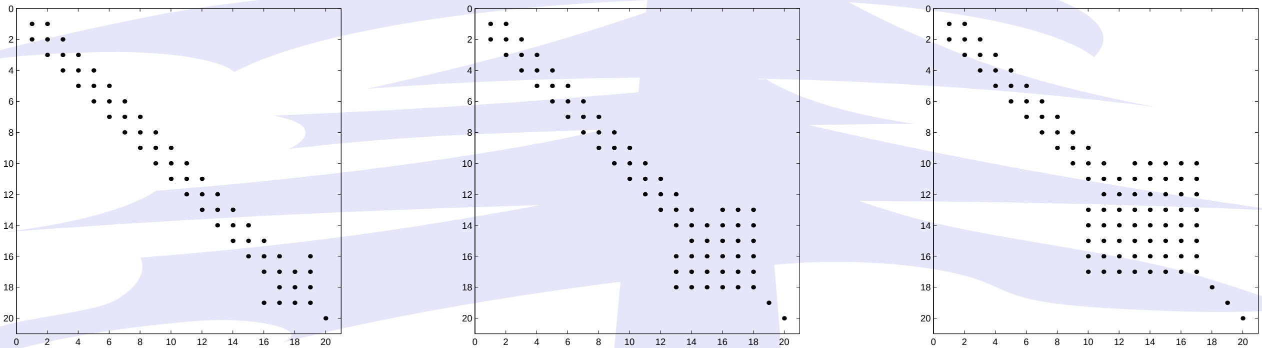


Main observations:

- Each extra order fills in two symmetric elements in the $\mathfrak{p}$-part.
- The fill-in in the $\mathfrak{k}$-part starts only at order 5.
- As long as the matrices P, K are tridiagonal, the commutators cost $\mathcal{O}(1)$.

The 'ugly' and the 'bad' fill-in

- The fill-in in the $\mathfrak{p}$-part is 'ugly' but not harmful: once once the $\mathfrak{p}$-term is computed up to desired order, one needs only compute the exponential.
- The fill-in in the $\mathfrak{k}$-part is much more dangerous: if not taken care of, it propagates and we lose the whole benefits of our tridiagonalization/reduction to Hessenberg



Therefore the fill-in elements in the $\mathfrak{k}$ part must be annihilated by, for instance, Givens rotations ($\mathcal{O}(1)$ computations)

Order 2	Full vector	Full matrix	Tridiag vector	Tridiag matrix
Tridiag.	–	–	n^3	n^3
order cond.	$\frac{2}{3}n^3$	$\frac{2}{3}n^3$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
assembly exp	$3n^2$	$2n^3$	$6pn$	$4pn^2$
total	$\frac{2}{3}n^3$	$2\frac{2}{3}n^3$	$n^3+\mathcal{O}(pn)$	n^3+4pn^2

Order 3	Full vector	Full matrix	Tridiag vector	Tridiag matrix
Tridiag.	–	–	n^3	n^3
order cond.	$2\frac{1}{2}n^3$	$2\frac{1}{2}n^3$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
assembly exp	$3n^2$	$2n^3$	$6pn$	$4pn^2$
total	$2\frac{1}{2}n^3$	$4\frac{1}{2}n^3$	$n^3+\mathcal{O}(pn)$	n^3+4pn^2

Order 4	Full vector	Full matrix	Tridiag vector	Tridiag matrix
Tridiag.	–	–	n^3	n^3
order cond.	$4n^3$	$4n^3$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
assembly exp	$3n^2$	$2n^3$	$6pn$	$4pn^2$
total	$4n^3$	$6n^3$	$n^3+\mathcal{O}(pn)$	n^3+4pn^2

Comparison of cost of the approximation of the exponential without (Full) and with reduction to tridiagonal form (Tridiag) for splittings of order $2, 3, 4$. Only dominant terms are reported.

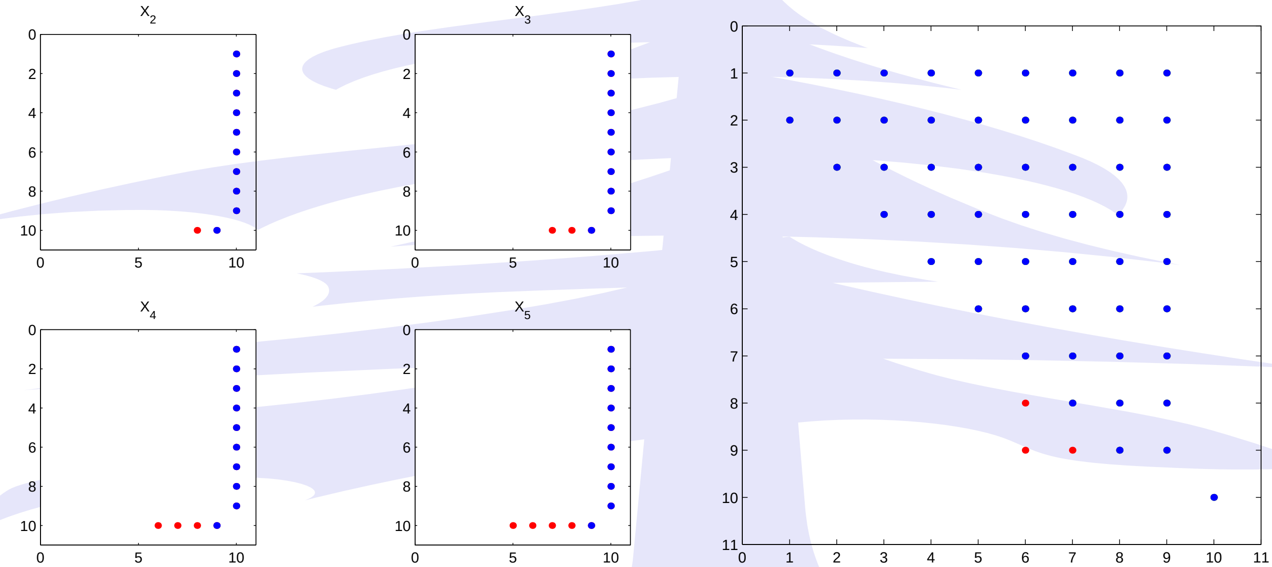
Matrices in $\mathrm{GL}(n), \mathrm{SL}(n)$

$$z \in \mathrm{GL}(n) \Leftrightarrow \det(z) \neq 0 \qquad z \in \mathrm{SL}(n) \Leftrightarrow \det(z) = 1$$

$$Z \in \mathfrak{gl}(n) \Leftrightarrow Z \text{ is } n \times n \qquad Z \in \mathfrak{sl}(n) \Leftrightarrow \mathrm{tr}(Z) = 1$$

- Tridiagonalization possible, but can be unstable.
- Reduce to Hessenberg form by orthogonal transformations.

Order 5 terms for matrices in Hessenberg form:



Back
Close

Symplectic matrices

$$z \in \mathrm{Sp}(n) \Leftrightarrow zJz^\top = J$$
$$Z \in^{(} n) \Leftrightarrow ZJ + JZ^\top = O,$$

where $J = \begin{bmatrix} O_n & -I_n \\ I_n & O_n \end{bmatrix}$

- Automorphisms $\sigma(z) = \mathcal{S}z\mathcal{S}$, where $\mathcal{S} = \begin{bmatrix} S & O_n \\ O_n & S \end{bmatrix}$

$$S = \left[\begin{array}{ccc|c} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \hline 0 & \cdots & 0 & -1 \end{array}\right].$$

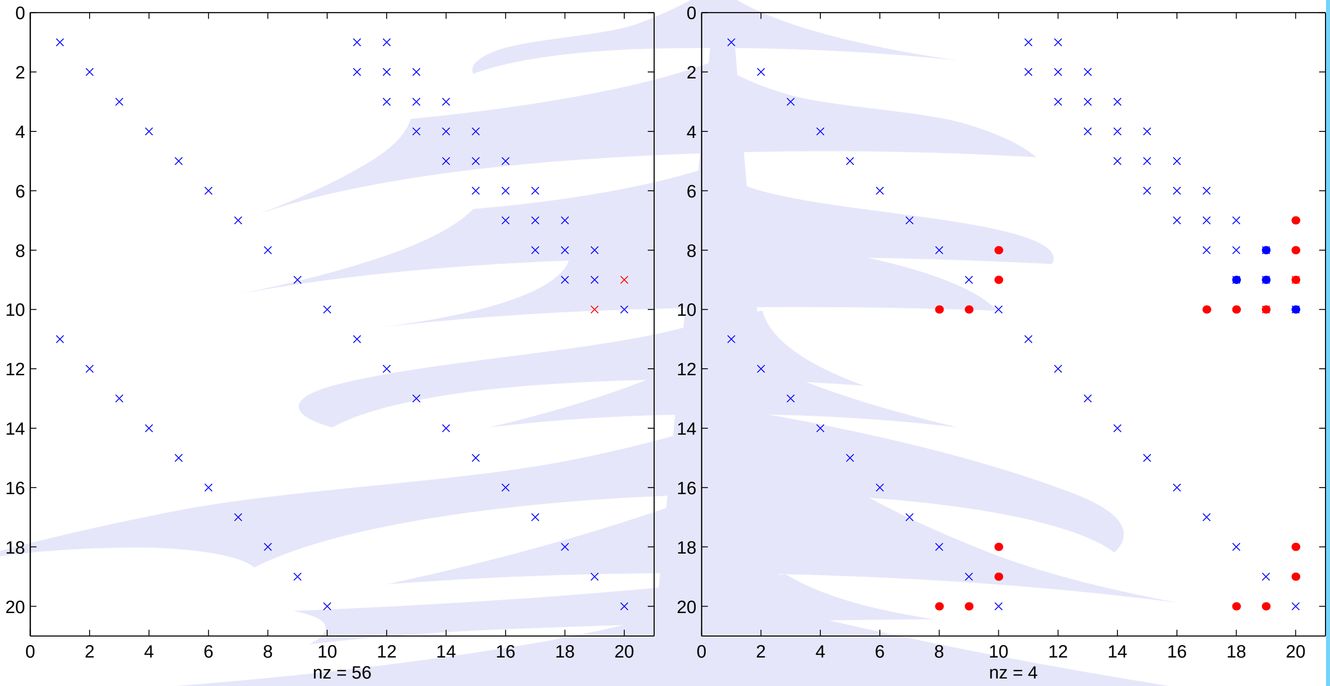
- Symplectic matrices are not necessarily closed under orthogonal transformations, hence cannot use Hessenberg/Tridiag forms
- Reduce to butterfly form by symplectic transformations: for matrices in the group

$$\begin{bmatrix} D_1 & T_1 \\ D_2 & T_2 \end{bmatrix},$$

D_i diagonal, T_i symmetric.

- At algebra level: $T_2 = -D_1^\top$.

Order 5 terms for matrices in butterfly form:



Reduction to tridiagonal/Hessenberg/ butterfly form

- For tridiagonalization/Hessenberg use Householder reflections:

$$H = I - \beta \mathbf{v}\mathbf{v}^T, \qquad \beta = \frac{2}{\|\mathbf{v}\|^2}$$

Then,

$$HZH = Z - \beta \mathbf{v}\mathbf{v}^T Z - \beta Z \mathbf{v}\mathbf{v}^T + \beta^2 \underbrace{\mathbf{v}\mathbf{v}^T Z \mathbf{v}\mathbf{v}^T}_{0 \text{ if } Z \text{ is skew}}$$

Cost:

- n^3 for symmetric/skew-symmetric matrices
- $\frac{10}{3}n^3$ for arbitrary matrices

- Reduction to butterfly form is done by means of **symplectic** transformations:
  - symplectic Givens/Householder (orthosymplectic),
  - symplectic Gauss transformations (not orthogonal)

  as proposed by Faßbender, Benner, Watkins (at the group level, for QR-like iterations).

  The basic idea of the algorithm is: at each step j,
  - bring the jth column of M into the desired form
  - bring the $(n+j)$th row of M into the desired form.

  The algorithm uses mostly the orthosymplectic transformations and only few non-orthogonal symplectic transformations, therefore one can expect relatively good stability properties.

- Cost of reduction to butterfly form

$$20\frac{2}{3}n^3$$

  for a matrix of dimension $2n$.

- **symplectic Givens transformations:,**

$$G = \left[\begin{array}{ccc|ccc} I_{k-1} & & & & & \\ & c & & & s & \\ & & I_{n-k} & & & \\ \hline & & & I_{k-1} & & \\ & -s & & & c & \\ & & & & & I_{n-k} \end{array}\right],$$

- **symplectic Householder transformations:**

$$H = \left[\begin{array}{cc|cc} I_{k-1} & & & \\ & Q & & \\ \hline & & I_{k-1} & \\ & & & Q \end{array}\right], \qquad Q = I_{n-k+1} - \beta \mathbf{v}\mathbf{v}^\top, \quad \beta = \frac{2}{\|v\|^2},$$

- **symplectic Gauss transformations:**

$$L = \left[\begin{array}{cccc|cccc} I_{k-2} & & & & & & & \\ & c & & & & & d & \\ & & c & & & d & & \\ & & & I_{n-k} & & & & \\ \hline & & & & I_{k-2} & & & \\ & & & & & c^{-1} & & \\ & & & & & & c^{-1} & \\ & & & & & & & I_{n-k} \end{array}\right].$$

Comparison of cost of the approximation of the exponential without (Full) and with reduction to Hessenberg form (Hess) for splittings of order $2, 3, 4$. Only dominant terms are reported.

Order	Full		Hess	
2	vector	matrix	vector	matrix
Hessenberg	–	–	$3\frac{1}{3}n^3$	$3\frac{1}{3}n^3$
order cond.	$1\frac{1}{3}n^3$	$1\frac{1}{3}n^3$	$\frac{1}{3}n^3$	$\frac{1}{3}n^3$
assembly exp	$3n^2$	$2n^3$	n^2	n^3
total	$1\frac{1}{3}n^3$	$2\frac{1}{3}n^3$	$3\frac{2}{3}n^3$	$4\frac{2}{3}n^3$

Order	Full		Hess	
3	vector	matrix	vector	matrix
Hessenberg	–	–	$3\frac{1}{3}n^3$	$3\frac{1}{3}n^3$
order cond.	$5n^3$	$5n^3$	$\frac{2}{3}n^3$	$\frac{2}{3}n^3$
assembly exp	$3n^2$	$2n^3$	n^2	n^3
total	$5n^3$	$7n^3$	$4n^3$	$5n^3$

Order	Full		Hess	
4	vector	matrix	vector	matrix
Hessenberg	–	–	$3\frac{1}{3}n^3$	$3\frac{1}{3}n^3$
order cond.	$7n^3$	$7n^3$	n^3	n^3
assembly exp	$3n^2$	$2n^3$	n^2	n^3
total	$7n^3$	$9n^3$	$4\frac{1}{3}n^3$	$5\frac{1}{3}n^3$

A divide and conquer approach

Consider $Z \in \mathfrak{so}(n)$ and assume that it is already in tridiagonal form. Our point of departure is to consider an inner automorphism $\sigma(Z) = SZS$ where

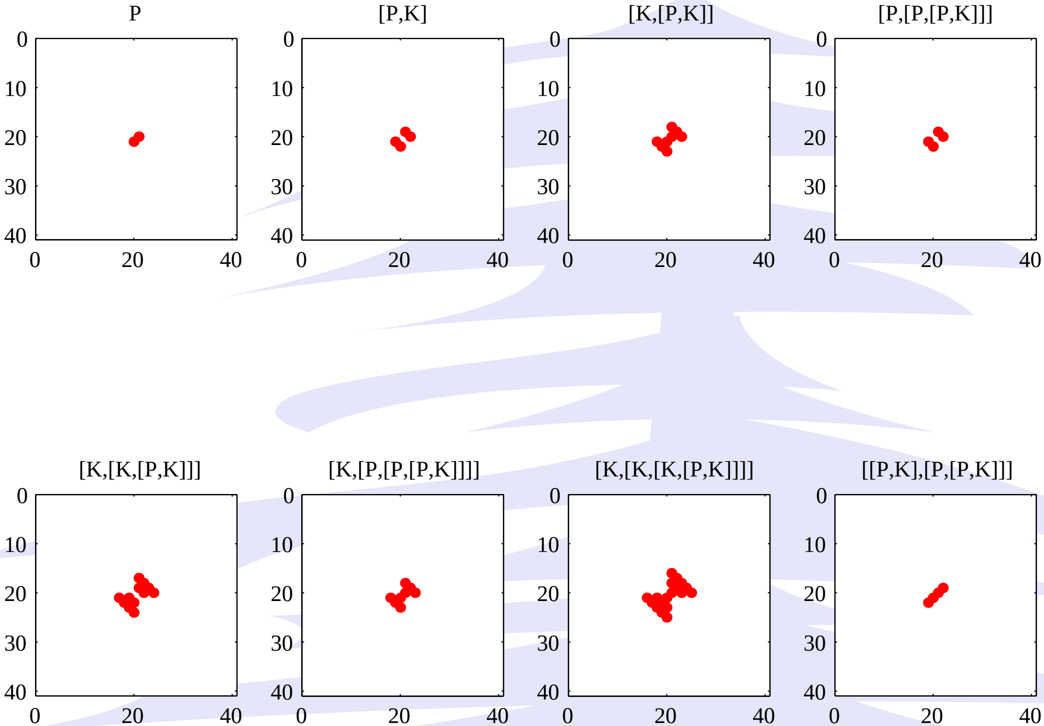
$$S = \begin{bmatrix} I_{n_1 \times n_1} & O_{n_1 \times n_2} \\ O_{n_2 \times n_1} & -I_{n_2 \times n_2} \end{bmatrix},$$

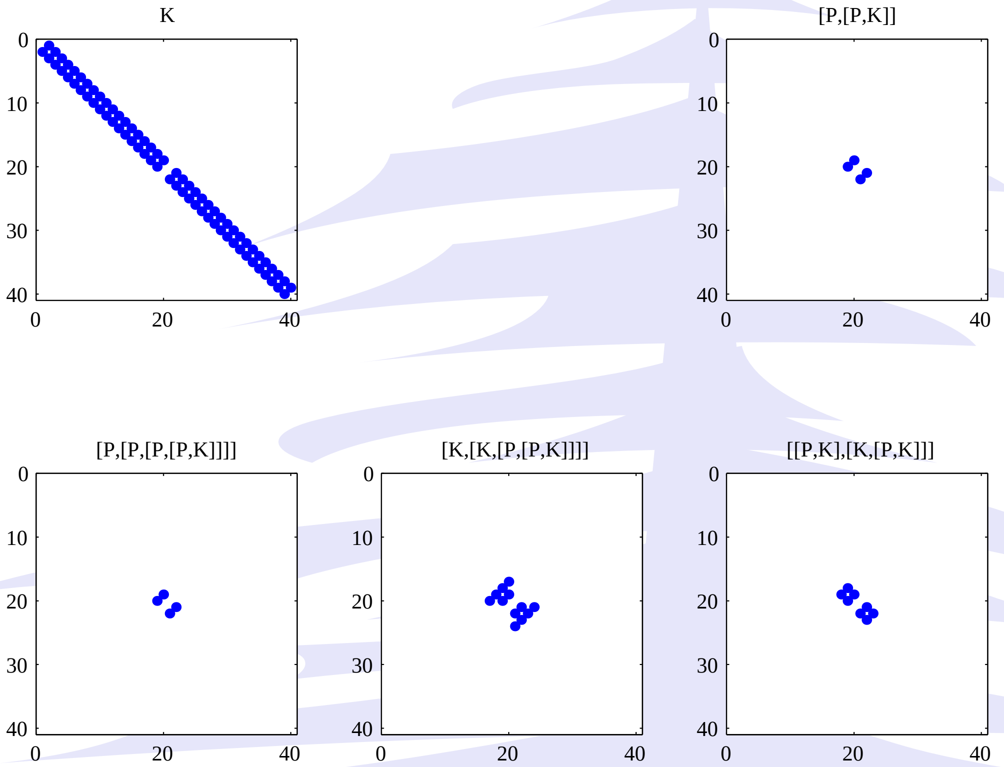
where $n_1 + n_2 = n$: an obvious choice is $n_1 = \lfloor \frac{n}{2} \rfloor$, however, other choices are possible, e.g. the index corresponding to the least off-diagonal element.

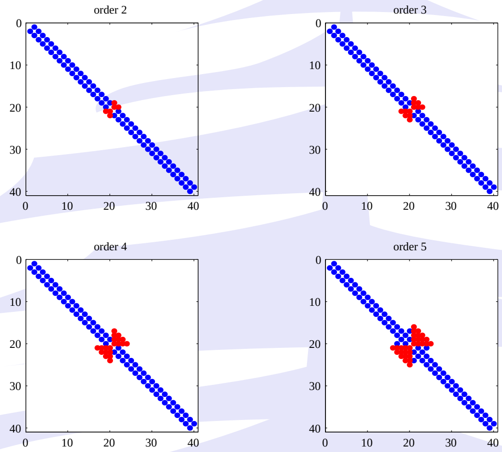
With this choice,

$$K = \begin{bmatrix} K_1 & O \\ O & K_2 \end{bmatrix}, \qquad P = \begin{bmatrix} O & P_1 \\ P_2 & O \end{bmatrix}.$$

- both K_1 and K_2 are tridiagonal
- P_1 and P_2 have a single nonzero entry, in the lower left and upper right corner respectively.


P
[P,K]
[K,[P,K]]
[P,[P,[P,K]]]
[K,[K,[P,K]]]
[K,[P,[P,[P,K]]]]
[K,[K,[K,[P,K]]]]
[[P,K],[P,[P,K]]]
0 10 20 30 40
0 20 40


K
[P,[P,K]]
[P,[P,[P,[P,K]]]]
[K,[K,[P,[P,K]]]]
[[P,K],[K,[P,K]]]
0
10
20
30
40
0
20
40

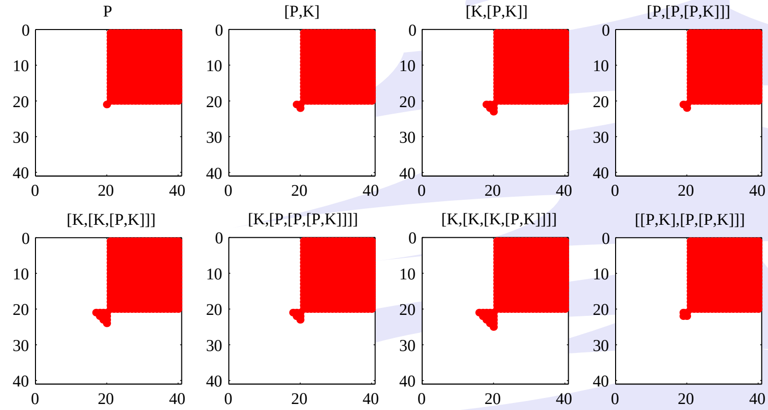


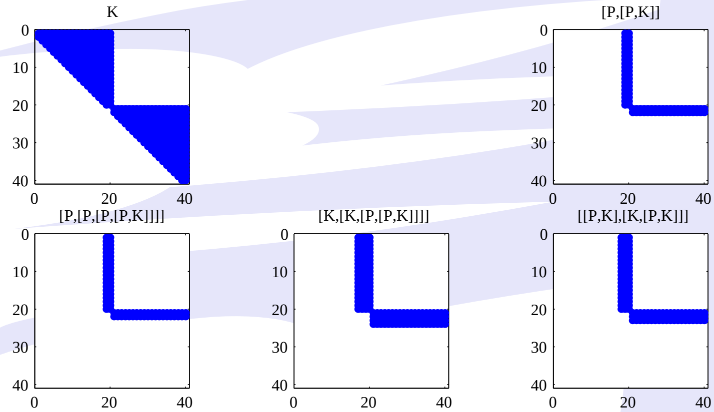
The following observations form the basis for an efficient divide-and-conquer algorithm to compute the exponential function in $\mathfrak{so}(n)$:

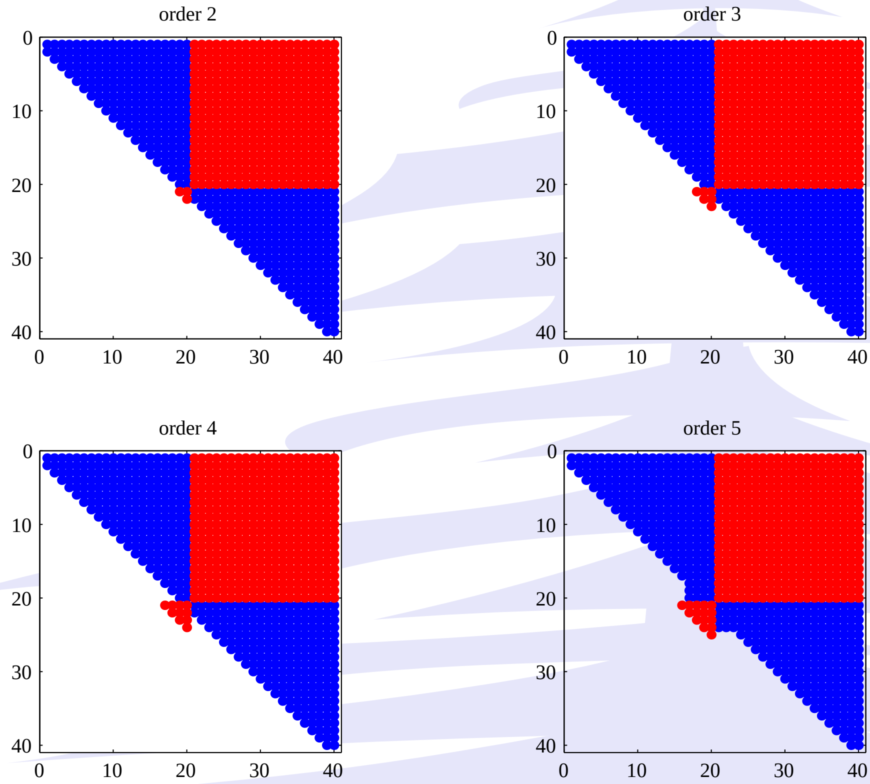
- All the commutators, hence $X(t)$ and $Y(t)$ (up to desired order) amount to $\mathcal{O}(1)$ flops.
- The exact exponential of X reduces to that of a small matrix and can be evaluated in $\mathcal{O}(1)$ flops.
- Y is a reducible matrix.
- The departure of Y from tridiagonal can be corrected in a small number of Givens rotations. Because of reducibility, we can act **separately** on each of the two components, hence the outcome is two tridiagonal matrices, of size $n_1 \times n_1$ and $n_2 \times n_2$, resp. Again, the cost is $\mathcal{O}(1)$ flops.

General matrices

Let $Z \in \mathfrak{gl}(n)$ or $\mathfrak{sl}(n)$ and suppose that we have already brought it to an upper Hessenberg form.






order 2
order 3
order 4
order 5
0
10
20
30
40

Cost of the algorithm

- Computation of the exponentials: We write

$$
\begin{aligned}
X_{[p]} &= X_1 t + \cdots + X_p t^p \\
&= Pt + \frac{1}{2}[P,K]t^2 + \cdots = \begin{bmatrix} O & B_1 \\ B_2 & O \end{bmatrix}
\end{aligned}
$$

where p is the order of the approximation. Roughly speaking,

order $p \Leftrightarrow$ fill-in of about $p \ll n$ elements.

Then:

$$
X_{[p]}^{2m} = \begin{bmatrix} V^m & O \\ O & W^m \end{bmatrix}, \quad X_{[p]}^{2m+1} = \begin{bmatrix} V^m & O \\ O & W^m \end{bmatrix} \begin{bmatrix} O & B_1 \\ B_2 & O \end{bmatrix}, \qquad V = B_1 B_2, W = B_2 B_1,
$$

and

$$
\exp(X_{[p]}) = \begin{bmatrix} C(V) & S(V)B_1 \\ S(W)B_2 & C(W) \end{bmatrix}, \qquad C(\theta) = \cosh \theta^{1/2}, \quad S(\theta) = \sinh \theta^{1/2}
$$

costs about $n_1 n_2 p \approx \frac{1}{4} p n^2$ for $n_1 = n_2 = n/2$.

If $n = s^2$, iterating on matrices of lower dimensions, the cost of computing all the exponentials reduces to

- $\frac{1}{2} p n^2$ on a serial machine
- $\frac{1}{3} p n^2$ on a parallel machine

- Computation of commutators: 3 "expensive" commutators for order 4, 7 for order 5, amounting to
  - $6 \sum_{s=1}^{\log_2(n/2)} 2^{\log_2(n/2)-s} 2^{3s} \approx n^3$ for order 4; $14 \sum_{s=1}^{\log_2(n/2)} 2^{\log_2(n/2)-s} 2^{3s} \approx 7/3 n^3$ for order 5 (serial machine)
  - $6 \sum_{s=1}^{\log_2(n/2)} 2^{3s} \approx 6/7 n^3$ for order 4; $14 \sum_{s=1}^{\log_2(n/2)} 2^{3s} \approx 2n^3$ for order 5 (parallel machine)

Error analysis

Although it is always true that

$$\|e^A\| \leq e^{\|A\|},$$

when dealing with error analysis of exponentials it is useful to consider the *logarithmic norm* μ of a matrix,

$$\mu(A) = \max\left\{\mu : \mu \text{ eigenvalue of } \frac{A + A^*}{2}\right\},$$

which is the derivative of $\|e^{tA}\|$: Set $\nu(t) = \|e^{tA}\|$. Then, by differentiation, it is easy verified that

$$\nu'(t) = \mu(A)\nu(t), \quad t > 0,$$

from which we deduce that $\nu(t) = e^{t\mu(A)}$, for all $t \geq 0$ $(\nu(0) = 1)$. Therefore, the growth or decay of $\|e^{tA}\|$ depends on whether the sign of $\mu(A)$ is **positive or negative**.
The logarithmic norm obeys the relations

$$\mu(A) \leq \|A\|,$$
$$\mu(A + B) \leq \mu(A) + \|B\|,$$
$$\|e^{\alpha A}\| \leq e^{\alpha\mu(A)},$$
$$\|e^{\alpha A}\| \geq e^{-\alpha\mu(-A)},$$

with $\alpha \geq 0$.

Example of an error bound

Assume that $Z = P_1 + \cdots + P_{m-1} + K_{m-1}$. Then, **for all** $t \geq 0$, it is true that

$$\|E_{2,m}(t)\| \;\leq\; \sum_{i=1}^{m-1} \Big(\frac{t^3}{3}\alpha_i + \frac{t^4}{4}\gamma_i(t) + \frac{t^5}{5}\delta_i(t) + \frac{t^6}{6}\eta_i(t) + \frac{t^7}{7}\theta_i(t) \Big) \times \mathrm{e}^{\sum_{k=0}^{i-1}(t\mu(P_k)+\frac{1}{2}t^2\|[P_k,K_k]\|)+\frac{1}{2}t^2\|[P_i,K_i]\|} \max\{\mathrm{e}^{t(\mu(P_i)+\mu(K_i))}, \mathrm{e}^{t\mu(P_i+K_i)}\},$$

where the P_i, K_i are the splitted terms of Z and $\alpha_i, \gamma_i(t), \ldots, \theta_i(t)$ are the analytic functions depending on the norm of commutators of P_i and K_i.

Similar bounds yield for other splittings.

For skew-symmetric matrices:
Given that $\exp(tZ)$ and $\exp(tP)$ and $\exp(tK)$ are orthogonal, one always has that the trivial global error bound

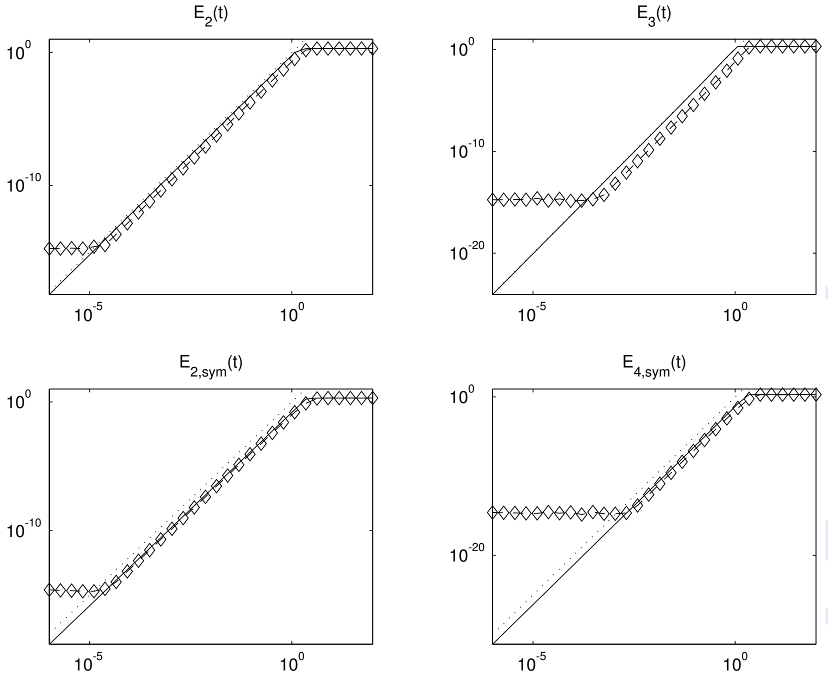
$$E(t) \leq \|\mathrm{e}^{tZ}\| + \|\mathrm{e}^{tP}\|\|\mathrm{e}^{tK}\| = 2$$

as a consequence of the triangle inequality. (The assertion is obviously true also for higher order of approximants and for a m-term splitting).
Moreover,

$$Z \in \mathfrak{so}(n) \Leftrightarrow \mu(Z) = 0,$$

hence the previous results mean essentially that:

For skew-symmetric matrices, the error is essentially local truncation error (until it reaches the trivial bound).



Estimated (solid line) and real errors (diamonds joined by dashed line) for 50×50 skew-symmetric matrices P, K versus $t \in [10^{-6}, 10^2]$.

Concluding remarks

- As long as $Z = P + K$, the GPD splittings are valid indipendently of the automorphism σ and can be thought as **inverse BCH** formulas:

$$\text{BCH formula:} \quad \exp(tB)\exp(tC) = \exp(t(B+C) + \frac{1}{2}t^2[B,C] + \cdots)$$

$$\text{inverse BCH:} \quad \exp(tZ) = \exp(X(t))\exp(Y(t))$$

  The advantage is that, if such automorphism σ determining P, K exists, then the factors have the properties discussed above.

- Symmetric GPD possible,

$$\exp(tZ) = \exp(X(t))\exp(Y(t))\exp(X(t))$$

$$\begin{aligned} X(t) &= \tfrac{1}{2}Pt + \tfrac{1}{24}[K,[P,K]]t^3 - (\tfrac{1}{1440}[K,[P,[P,[P,K]] \\ &\quad + \tfrac{1}{240}[K,[K,[K,[P,K]]]] + \tfrac{1}{360}[[P,K],[P,[P,K]]])t^5 + \cdots \\ Y(t) &= Kt + \tfrac{1}{24}[P,[P,K]]t^3 + (\tfrac{1}{1920}[P,[P,[P,[P,K]]]] \\ &\quad - \tfrac{13}{1440}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]])t^5 + \cdots \end{aligned}$$

and both $X(t)$ and $Y(t)$ expand in odd powers of t only.

Back
Close

What we got so far . . .

- Methods that map Lie algebras to the corresponding Lie group
- They are 'cheap'
  - $\mathcal{O}(n^3)$ with small constant, w.r.t. standard exponential routines.
- We can tackle most groups choosing the appropriate automorphisms
- High order is easily achievable
  - Once the sparse form is obtained, the commutators that give the order conditions are cheap to compute $\mathcal{O}(1)$
- Easy scaling and squaring:

$$\exp(Z) = Q \exp(\tilde{Z}) Q^\top, \qquad \tilde{Z} \text{ sparse}$$

Apply scaling and squaring to $\exp(\tilde{Z})$: Divide $\tilde{Z}$ by 2^j, compute splitting to desired order, apply to the vector(s) j-times. The exponentials$\times$vector are $\mathcal{O}(n^2)$.

Things that need to be explored further. . .

- Symmetric GPD
- Divide and conquer

Back
Close

- Stiff problems
  - $\exp(Z) = \exp(\frac{\mathrm{tr}(Z)}{n}) \exp(Z - \frac{\mathrm{tr}(Z)}{n} I)$
  - Commutators may introduce extra stiffness
  - Commutator-free methods, Exponential Time Differentiation (ETD) methods
- Different choices of automorphisms (or, simply P, K) that lead to different splittings
  - Example: Krylov. $\exp(A)\mathbf{v} \approx \beta V_m \exp(H_m)\mathbf{e}_1$.
    Choose $\mathrm{d}\sigma_{\mathbf{v}}(A) = (I - V_m V_m^\top)^\top A (I - V_m V_m^\top)$. Then

    $$A = P + K,$$

    has the property that

    $$\exp(K)\mathbf{v} = \mathbf{v}, \qquad \exp(P) = \beta V_m \exp(H_m)\mathbf{e}_1,$$

    hence,

    $$\beta V_m \exp(H_m)\mathbf{e}_1 = \exp(P)\exp(K)\mathbf{v}.$$

    The commutators 'destroy' the approximation (order theory recovered, worse approximation).
- Very large problems (for which $\mathcal{O}(n^3)$ is not feasible): can something *a la'* Krylov be done?

References

On the theory of GPD:

Munthe-Kaas, Quispel and Zanna (2001), 'Generalized polar decompositions on Lie groups with involutive automorphisms', *Found. Comp. Math.* **1**(3), 297–324.

Zanna (2000) 'Recurrence relations and convergence theory for the generalized polar decomposition on Lie groups', to appear in *Math. Comp.*

Lawson (1994), 'Polar and Ol'shanskii decompositions', *J. Reine Angew. Math.* **448**, 191–219.

GPD for the matrix exponential:

Zanna and Munthe-Kaas (2002), 'Generalized polar decompositions for the approximation of the matrix exponential', *SIAM J. Num. Anal.* **23**(3), 840–862.

Iserles and Zanna (2002), 'Efficient computation of the matrix exponential by generalized polar decompositions', to appear in *SIAM J. Num. Anal.*

SR for symplectic matrices:

Faßbender (2000), 'The parametrized SR algorithm for symplectic butterfly matrices', *Math. Comp.* **70**(236), 1515–1541.

Stability and error analysis:

Zanna (2001), 'Error analysis for exponential splitting based on Generalized Polar Decompositions', Report in Informatics 220, University of Bergen, Norway.

Iserles and Zanna (2003), 'On the spectra of certain matrices generated by involutive automorphisms', to appear in *SIAM J. Matrix Anal.*

Sheng (1993), 'Global error estimates for exponential splitting', *IMA J. Numer. Anal.*, **14**, 27–57.

General theory for computation of $\exp$:

Moler and van Loan (2003), 'Nineteen dubious ways to compute the matrix exponential — 25 years later', *SIAM Rev.*

Golub and van Loan (1989), 'Matrix Computation', John Hopkins, Baltimore.