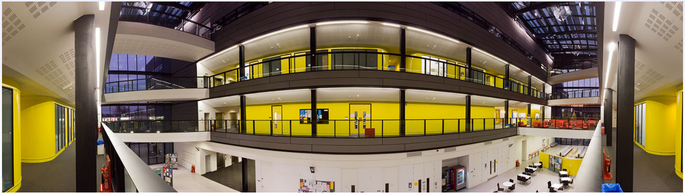
How and How Not to Compute the Exponential of a Matrix

Nick Higham
School of Mathematics
The University of Manchester

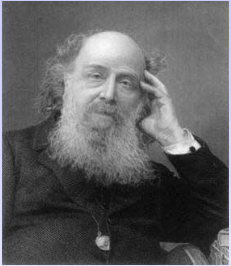
higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/



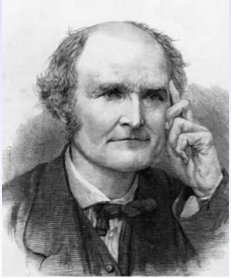
Outline

Cayley and Sylvester

- Term **"matrix"** coined in 1850 by James Joseph Sylvester, FRS (1814–1897).



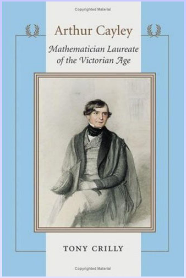
- **Matrix algebra** developed by Arthur Cayley, FRS (1821–1895).
  Memoir on the Theory of Matrices (1858).



Cayley and Sylvester on Matrix Functions

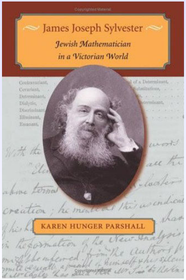
- Cayley considered matrix square roots in his 1858 memoir.

  Tony Crilly, *Arthur Cayley: Mathematician Laureate of the Victorian Age*, 2006.



- Sylvester (1883) gave first definition of $f(A)$ for general f.

  Karen Hunger Parshall, *James Joseph Sylvester. Jewish Mathematician in a Victorian World*, 2006.



Laguerre (1867):

En particulier, si nous définissons e^{X}, X étant un système d'ordre quelconque, comme étant la somme de la série

$$\Omega + \mathrm{X} + \frac{\mathrm{X}^2}{1.2} + \frac{\mathrm{X}^3}{1.2.3} + \ldots,$$

e^{X} sera une fonction de la variable X; mais il est à remarquer qu'en général on n'aura pas

$$e^{\mathrm{X}}.e^{\mathrm{Y}} = e^{\mathrm{X}+\mathrm{Y}}.$$

Peano (1888):

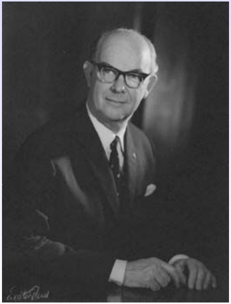
$$\mathbf{x} = \left[1 + \mathbf{R}t + \frac{1}{2!}(\mathbf{R}t)^2 + \cdots\right]\mathbf{a},$$

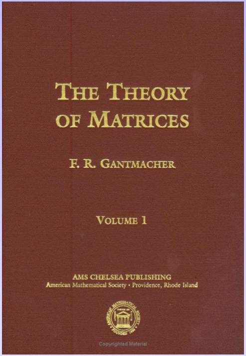
ou, en posant $e^{\mathbf{R}} = 1 + \mathbf{R} + \frac{1}{2!}\mathbf{R}^2 + \cdots,$

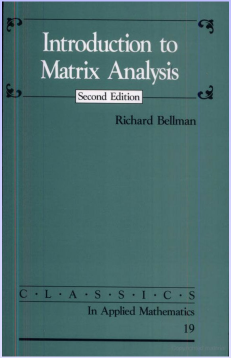
$$\mathbf{x} = e^{\mathbf{R}t}\,\mathbf{a}.$$

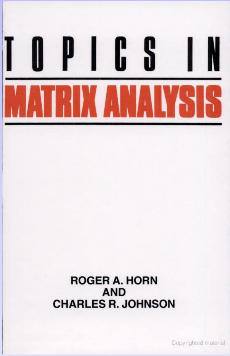
Matrices in Applied Mathematics

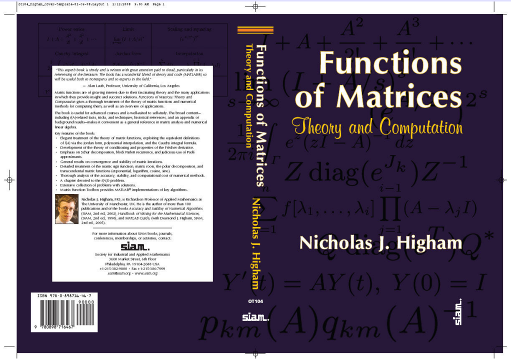
- Frazer, Duncan & Collar, Aerodynamics Division of NPL: aircraft flutter, matrix structural analysis.
- **Elementary Matrices & Some Applications to Dynamics and Differential Equations, 1938**. Emphasizes importance of e^A.
- Arthur Roderick Collar, FRS (1908–1986): *"First book to treat matrices as a branch of applied mathematics"*.




THE THEORY OF MATRICES
F. R. GANTMACHER
VOLUME 1
AMS CHELSEA PUBLISHING
American Mathematical Society • Providence, Rhode Island
Copyrighted Material


Introduction to Matrix Analysis
Second Edition
Richard Bellman
C·L·A·S·S·I·C·S
In Applied Mathematics
19


TOPICS IN
MATRIX ANALYSIS
ROGER A. HORN
AND
CHARLES R. JOHNSON
Copyrighted material


Functions of Matrices
Theory and Computation
Nicholas J. Higham
OT104
siam.

Formulae

$A \in \mathbb{C}^{n\times n}$:

Power series $I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$	**Limit** $\lim_{s\to\infty} (I + A/s)^s$	**Scaling and squaring** $(e^{A/2^s})^{2^s}$
Cauchy integral $\frac{1}{2\pi i}\int_\Gamma e^z (zI - A)^{-1}\, dz$	**Jordan form** $Z\mathrm{diag}(e^{J_k})Z^{-1}$	**Interpolation** $\sum_{i=1}^{n} f[\lambda_1,\ldots,\lambda_i]\prod_{j=1}^{i-1}(A - \lambda_j I)$
Differential system $Y'(t) = AY(t),\ \ Y(0) = I$	**Schur form** $Q\mathrm{diag}(e^T)Q^*$	**Padé approximation** $p_{km}(A)q_{km}(A)^{-1}$

Exponential of Sum

Theorem

For $A, B \in \mathbb{C}^{n\times n}$, $e^{(A+B)t} = e^{At}e^{Bt}$ for all t if and only if $AB = BA$.

Theorem (Wermuth)

Let $A, B \in \mathbb{C}^{n\times n}$ have algebraic elements and let $n \geq 2$. Then $e^A e^B = e^B e^A$ if and only if $AB = BA$.

Exponential of Sum

Theorem

For $A, B \in \mathbb{C}^{n\times n}$, $e^{(A+B)t} = e^{At}e^{Bt}$ for all t if and only if $AB = BA$.

Theorem (Wermuth)

Let $A, B \in \mathbb{C}^{n\times n}$ have algebraic elements and let $n \geq 2$. Then $e^A e^B = e^B e^A$ if and only if $AB = BA$.

Theorem

Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then $e^{A\oplus B} = e^A \otimes e^B$, where $A \oplus B = A \otimes I_m + I_n \otimes B$.

Taylor Series

Theorem (Suzuki)

For $A \in \mathbb{C}^{n\times n}$, let

$$T_{r,s} = \left[\sum_{i=0}^{r} \frac{1}{i!}\left(\frac{A}{s}\right)^i\right]^s.$$

Then

$$\|e^A - T_{r,s}\| \le \frac{\|A\|^{r+1}}{s^r(r+1)!} e^{\|A\|}$$

and $\lim_{r\to\infty} T_{r,s}(A) = \lim_{s\to\infty} T_{r,s}(A) = e^A$.

Outline

1. History & Properties
2. Applications
3. Methods

Application: Control Theory

Convert **continuous-time system**

$$\begin{aligned}\frac{dx}{dt} &= Fx(t) + Gu(t),\\ y &= Hx(t) + Ju(t),\end{aligned}$$

to **discrete-time state-space system**

$$\begin{aligned}x_{k+1} &= Ax_k + Bu_k,\\ y_k &= Hx_k + Ju_k.\end{aligned}$$

Have

$$A = e^{F\tau}, \qquad B = \left(\int_0^\tau e^{Ft}\,dt\right) G,$$

where τ is the sampling period.
MATLAB Control System Toolbox: `c2d` and `d2c`.

Phi Functions: Definition

$$\varphi_0(z) = e^z, \quad \varphi_1(z) = \frac{e^z - 1}{z}, \quad \varphi_2(z) = \frac{e^z - 1 - z}{z^2}, \ldots$$

$$\varphi_{k+1}(z) = \frac{\varphi_k(z) - 1/k!}{z}.$$

$$\boxed{\varphi_k(z) = \sum_{j=0}^{\infty} \frac{z^j}{(j+k)!}.}$$

Phi Functions: Solving DEs

$y \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$.

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0.$$

Phi Functions: Solving DEs

$y \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$.

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0.$$

$$\frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t\,\varphi_1(tA)b.$$

Phi Functions: Solving DEs

$y \in \mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$.

$$\frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At} y_0.$$

$$\frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t\,\varphi_1(tA)b.$$

$$\frac{dy}{dt} = Ay + ct, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t^2\,\varphi_2(tA)c.$$

$$\vdots$$

Exponential Integrators

Consider

$$y' = Ly + N(y).$$

$N(y(t)) \approx N(y(0))$ implies

$$y(t) \approx e^{tL} y_0 + t\varphi_1(tL)N(y(0)).$$

Exponential Euler method:

$$\boxed{y_{n+1} = e^{hL} y_n + h\varphi_1(hL)N(y_n).}$$

Lawson (1967); recent resurgence.

The Average Eye

First order character of optical system characterized by **transference matrix**

$$T = \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{5\times 5},$$

where $S \in \mathbb{R}^{4\times 4}$ is **symplectic**:

$$S^T J S = J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}.$$

Average $m^{-1}\sum_{i=1}^{m} T_i$ is not a transference matrix.

Harris (2005) proposes the average $\exp(m^{-1}\sum_{i=1}^{m}\log(T_i))$.

Beyond Matrices

- GluCat library: generic library of C++ templates for universal Clifford algebras: exp, log, square root, trig functions.
  `http://glucat.sourceforge.net`.
- Group exponential of a diffeomorphism in computational anatomy to study variability among medical images (Bossa et al., 2008).

Outline

1. History & Properties
2. Applications
3. Methods

Cayley–Hamilton Theorem

Theorem (Cayley, 1857)

If $A, B \in \mathbb{C}^{n\times n}$, $AB = BA$, *and* $f(x, y) = \det(xA - yB)$ *then* $f(B, A) = 0$.

- $p(t) = \det(tI - A)$ implies $p(A) = 0$.
- $A^n = \sum_{k=0}^{n-1} c_n A^k$.
- $e^A = \sum_{k=0}^{n-1} d_n A^k$.

Walz's Method

Walz (1988) proposed computing

$$C_k = (I + 2^{-k}A)^{2^k}$$

with Richardson extrapolation to accelerate cgce of the C_k.

Numerically unstable in practice (Parks, 1994).

Diagonalization (1)

$A = Z\mathrm{diag}(\lambda_i)Z^{-1}$ implies $f(A) = Z\mathrm{diag}(f(\lambda_i))Z^{-1}$.

But

- Z may be ill conditioned ($\kappa(Z) = \|Z\|\|Z^{-1}\| \gg 1$).
- A may not be diagonalizable.

Diagonalization (2)

```
>> A = [3 -1; 1 1]; X = funm_ev(A,@exp)
X =
   14.7781   -7.3891
    7.3891         0
>> norm(X - expm(A))/norm(expm(A))
ans =  1.3431e-009
>>  expm_cond(A)
ans =  3.4676

>> [Z,D]=eig(A)
Z =                        D =
    0.7071    0.7071           2.0000         0
    0.7071    0.7071                0    2.0000
```

Scaling and Squaring Method

- $B \leftarrow A/2^s$ so $\|B\|_\infty \approx 1$
- $r_m(B) = [m/m]$ Padé approximant to e^B
- $X = r_m(B)^{2^s} \approx e^A$

- Originates with **Lawson (1967)**.
- **Ward (1977)**: algorithm, with rounding error analysis and a posteriori error bound.
- **Moler & Van Loan (1978)**: give backward error analysis allowing choice of s and m.
- **H (2005)**: sharper analysis giving optimal s and m. MATLAB's `expm`, Mathematica, NAG Library Mark 22.

Padé Approximants r_m to e^x

$r_m(x) = p_m(x)/q_m(x)$ known explicitly:

$$p_m(x) = \sum_{j=0}^{m} \frac{(2m-j)!\,m!}{(2m)!\,(m-j)!} \frac{x^j}{j!}$$

and $q_m(x) = p_m(-x)$.

Henri Padé
1863–1953

Error satisfies

$$e^x - r_m(x) = (-1)^m \frac{(m!)^2}{(2m)!(2m+1)!} x^{2m+1} + O(x^{2m+2}).$$

Truncation Analysis

$$h_{2m+1}(X) := \log(e^{-X} r_m(X)) = \sum_{k=2m+1}^{\infty} c_k X^k.$$

Then $r_m(X) = e^{X+h_{2m+1}(X)}$. Hence

$$\boxed{r_m(2^{-s}A)^{2^s} = e^{A+2^s h_{2m+1}(2^{-s}A)} =: e^{A+\Delta A}.}$$

Want $\|\Delta A\|/\|A\| \le u$.

- Moler & Van Loan (1978): a priori bound for h_{2m+1}; $m = 6$, $\|2^{-s}A\| \le 1/2$ in MATLAB.
- H (2005): sharp normwise bound using symbolic arithmetic and high precision. Choose (s, m) to minimize computational cost.

Scaling & Squaring Algorithm (H, 2005)

m	3	5	7	9	13
θ_m	0.015	0.25	0.95	2.1	5.4

for $m = [3\ 5\ 7\ 9\ 13]$
    if $\|A\|_1 \le \theta_m$, $X = r_m(A)$, quit, end
end

Scaling & Squaring Algorithm (H, 2005)

m	3	5	7	9	13
θ_m	0.015	0.25	0.95	2.1	5.4

for $m = [3\ 5\ 7\ 9\ 13]$
    if $\|A\|_1 \le \theta_m$, $X = r_m(A)$, quit, end
end

$A \leftarrow A/2^s$ with $s \ge 0$ minimal s.t. $\|A/2^s\|_1 \le \theta_{13} = 5.4$

Scaling & Squaring Algorithm (H, 2005)

m	3	5	7	9	13
θ_m	0.015	0.25	0.95	2.1	5.4

for $m = [3\ 5\ 7\ 9\ 13]$
    if $\|A\|_1 \le \theta_m$, $X = r_m(A)$, quit, end
end

$A \leftarrow A/2^s$ with $s \ge 0$ minimal s.t. $\|A/2^s\|_1 \le \theta_{13} = 5.4$

$A_2 = A^2$, $A_4 = A_2^2$, $A_6 = A_2 A_4$

$$U = A\big[A_6(b_{13}A_6 + b_{11}A_4 + b_9A_2) + b_7A_6 + b_5A_4 + b_3A_2 + b_1 I\big]$$

$$V = A_6(b_{12}A_6 + b_{10}A_4 + b_8A_2) + b_6A_6 + b_4A_4 + b_2A_2 + b_0 I$$

Solve $(-U + V)r_{13} = U + V$ for r_{13}.

$X = {r_{13}}^{2^s}$ by repeated squaring.

Example

$$A = \begin{bmatrix} 1 & b \\ 0 & -1 \end{bmatrix}, \qquad e^A = \begin{bmatrix} e & \frac{b}{2}(e - e^{-1}) \\ 0 & e^{-1} \end{bmatrix}.$$

b	`expm`(A)	s	`funm`(A)
10^3	1.7e-15	8	1.9e-16
10^4	1.8e-13	11	3.8e-20
10^5	7.5e-13	15	1.2e-16
10^6	1.3e-11	18	2.0e-16
10^7	7.2e-11	21	1.6e-16
10^8	3.0e-12	25	1.3e-16

Overscaling

Kenney & Laub (1998); Dieci & Papini (2000).

A large $\|A\|$ causes a larger than necessary s to be chosen, with a harmful effect on accuracy.

Overscaling

Kenney & Laub (1998); Dieci & Papini (2000).

A large $\|A\|$ causes a larger than necessary s to be chosen, with a harmful effect on accuracy.

x	$e^x - (1+x)$	$e^x - (1+x/2)^2$
9.9e-9	2.2e-16	6.7e-16
8.9e-9	0	6.7e-16

Overscaling

Kenney & Laub (1998); Dieci & Papini (2000).

A large $\|A\|$ causes a larger than necessary s to be chosen, with a harmful effect on accuracy.

x	$e^x - (1+x)$	$e^x - (1+x/2)^2$
9.9e-9	2.2e-16	6.7e-16
8.9e-9	0	6.7e-16

$$\exp\left(\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\right) = \begin{bmatrix} e^{A_{11}} & \int_0^1 e^{A_{11}(1-s)} A_{12} e^{A_{22}s}\,ds \\ 0 & e^{A_{22}} \end{bmatrix}.$$

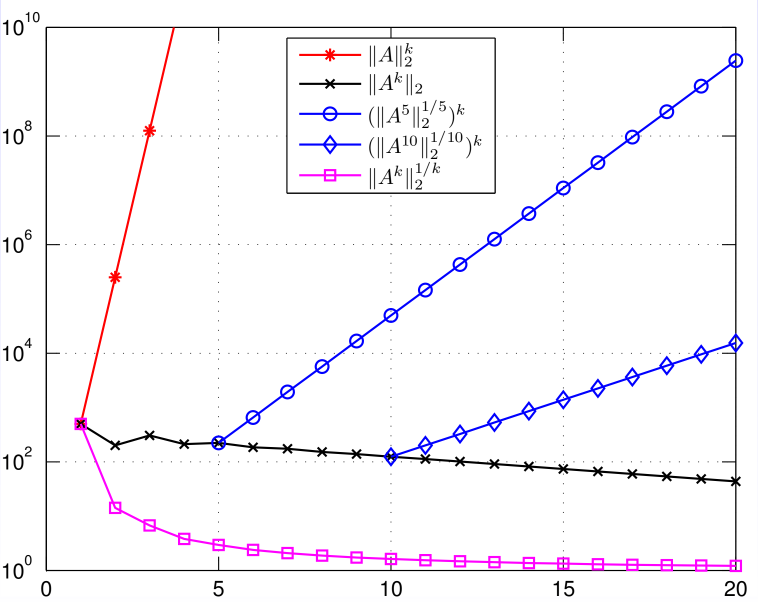
Insight

- Non-normality implies $\rho(A) \ll \|A\|$.
- Note that

$$\rho(A) \le \|A^k\|^{1/k} \le \|A\|, \qquad k = 1\colon \infty.$$

  and $\lim_{k\to\infty} \|A^k\|^{1/k} = \rho(A)$.
- Use $\|A^k\|^{1/k}$ instead of $\|A\|$ in the truncation bounds.

$$A = \begin{bmatrix} 0.9 & 500 \\ 0 & -0.5 \end{bmatrix}.$$



New Alg (Al-Mohy & H, 2009)

- Truncation bounds use $\|A^k\|^{1/k}$ rather than $\|A\|$, leading to major benefits in speed and accuracy.

- Special treatment of triangular matrices leads to more stable squaring phase.

Example

$$A = \begin{bmatrix} 1 & b \\ 0 & -1 \end{bmatrix}, \quad A^{2k} \equiv I, \quad A^{2k+1} \equiv A.$$

New alg selects $s = 0$, $m = 9$ for all b.

b	expm	expm_new	cost: $\frac{\texttt{expm}}{\texttt{expm\_new}}$
10^3	1.7e-15	0	2.8
10^4	1.8e-13	0	3.4
10^5	7.5e-13	0	4.2
10^6	1.3e-11	2.0e-16	4.8
10^7	7.2e-11	0	5.4
10^8	3.0e-12	1.3e-16	6.2
10^{17}	1.5e-1	1.4e-16	12.2

Summary of New Algorithm

- Overscaling problem alleviated.
- New algorithm is no slower than **`expm`**, potentially faster, potentially more accurate.

Summary of New Algorithm

- Overscaling problem alleviated.
- New algorithm is no slower than `expm`, potentially faster, potentially more accurate.

Stability of squaring phase for full matrices remains an **open problem**.

The $e^A B$ Problem

Exploit, for integer s,

$$e^A B = (e^{s^{-1}A})^s B = \underbrace{e^{s^{-1}A} e^{s^{-1}A} \cdots e^{s^{-1}A}}_{s \text{ times}} B.$$

Choose s so $T_m(s^{-1}A) = \sum_{j=0}^{m} \dfrac{(s^{-1}A)^j}{j!} \approx e^{s^{-1}A}$. Then

$$B_{i+1} = T_m(s^{-1}A)B_i, \quad i = 0\colon s-1, \qquad B_0 = B$$

yields $B_s \approx e^A B$.

The $e^A B$ Problem

Exploit, for integer s,

$$e^A B = (e^{s^{-1}A})^s B = \underbrace{e^{s^{-1}A} e^{s^{-1}A} \cdots e^{s^{-1}A}}_{s \text{ times}} B.$$

Choose s so $T_m(s^{-1}A) = \sum_{j=0}^{m} \dfrac{(s^{-1}A)^j}{j!} \approx e^{s^{-1}A}$. Then

$$B_{i+1} = T_m(s^{-1}A)B_i, \quad i = 0\colon s-1, \qquad B_0 = B$$

yields $B_s \approx e^A B$.

Choose s and m using $\|A^k\|^{1/k}$ approach.

Algorithm for $F = e^{tA}B$ (Al-Mohy & H, 2010)

$\mu = \mathrm{trace}(A)/n$

$A = A - \mu I$

$[m, s] = \mathrm{parameters}(tA)$

$F = B,\ \eta = e^{t\mu/s}$

for $i = 1\colon s$

  $c_1 = \|B\|_\infty$

  for $j = 1\colon m$

    $B = tAB/(sj),\ c_2 = \|B\|_\infty$

    $F = F + B$

    if $c_1 + c_2 \le \mathrm{tol}\|F\|_\infty$, quit, end

    $c_1 = c_2$

  end

  $F = \eta F,\ B = F$

end

The Sixth Dubious Way

Moler & Van Loan (1978, 2003)

METHOD 6. SINGLE STEP O.D.E. METHODS. Two of the classical techniques for the solution of differential equations are the fourth order Taylor and Runge–Kutta methods with fixed step size. For our particular equation they become

$$x_{j+1} = \left(I + hA + \cdots + \frac{h^4}{4!}A^4\right)x_j = T_4(hA)x_j$$

and

$$x_{j+1} = x_j + \tfrac{1}{6}k_1 + \tfrac{1}{3}k_2 + \tfrac{1}{3}k_3 + \tfrac{1}{6}k_4,$$

where $k_1 = hAx_j$, $k_2 = hA(x_j + \frac{1}{2}k_1)$, $k_3 = hA(x_j + \frac{1}{2}k_2)$, and $k_4 = hA(x_j + k_3)$. A little manipulation reveals that in this case, the two methods would produce identical results were it not for roundoff error. As long as the step size is fixed, the matrix $T_4(hA)$ need be computed just once and then x_{j+1} can be obtained from x_j with just one matrix-vector multiplication. The standard Runge–Kutta method would require 4 such multiplications per step.

Let us consider $x(t)$ for one particular value of t, say $t = 1$. If $h = 1/m$, then

$$x(1) = x(mh) \simeq x_m = [T_4(hA)]^m x_0.$$

Advantages of New Algorithm

Versus **one-step ODE integrators**:

- Fully exploits the **linearity** of the ODE.
- **Variable order**, up to $m = 55$.
- **Backward error** based; ODE integrator controls local (forward) errors.
- **Overscaling** avoided.

Versus **Krylov methods**:

- Very competitive in cost and storage.
- Cost dominated by matrix–vector multiplications.
- Black box: no need to choose/tune parameters.

Frechét Derivative

$$f(A+E) - f(A) - L(A,E) = o(\|E\|).$$

$$L_{\exp}(A,E) = \int_0^1 e^{A(1-s)} E e^{As}\, ds.$$

- Method based on

$$f\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right) = \begin{bmatrix} f(X) & L(X,E) \\ 0 & f(X) \end{bmatrix}.$$

- Kenney & Laub (1998): Kronecker–Sylvester alg, Padé of $\tanh(x)/x$: **$538n^3$** (complex) flops.
- **Al-Mohy & H (2009)**: e^A and $L(A,E)$ in only **$48n^3$** flops.

In Conclusion

- Growing number of applications of $f(A)$.
- $f(A)$ algorithms ready for library deployment.
- Need better understanding of conditioning of $f(A)$.
- How to exploit structure?
- More work needed on $f(A)b$ problem.

References I

A. H. Al-Mohy and N. J. Higham.
Computing the Fréchet derivative of the matrix exponential, with an application to condition number estimation.
SIAM J. Matrix Anal. Appl., 30(4):1639–1657, 2009.

A. H. Al-Mohy and N. J. Higham.
A new scaling and squaring algorithm for the matrix exponential.
SIAM J. Matrix Anal. Appl., 31(3):970–989, 2009.

References II

A. H. Al-Mohy and N. J. Higham.
Computing the action of the matrix exponential, with an application to exponential integrators.
MIMS EPrint 2010.30, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Mar. 2010.
23 pp.

References III

R. Bellman.
Introduction to Matrix Analysis.
McGraw-Hill, New York, second edition, 1970. xxiii+403 pp. Reprinted by Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997. ISBN 0-89871-399-4.

M. Bossa, E. Zacur, and S. Olmos.
Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation.
In *Computer Vision and Pattern Recognition Workshops, 2008 (CVPRW '08)*, pages 1–8. IEEE Computer Society, 2008.

References IV

T. Crilly.
Cayley's anticipation of a generalised Cayley–Hamilton theorem.
Historia Mathematica, 5:211–219, 1978.

T. Crilly.
Arthur Cayley: Mathematician Laureate of the Victorian Age.
Johns Hopkins University Press, Baltimore, MD, USA, 2006.
ISBN 0-8018-8011-4.
xxi+610 pp.

References V

L. Dieci and A. Papini.
Padé approximation for the exponential of a block triangular matrix.
Linear Algebra Appl., 308:183–202, 2000.

R. A. Frazer, W. J. Duncan, and A. R. Collar.
Elementary Matrices and Some Applications to Dynamics and Differential Equations.
Cambridge University Press, Cambridge, UK, 1938.
xviii+416 pp.
1963 printing.

References VI

Glucat: Generic library of universal Clifford algebra templates.
|http://glucat.sourceforge.net/|.

N. J. Higham.
The Matrix Function Toolbox.
http://www.ma.man.ac.uk/~higham/mftoolbox.

References VII

N. J. Higham.
Functions of Matrices: Theory and Computation.
Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
ISBN 978-0-898716-46-7.
xx+425 pp.

N. J. Higham and A. H. Al-Mohy.
Computing matrix functions.
Acta Numerica, 19:159–208, 2010.

References VIII

- N. J. Higham and F. Tisseur.
  A block algorithm for matrix 1-norm estimation, with an application to 1-norm pseudospectra.
  SIAM J. Matrix Anal. Appl., 21(4):1185–1201, 2000.

- R. A. Horn and C. R. Johnson.
  Topics in Matrix Analysis.
  Cambridge University Press, Cambridge, UK, 1991.
  ISBN 0-521-30587-X.
  viii+607 pp.

References IX

- C. S. Kenney and A. J. Laub.
  A Schur–Fréchet algorithm for computing the logarithm and exponential of a matrix.
  SIAM J. Matrix Anal. Appl., 19(3):640–663, 1998.

- C. B. Moler and C. F. Van Loan.
  Nineteen dubious ways to compute the exponential of a matrix.
  SIAM Rev., 20(4):801–836, 1978.

- C. B. Moler and C. F. Van Loan.
  Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later.
  SIAM Rev., 45(1):3–49, 2003.

References X

M. J. Parks.
A Study of Algorithms to Compute the Matrix Exponential.
PhD thesis, Mathematics Department, University of California, Berkeley, CA, USA, 1994.

K. H. Parshall.
James Joseph Sylvester. Jewish Mathematician in a Victorian World.
Johns Hopkins University Press, Baltimore, MD, USA, 2006.
ISBN 0-8018-8291-5.
xiii+461 pp.

References XI

G. Walz.
Computing the matrix exponential and other matrix functions.
J. Comput. Appl. Math., 21:119–123, 1988.

E. M. E. Wermuth.
Two remarks on matrix exponentials.
Linear Algebra Appl., 117:127–132, 1989.