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

1. History & Properties
2. Applications
3. Methods
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 $f$.

Laguerre (1867):

\[
\text{En particulier, si nous définissons } e^x, \text{ } X \text{ étant un système d'ordre quelconque, comme étant la somme de la série}
\]
\[
\Omega + x + \frac{x^2}{1 \cdot 2} + \frac{x^3}{1 \cdot 2 \cdot 3} + \ldots,
\]
\[
e^x \text{ sera une fonction de la variable } X; \text{ mais il est à remarquer qu'en général on n'aura pas}
\]
\[
e^x \cdot e^y = e^{x+y}.
\]

Peano (1888):

\[
x = \left[ 1 + R \cdot t + \frac{1}{2!} \cdot (R \cdot t)^2 + \cdots \right] a,
\]
\[
\text{ou, en posant } e^R = 1 + R + \frac{1}{2!} \cdot R^2 + \cdots,
\]
\[
x = e^{R \cdot t} a.
\]
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".
### Formulae

\[ A \in \mathbb{C}^{n \times n}: \]

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

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 $AB = BA$. 
Properties (1)

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 $A \oplus B = A \otimes I_m + I_n \otimes B$. 
Properties (2)

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

\[
\| e^A - T_{r,s} \| \leq \frac{\| A \|^{r+1}}{s^r (r + 1)!} e^{\| A \|}
\]

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

\[
\frac{dx}{dt} = Fx(t) + Gu(t),
\]
\[
y = Hx(t) + Ju(t),
\]

to discrete-time state-space system

\[
x_{k+1} = Ax_k + Bu_k,
\]
\[
y_k = Hx_k + Ju_k.
\]

Have

\[
A = e^{F\tau}, \quad B = \left( \int_{0}^{\tau} e^{Ft} dt \right) G,
\]

where \( \tau \) is the sampling period.

MATLAB Control System Toolbox: \texttt{c2d} and \texttt{d2c}. 
Ψ Functions: Definition

$\psi_0(z) = e^z, \quad \psi_1(z) = \frac{e^z - 1}{z}, \quad \psi_2(z) = \frac{e^z - 1 - z}{z^2}, \ldots$

$\psi_{k+1}(z) = \frac{\psi_k(z) - 1/k!}{z}$

$\psi_k(z) = \sum_{j=0}^{\infty} \frac{z^j}{(j + k)!}.$
Psi Functions: Solving DEs

\[ y \in \mathbb{C}^n, \quad A \in \mathbb{C}^{n \times n}. \]

\[ \frac{dy}{dt} = Ay, \quad y(0) = y_0 \quad \Rightarrow \quad y(t) = e^{At}y_0. \]
Psi Functions: Solving DEs

\[ y \in \mathbb{C}^n, \quad A \in \mathbb{C}^{n \times n}. \]

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

\[ \frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t \psi_1(tA)b. \]
Psi Functions: Solving DEs

\[ y \in \mathbb{C}^n, \; A \in \mathbb{C}^{n \times n}. \]

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

\[
\frac{dy}{dt} = Ay + b, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t\psi_1(tA)b.
\]

\[
\frac{dy}{dt} = Ay + ct, \quad y(0) = 0 \quad \Rightarrow \quad y(t) = t^2\psi_2(tA)c.
\]

\vdots
Consider

\[ y' = Ly + N(y). \]

\[ N(y(t)) \approx N(y(0)) \implies y(t) \approx e^{tL}y_0 + t\psi_1(tL)N(y(0)). \]

**Exponential Euler method:**

\[ y_{n+1} = e^{hL}y_n + h\psi_1(hL)N(y_n). \]

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:

\[ S^T J S = J, \]

where \( J = \begin{bmatrix} 0 & l_2 \\ -l_2 & 0 \end{bmatrix} \).

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

Harris (2005) proposes the average \( \exp(m^{-1} \sum_{i=1}^{m} \log(T_i)) \).
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: $S^TJS = J$, where $J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}$.

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

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

For Hermitian pos def $A$ and $B$, Arsigny et al. (2007) define the log-Euclidean mean

$$E(A, B) = \exp\left(\frac{1}{2} (\log(A) + \log(B))\right)$$. 
Beyond Matrices

- GluCat library: generic library of C++ templates for universal Clifford algebras: exp, log, square root, trig functions.
- Group exponential of a diffeomorphism in computational anatomy to study variability among medical images (Bossa et al., 2008).
Note that $e^T$ is a power series in $T$. This means that a wide variety of methods in linear algebra can also be used to evaluate $e^T$. . . . brute force evaluation of the power series, . . . matrix decomposition methods or polynomial methods based on the Cayley-Hamilton theorem . . .
Note that $e^T$ is a power series in $T$. This means that a wide variety of methods in linear algebra can also be used to evaluate $e^T$. . . . brute force evaluation of the power series, . . . matrix decomposition methods or polynomial methods based on the Cayley-Hamilton theorem . . .

Since evaluation of functions of matrices may be fraught with difficulties (such as roundoff and truncation errors, ill conditioning, near confluence of eigenvalues, etc.), there is a distinct advantage in having a rich class of solution techniques available for finding $e^T$. If one method fails to find an accurate answer, one can always fall back on a different method.
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)

\[ A = Z \text{diag}(\lambda_i) Z^{-1} \text{ implies } f(A) = Z \text{diag}(f(\lambda_i)) Z^{-1}. \]

But
- \( Z \) may be ill conditioned (\( \kappa(Z) = \|Z\|\|Z^{-1}\| \gg 1 \)).
- \( A \) may not be diagonalizable.
>> 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))  
an = 1.3431e-009

>> expm_cond(A)  
an = 3.4676

>> [Z,D]=eig(A)  
Z =
  0.7071  0.7071
  0.7071  0.7071
D =
  2.0000  0
     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).
- 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:

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

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

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

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

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

\[ r_m(2^{-s} A)^{2^s} = e^{A + 2^s h_{2m+1}(2^{-s} A)} = e^{A + \Delta A}. \]

Want \( \|\Delta A\|/\|A\| \leq u \).

- Moler & Van Loan (1978): a priori bound for \( h_{2m+1} \);
  \( m = 6, \|2^{-s} A\| \leq 1/2 \) 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)

<table>
<thead>
<tr>
<th>$m$</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_m$</td>
<td>0.015</td>
<td>0.25</td>
<td>0.95</td>
<td>2.1</td>
<td>5.4</td>
</tr>
</tbody>
</table>

for $m = [3 \ 5 \ 7 \ 9 \ 13]$

if $\|A\|_1 \leq \theta_m$, $X = r_m(A)$, quit, end

end
Scaling & Squaring Algorithm (H, 2005)

<table>
<thead>
<tr>
<th>$m$</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_m$</td>
<td>0.015</td>
<td>0.25</td>
<td>0.95</td>
<td>2.1</td>
<td>5.4</td>
</tr>
</tbody>
</table>

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

$A \leftarrow A/2^s$ with $s \geq 0$ minimal s.t. $\|A/2^s\|_1 \leq \theta_{13} = 5.4$
### Scaling & Squaring Algorithm (H, 2005)

<table>
<thead>
<tr>
<th>$m$</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_m$</td>
<td>0.015</td>
<td>0.25</td>
<td>0.95</td>
<td>2.1</td>
<td>5.4</td>
</tr>
</tbody>
</table>

for $m = [3 \ 5 \ 7 \ 9 \ 13]$

if $\|A\|_1 \leq \theta_m$, $X = r_m(A)$, quit, end

end

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

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

$U = A \left[ A_6(b_{13} A_6 + b_{11} A_4 + b_9 A_2) + b_7 A_6 + b_5 A_4 + b_3 A_2 + b_1 I \right]$

$V = A_6(b_{12} A_6 + b_{10} A_4 + b_8 A_2) + b_6 A_6 + b_4 A_4 + b_2 A_2 + b_0 I$

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

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

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

<table>
<thead>
<tr>
<th>( b )</th>
<th>( \expm(A) )</th>
<th>( s )</th>
<th>( \expm(A)^\dagger )</th>
<th>( s )</th>
<th>( \funm(A) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^3 )</td>
<td>1.7e-15</td>
<td>8</td>
<td>1.9e-16</td>
<td>0</td>
<td>1.9e-16</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>1.8e-13</td>
<td>11</td>
<td>7.6e-20</td>
<td>0</td>
<td>3.8e-20</td>
</tr>
<tr>
<td>( 10^5 )</td>
<td>7.5e-13</td>
<td>15</td>
<td>1.2e-16</td>
<td>0</td>
<td>1.2e-16</td>
</tr>
<tr>
<td>( 10^6 )</td>
<td>1.3e-11</td>
<td>18</td>
<td>2.0e-16</td>
<td>0</td>
<td>2.0e-16</td>
</tr>
<tr>
<td>( 10^7 )</td>
<td>7.2e-11</td>
<td>21</td>
<td>1.6e-16</td>
<td>0</td>
<td>1.6e-16</td>
</tr>
<tr>
<td>( 10^8 )</td>
<td>3.0e-12</td>
<td>25</td>
<td>1.3e-16</td>
<td>0</td>
<td>1.3e-16</td>
</tr>
</tbody>
</table>

For \( b = 10^8 \), \( r_m(x)^{2^{25}} \approx \left( (1 + \frac{1}{2}x)/(1 - \frac{1}{2}x) \right)^{2^{25}} \) with
\[ x = \pm 2^{-25} \approx \pm 10^{-8}. \]
Overscaling


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

\[
\exp \left( \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \right) = \begin{bmatrix} e^{A_{11}} & \int_0^1 e^{A_{11}(1-s)} A_{12} e^{A_{22}s} \, ds \\ 0 & e^{A_{22}} \end{bmatrix}.
\]
Al-Mohy & H (2009):
Existing method based on analysis in terms of $\|A\|$. Why not instead use $\|A^k\|^{1/k}$?

$$\rho(A) \leq \|A^k\|^{1/k} \leq \|A\|, \quad k = 1: \infty,$$

$$\lim_{k \to \infty} \|A^k\|^{1/k} = \rho(A).$$
\[ A = \begin{bmatrix} 0.9 & 500 \\ 0 & -0.5 \end{bmatrix}. \]
**Theorem**

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

$$\left\| \sum_{k=\ell}^{\infty} c_k A^k \right\| \leq \sum_{k=\ell}^{\infty} |c_k| \left( \|A^t\|^{1/t} \right)^k$$

where $\|A^t\|^{1/t} = \max\{\|A^k\|^{1/k} : k \geq \ell, \ c_k \neq 0\}$.

**Proof.** Use $\|A^k\| = (\|A^k\|^{1/k})^k \leq (\|A^t\|^{1/t})^k$. □
Lemma

If \( k = pm_1 + qm_2 \) with \( p, q \in \mathbb{N} \) and \( m_1, m_2 \in \mathbb{N} \cup \{0\} \),

\[
\| A^k \|^{1/k} \leq \max(\| A^p \|^{1/p}, \| A^q \|^{1/q}).
\]

Proof. Let \( \delta = \max(\| A^p \|^{1/p}, \| A^q \|^{1/q}) \). Then

\[
\| A^k \| \leq \| A^{pm_1} \| \| A^{qm_2} \|
\leq (\| A^p \|^{1/p})^{pm_1} (\| A^q \|^{1/q})^{qm_2}
\leq \delta^{pm_1} \delta^{qm_2} = \delta^k.
\]

Take \( \{p, q\} = \{r, r + 1\} \) for \( k \geq r(r - 1) \).
**New Scaling and Squaring Algorithm**

- Truncation bounds use $\|A^k\|^{1/k}$ instead of $\|A\|$.
- Roundoff considerations: correction to chosen $m$.
- Use *estimates* of $\|A^k\|$ where necessary (alg of H & Tisseur (2000)).
- Special treatment of triangular matrices to ensure accurate diagonal.
- New alg no slower than $\text{expm}$, potentially faster, potentially more accurate.
Summary of New Alg

- Major benefits in speed and accuracy through using $\|A^k\|^{1/k}$ in place of $\|A\|^k$.
- Overscaling problem “solved”.
- Stability of squaring phase remains an open question.
Frechét Derivative

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

\[ L(A, E) = \int_0^1 e^{A(1-s)} E e^{As} \, ds. \]

- Method based on

\[
\begin{bmatrix}
X & E \\
0 & X
\end{bmatrix}
\]

\[
\begin{bmatrix}
f(X) & L(X, E) \\
0 & f(X)
\end{bmatrix}
\]

- 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

- Many applications of $f(A)$, e.g. control theory, Markov chains, theoretical physics.
- Need better understanding of conditioning of $f(A)$.
- How to exploit structure?
- Need “factorization-free” methods for large, sparse $A$.
- Specialize to $f(A)b$ problem.


T. Crilly. 
Cayley’s anticipation of a generalised Cayley–Hamilton theorem. 

T. Crilly. 
*Arthur Cayley: Mathematician Laureate of the Victorian Age.* 
xxi+610 pp.
L. Dieci and A. Papini. Padé approximation for the exponential of a block triangular matrix.

References V

Glucat: Generic library of universal Clifford algebra templates.

N. J. Higham.
The Matrix Function Toolbox.
http://www.ma.man.ac.uk/~higham/mftoolbox.
N. J. Higham.  
*Functions of Matrices: Theory and Computation.*  
xx+425 pp.

R. A. Horn and C. R. Johnson.  
*Topics in Matrix Analysis.*  
ISBN 0-521-30587-X.  
viii+607 pp.
C. S. Kenney and A. J. Laub.
A Schur–Fréchet algorithm for computing the logarithm and exponential of a matrix.

M. J. Parks.
A Study of Algorithms to Compute the Matrix Exponential.
References VIII

*James Joseph Sylvester. Jewish Mathematician in a Victorian World.*  
xiii+461 pp.

B. Sengupta.  
Markov processes whose steady state distribution is matrix-exponential with an application to the $GI/PH/1$ queue.  
G. Walz.
Computing the matrix exponential and other matrix functions.