1. Properties of the Matrix Exponential

Let A be a real or complex $n \times n$ matrix. The exponential of A is defined via its Taylor series,

$$e^A = I + \sum_{n=1}^{\infty} \frac{A^n}{n!}\,, \tag{1}$$

where I is the $n \times n$ identity matrix. The radius of convergence of the above series is infinite. Consequently, eq. (1) converges for all matrices A. In these notes, we discuss a number of key results involving the matrix exponential and provide proofs of three important theorems. First, we consider some elementary properties.

Property 1: If $\bigl[A\,,\,B\bigr] \equiv AB - BA = 0$, then

$$e^{A+B} = e^A e^B = e^B e^A\,. \tag{2}$$

This result can be proved directly from the definition of the matrix exponential given by eq. (1). The details are left to the ambitious reader.

Remarkably, the converse of property 1 is FALSE. One counterexample is sufficient. Consider the 2×2 complex matrices

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 2\pi i \end{pmatrix}\,, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 2\pi i \end{pmatrix}\,. \tag{3}$$

An elementary calculation yields

$$e^A = e^B = e^{A+B} = I\,, \tag{4}$$

where I is the 2×2 identity matrix. Hence, eq. (2) is satisfied. Nevertheless, it is a simple matter to check that $AB \neq BA$, i.e., $[A\,,\,B] \neq 0$.

Indeed, one can use the above counterexample to construct a second counterexample that employs only real matrices. Here, we make use of the well known isomorphism between the complex numbers and real 2×2 matrices, which is given by the mapping

$$z = a + ib \quad \longmapsto \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\,. \tag{5}$$

It is straightforward to check that this isomorphism respects the multiplication law of two complex numbers. Using eq. (5), we can replace each complex number in eq. (3) with the corresponding real 2×2 matrix,

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\pi \\ 0 & 0 & -2\pi & 0 \end{pmatrix}\,, \qquad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2\pi \\ 0 & 1 & -2\pi & 0 \end{pmatrix}\,.$$

One can again check that eq. (4) is satisfied, where I is now the 4×4 identity matrix, whereas $AB \neq BA$ as before.

It turns out that a small modification of Property 1 is sufficient to avoid any such counterexamples.

Property 2: If $e^{t(A+B)} = e^{tA}e^{tB} = e^{tB}e^{tA}$, where $t \in (a,b)$ (where $a < b$) lies within some open interval of the real line, then it follows that $[A\,,\,B] = 0$.

Property 3: If S is a non-singular matrix, then for any matrix A,

$$\exp\bigl\{SAS^{-1}\bigr\} = Se^{A}S^{-1}\,. \tag{6}$$

The above result can be derived simply by making use of the Taylor series definition [cf. eq. (1)] for the matrix exponential.

Property 4: If $[A(t)\,,\,dA/dt] = 0$, then

$$\frac{d}{dt}e^{A(t)} = e^{A(t)}\frac{dA(t)}{dt} = \frac{dA(t)}{dt}e^{A(t)}\,.$$

This result should be self evident since it replicates the well known result for ordinary (commuting) functions. Note that Theorem 2 below generalizes this result in the case of $[A(t)\,,\,dA/dt] \neq 0$

Property 5: If $\bigl[A\,,\,[A\,,\,B]\bigr] = 0$, then $e^{A}\,B\,e^{-A} = B + [A\,,\,B]$.

To prove this result, we define

$$B(t) \equiv e^{tA}Be^{-tA}\,,$$

and compute

$$\frac{dB(t)}{dt} = Ae^{tA}Be^{-tA} - e^{tA}Be^{-tA}A = [A\,,\,B(t)]\,,$$

$$\frac{d^2B(t)}{dt^2} = A^2e^{tA}Be^{-tA} - 2Ae^{tA}Be^{-tA}A + e^{tA}Be^{-tA}A^2 = \bigl[A\,,\,[A\,,\,B(t)]\bigr]\,.$$

By assumption, $\bigl[A\,,\,[A\,,\,B]\bigr] = 0$, which must also be true if one replaces $A \to tA$ for any number t. Hence, it follows that $\bigl[A\,,\,[A\,,\,B(t)]\bigr] = 0$, and we can conclude that $d^2B(t)/dt^2 = 0$. It then follows that $B(t)$ is a linear function of t, which can be written as

$$B(t) = B(0) + t\left(\frac{dB(t)}{dt}\right)_{t=0}\,.$$

Noting that $B(0) = B$ and $(dB(t)/dt)_{t=0} = [A\,,\,B]$, we end up with

$$e^{tA}\,B\,e^{-tA} = B + t[A\,,\,B]\,. \tag{7}$$

By setting $t = 1$, we arrive at the desired result. If the double commutator does not vanish, then one obtains a more general result, which is presented in Theorem 1 below.

If $\big[A\,,\,B\big] \neq 0$, the $e^A e^B \neq e^{A+B}$. The general result is called the Baker-Campbell-Hausdorff formula, which will be proved in Theorem 4 below. Here, we shall prove a somewhat simpler version,

Property 6: If $\big[A\,,\,[A\,,\,B]\big] = \big[B\,,\,[A\,,\,B]\big] = 0$, then

$$e^A e^B = \exp\big\{A + B + \tfrac{1}{2}[A\,,\,B]\big\}\,. \tag{8}$$

To prove eq. (8), we define a function,

$$F(t) = e^{tA} e^{tB}\,.$$

We shall now derive a differential equation for $F(t)$. Taking the derivative of $F(t)$ with respect to t yields

$$\frac{dF}{dt} = Ae^{tA}e^{tB} + e^{tA}e^{tB}\ B = AF(t) + e^{tA}Be^{-tA}F(t) = \big\{A + B + t[A\,,\,B]\big\}F(t)\,, \tag{9}$$

after noting that B commutes with e^{Bt} and employing eq. (7). By assumption, both A and B, and hence their sum, commutes with $[A\,,\,B]$. Thus, in light of Property 4 above, it follows that the solution to eq. (9) is

$$F(t) = \exp\big\{t(A + B) + \tfrac{1}{2}t^2[A\,,\,B]\big\}F(0)\,.$$

Setting $t = 0$, we identify $F(0) = I$, where I is the identity matrix. Finally, setting $t = 1$ yields eq. (8).

Property 7: For any matrix A,

$$\det \exp A = \exp\big\{\mathrm{Tr}A\big\}\,. \tag{10}$$

If A is diagonalizable, then one can use Property 3, where S is chosen to diagonallize A. In this case, $D = SAS^{-1} = \mathrm{diag}(\lambda_1\,,\,\lambda_2\,,\,\ldots\,,\,\lambda_n)$, where the λ_i are the eigenvalues of A (allowing for degeneracies among the eigenvalues if present). It then follows that

$$\det e^A = \prod_i e^{\lambda_i} = e^{\lambda_1+\lambda_2+\ldots+\lambda_n} = \exp\big\{\mathrm{Tr}A\big\}\,.$$

However, not all matrices are diagonalizable. One can modify the above derivation by employing the Jordan canonical form. But, here I prefer another technique that is applicable to all matrices whether or not they are diagonalizable. The idea is to define a function

$$f(t) = \det e^{At}\,,$$

and then derive a differential equation for $f(t)$. If $|\delta t/t| \ll 1$, then

$$\det e^{A(t+\delta t)} = \det(e^{At}e^{A\delta t}) = \det e^{At}\det e^{A\delta t} = \det e^{At}\det(I + A\delta t)\,, \tag{11}$$

after expanding out $e^{A\delta t}$ to linear order in δt.

We now consider

$$\det(I + A\delta t) = \det \begin{pmatrix} 1 + A_{11}\delta t & A_{12}\delta t & \dots & A_{1n}\delta t \\ A_{21}\delta t & 1 + A_{22}\delta t & \dots & A_{2n}\delta t \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}\delta t & A_{n2}\delta t & \dots & 1 + A_{nn}\delta \end{pmatrix}$$

$$= (1 + A_{11}\delta t)(1 + A_{22}\delta t) \cdots (1 + A_{nn}\delta t) + \mathcal{O}\big((\delta t)^2\big)$$

$$= 1 + \delta t(A_{11} + A_{22} + \cdots + A_{nn}) + \mathcal{O}\big((\delta t)^2\big) = 1 + \delta t \,\mathrm{Tr} A + \mathcal{O}\big((\delta t)^2\big)\,.$$

Inserting this result back into eq. (11) yields

$$\frac{\det e^{A(t+\delta t)} - \det e^{At}}{\delta t} = \mathrm{Tr} A \, \det e^{At} + \mathcal{O}(\delta t)\,.$$

Taking the limit as $\delta t \to 0$ yields the differential equation,

$$\frac{d}{dt} \det e^{At} = \mathrm{Tr} A \, \det e^{At}\,. \tag{12}$$

The solution to this equation is

$$\ln \det e^{At} = t \,\mathrm{Tr} A\,, \tag{13}$$

where the constant of integration has been determined by noting that $(\det e^{At})_{t=0} = \det I = 1$. Exponentiating eq. (13), we end up with

$$\det e^{At} = \exp\big\{t \,\mathrm{Tr} A\big\}\,.$$

Finally, setting $t = 1$ yields eq. (10).

Note that this last derivation holds for any matrix A (including matrices that are singular and/or are not diagonalizable).

Remark: For any invertible matrix function $A(t)$, Jacobi's formula is

$$\frac{d}{dt} \det A(t) = \det A(t) \,\mathrm{Tr} \left(A^{-1}(t) \frac{dA(t)}{dt} \right)\,. \tag{14}$$

Note that for $A(t) = e^{At}$, eq. (14) reduces to eq. (12) derived above.

2. Four Important Theorems Involving the Matrix Exponential

The adjoint operator ad_A, which is a linear operator acting on the vector space of $n \times n$ matrices, is defined by

$$\mathrm{ad}_A(B) = [A, B] \equiv AB - BA\,. \tag{15}$$

Note that

$$(\mathrm{ad}_A)^n(B) = \underbrace{\big[A, \cdots [A, [A, B]] \cdots\big]}_{n} \tag{16}$$

involves n nested commutators.

Theorem 1:

$$e^A B e^{-A} = \exp(\mathrm{ad}_A)(B) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} (\mathrm{ad}_A)^n (B) = B + [A, B] + \tfrac{1}{2} [A, [A, B]] + \cdots \,. \qquad (17)$$

Proof: Define

$$B(t) \equiv e^{tA} B e^{-tA} \,, \qquad (18)$$

and compute the Taylor series of $B(t)$ around the point $t = 0$. A simple computation yields $B(0) = B$ and

$$\frac{dB(t)}{dt} = A e^{tA} B e^{-tA} - e^{tA} B e^{-tA} A = [A, B(t)] = \mathrm{ad}_A (B(t)) \,. \qquad (19)$$

Higher derivatives can also be computed. It is a simple exercise to show that:

$$\frac{d^n B(t)}{dt^n} = (\mathrm{ad}_A)^n (B(t)) \,. \qquad (20)$$

Theorem 1 then follows by substituting $t = 1$ in the resulting Taylor series expansion of $B(t)$.

We now introduce two auxiliary functions that are defined by their power series:

$$f(z) = \frac{e^z - 1}{z} = \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!} \,, \qquad |z| < \infty \,, \qquad (21)$$

$$g(z) = \frac{\ln z}{z - 1} = \sum_{n=0}^{\infty} \frac{(1-z)^n}{n+1} \,, \qquad |1 - z| < 1 \,. \qquad (22)$$

These functions satisfy:

$$f(\ln z)\, g(z) = 1 \,, \qquad \text{for } |1 - z| < 1 \,, \qquad (23)$$

$$f(z)\, g(e^z) = 1 \,, \qquad \text{for} |z| < \infty \,. \qquad (24)$$

Theorem 2:

$$e^{A(t)} \frac{d}{dt} e^{-A(t)} = -f(\mathrm{ad}_A) \left(\frac{dA}{dt} \right) \,, \qquad (25)$$

where $f(z)$ is defined via its Taylor series in eq. (21). Note that in general, $A(t)$ does not commute with dA/dt. A simple example, $A(t) = A + tB$ where A and B are independent of t and $[A, B] \neq 0$, illustrates this point. In the special case where $[A(t), dA/dt] = 0$, eq. (25) reduces to

$$e^{A(t)} \frac{d}{dt} e^{-A(t)} = -\frac{dA}{dt} \,, \qquad \text{if} \quad \left[A(t), \frac{dA}{dt} \right] = 0 \,. \qquad (26)$$

Proof: Define

$$B(s, t) \equiv e^{sA(t)} \frac{d}{dt} e^{-sA(t)} \,, \qquad (27)$$

and compute the Taylor series of $B(s,t)$ around the point $s = 0$. It is straightforward to verify that $B(0,t) = 0$ and

$$\left.\frac{d^n B(s,t)}{ds^n}\right|_{s=0} = -(\mathrm{ad}_{A(t)})^{n-1}\left(\frac{dA}{dt}\right), \tag{28}$$

for all positive integers n. Assembling the Taylor series for $B(s,t)$ and inserting $s = 1$ then yields Theorem 2. Note that if $[A(t), dA/dt] = 0$, then $(d^n B(s,t)/ds^n)_{s=0} = 0$ for all $n \geq 2$, and we recover the result of eq. (26).

Theorem 3:

$$\frac{d}{dt}e^{-A(t)} = -\int_0^1 e^{-sA}\,\frac{dA}{dt}\,e^{-(1-s)A}\,ds\,. \tag{29}$$

This integral representation is an alternative version of Theorem 2.

Proof: Consider

$$\begin{aligned}\frac{d}{ds}\left(e^{-sA}\,e^{-(1-s)B}\right) &= -Ae^{-sA}\,e^{-(1-s)B} + e^{-sA}\,e^{-(1-s)B}B \\ &= e^{-sA}(B-A)e^{-(1-s)B}\,.\end{aligned} \tag{30}$$

Integrate eq. (30) from $s = 0$ to $s = 1$.

$$\int_0^1 \frac{d}{ds}\left(e^{-sA}\,e^{-(1-s)B}\right) = e^{-sA}\,e^{-(1-s)B}\Big|_0^1 = e^{-A} - e^{-B}\,. \tag{31}$$

Using eq. (30), it follows that:

$$e^{-A} - e^{-B} = \int_0^1 ds\,e^{-sA}(B-A)e^{-(1-s)B}\,. \tag{32}$$

In eq. (32), we can replace $B \longrightarrow A + hB$, where h is an infinitesimal quantity:

$$e^{-A} - e^{-(A+hB)} = h\int_0^1 ds\,e^{-sA}Be^{-(1-s)(A+hB)}\,. \tag{33}$$

Taking the limit as $h \to 0$,

$$\lim_{h\to 0}\frac{1}{h}\left[e^{-(A+hB)} - e^{-A}\right] = -\int_0^1 ds\,e^{-sA}Be^{-(1-s)A}\,. \tag{34}$$

Finally, we note that the definition of the derivative can be used to write:

$$\frac{d}{dt}e^{-A(t)} = \lim_{h\to 0}\frac{e^{-A(t+h)} - e^{-A(t)}}{h}\,. \tag{35}$$

Using

$$A(t+h) = A(t) + h\frac{dA}{dt} + \mathcal{O}(h^2)\,, \tag{36}$$

it follows that:

$$\frac{d}{dt}e^{-A(t)} = \lim_{h\to 0} \frac{\exp\left[-\left(A(t) + h\dfrac{dA}{dt}\right)\right] - \exp[-A(t)]}{h}\,. \tag{37}$$

Thus, we can use the result of eq. (34) with $B = dA/dt$ to obtain

$$\frac{d}{dt}e^{-A(t)} = -\int_0^1 e^{-sA}\,\frac{dA}{dt}\,e^{-(1-s)A}\,ds\,, \tag{38}$$

which is the result quoted in Theorem 3.

Second proof of Theorem 2: One can now derive Theorem 2 directly from Theorem 3. Multiply eq. (29) by $e^{A(t)}$ to obtain:

$$e^{A(t)}\frac{d}{dt}e^{-A(t)} = -\int_0^1 e^{(1-s)A}\,\frac{dA}{dt}\,e^{-(1-s)A}\,ds\,. \tag{39}$$

Using Theorem 1 [see eq. (17)],

$$\begin{aligned} e^{A(t)}\frac{d}{dt}e^{-A(t)} &= -\int_0^1 \exp\left[\mathrm{ad}_{(1-s)A}\right]\left(\frac{dA}{dt}\right)\,ds \\ &= -\int_0^1 e^{(1-s)\mathrm{ad}_A}\left(\frac{dA}{dt}\right)\,ds\,. \end{aligned} \tag{40}$$

Changing variables $s \longrightarrow 1-s$, it follows that:

$$e^{A(t)}\frac{d}{dt}e^{-A(t)} = -\int_0^1 e^{s\,\mathrm{ad}_A}\left(\frac{dA}{dt}\right)\,ds\,. \tag{41}$$

The integral over s is trivial, and one finds:

$$e^{A(t)}\frac{d}{dt}e^{-A(t)} = \frac{1-e^{\,\mathrm{ad}_A}}{\mathrm{ad}_A}\left(\frac{dA}{dt}\right) = -f(\mathrm{ad}_A)\left(\frac{dA}{dt}\right)\,, \tag{42}$$

which coincides with Theorem 2.

Theorem 4: The Baker-Campbell-Hausdorff (BCH) formula

$$\ln\left(e^A e^B\right) = B + \int_0^1 g\left[\exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B)\right](A)\,dt\,, \tag{43}$$

where $g(z)$ is defined via its Taylor series in eq. (22). Since $g(z)$ is only defined for $|1-z| < 1$, it follows that the BCH formula for $\ln\left(e^A e^B\right)$ converges provided that $\|e^A e^B - I\| < 1$, where I is the identity matrix and $\|\cdots\|$ is a suitably defined matrix norm. Expanding the BCH formula, using the Taylor series definition of $g(z)$, yields:

$$e^A e^B = \exp\left(A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{12}\,[A,[A,B]] + \tfrac{1}{12}\,[B,[B,A]] + \ldots\right)\,, \tag{44}$$

assuming that the resulting series is convergent. An example where the BCH series does not converge occurs for the following elements of SL(2,$\mathbb{R}$):

$$M = \begin{pmatrix} -e^{-\lambda} & 0 \\ 0 & -e^{\lambda} \end{pmatrix} = \exp\left[\lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] \exp\left[\pi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right], \tag{45}$$

where λ is any nonzero real number. It is easy to prove[1] that no matrix C exists such that $M = \exp C$. Nevertheless, the BCH formula is guaranteed to converge in a neighborhood of the identity of any Lie group.

Proof of the BCH formula: Define

$$C(t) = \ln(e^{tA}e^{B}) . \tag{46}$$

or equivalently,

$$e^{C(t)} = e^{tA}e^{B} . \tag{47}$$

Using Theorem 1, it follows that for any complex $n \times n$ matrix H,

$$\begin{aligned} \exp\left[\mathrm{ad}_{C(t)}\right](H) &= e^{C(t)} H e^{-C(t)} = e^{tA}e^{B} H e^{-tA}e^{-B} \\ &= e^{tA}\left[\exp(\mathrm{ad}_B)(H)\right] e^{-tA} \\ &= \exp(\mathrm{ad}_{tA}) \exp(\mathrm{ad}_B)(H) . \end{aligned} \tag{48}$$

Hence, the following operator equation is valid:

$$\exp\left[\mathrm{ad}_{C(t)}\right] = \exp(t\, \mathrm{ad}_A) \exp(\mathrm{ad}_B) , \tag{49}$$

after noting that $\exp(\mathrm{ad}_{tA}) = \exp(t\, \mathrm{ad}_A)$. Next, we use Theorem 2 to write:

$$e^{C(t)} \frac{d}{dt} e^{-C(t)} = -f(\mathrm{ad}_{C(t)}) \left(\frac{dC}{dt}\right) . \tag{50}$$

However, we can compute the left-hand side of eq. (50) directly:

$$e^{C(t)} \frac{d}{dt} e^{-C(t)} = e^{tA}e^{B} \frac{d}{dt} e^{-B}e^{-tA} = e^{tA} \frac{d}{dt} e^{-tA} = -A , \tag{51}$$

[1]The characteristic equation for any 2×2 matrix A is given by

$$\lambda^2 - (\mathrm{Tr}\ A)\lambda + \det\ A = 0 .$$

Hence, the eigenvalues of any 2×2 traceless matrix $A \in \mathfrak{sl}(2,\mathbb{R})$ [that is, A is an element of the Lie algebra of SL(2,$\mathbb{R}$)] are given by $\lambda_\pm = \pm(-\det\ \mathrm{A})^{1/2}$. Then,

$$\mathrm{Tr}\ e^A = \exp(\lambda_+) + \exp(\lambda_-) = \begin{cases} 2\cosh|\det\ A|^{1/2} , & \text{if } \det\ A \leq 0 , \\ 2\cos|\det\ A|^{1/2} , & \text{if } \det\ A > 0 . \end{cases}$$

Thus, if $\det\ A \leq 0$, then $\mathrm{Tr}\ e^A \geq 2$, and if $\det\ A > 0$, then $-2 \leq \mathrm{Tr}\ e^A < 2$. It follows that for any $A \in \mathfrak{sl}(2,\mathbb{R})$, $\mathrm{Tr}\ e^A \geq -2$. For the matrix M defined in eq. (45), $\mathrm{Tr}\ M = -2\cosh\lambda < -2$ for any nonzero real λ. Hence, no matrix C exists such that $M = \exp C$.

since B is independent of t, and tA commutes with $\frac{d}{dt}(tA)$. Combining the results of eqs. (50) and (51),

$$A = f(\mathrm{ad}_{C(t)})\left(\frac{dC}{dt}\right) . \tag{52}$$

Multiplying both sides of eq. (52) by $g(\exp \mathrm{ad}_{C(t)})$ and using eq. (24) yields:

$$\frac{dC}{dt} = g(\exp \mathrm{ad}_{C(t)})(A) . \tag{53}$$

Employing the operator equation, eq. (49), one may rewrite eq. (53) as:

$$\frac{dC}{dt} = g(\exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B))(A) , \tag{54}$$

which is a differential equation for $C(t)$. Integrating from $t = 0$ to $t = T$, one easily solves for C. The end result is

$$C(T) = B + \int_0^T g(\exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B))(A)\,dt , \tag{55}$$

where the constant of integration, B, has been obtained by setting $T = 0$. Finally, setting $T = 1$ in eq. (55) yields the BCH formula.

Finally, we shall use eq. (43) to obtain the terms exhibited in eq. (44). In light of the series definition of $g(z)$ given in eq. (22), we need to compute

$$\begin{aligned} I - \exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B) &= I - (I + t\,\mathrm{ad}_A + \tfrac{1}{2}t^2\,\mathrm{ad}_A^2)(I + \mathrm{ad}_B + \tfrac{1}{2}\mathrm{ad}_B^2) \\ &= -\mathrm{ad}_B - t\,\mathrm{ad}_A - t\,\mathrm{ad}_A\,\mathrm{ad}_B - \tfrac{1}{2}\mathrm{ad}_B^2 - \tfrac{1}{2}t^2\,\mathrm{ad}_A^2 , \end{aligned} \tag{56}$$

and

$$\big[I - \exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B)\big]^2 = \mathrm{ad}_B^2 + t\,\mathrm{ad}_A\mathrm{ad}_B + t\,\mathrm{ad}_B\mathrm{ad}_A + t^2\,\mathrm{ad}_A^2 , \tag{57}$$

after dropping cubic terms and higher. Hence, using eq. (22),

$$g(\exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B)) = I - \tfrac{1}{2}\mathrm{ad}_B - \tfrac{1}{2}t\,\mathrm{ad}_A - \tfrac{1}{6}t\,\mathrm{ad}_A\,\mathrm{ad}_B + \tfrac{1}{3}t\,\mathrm{ad}_B\,\mathrm{ad}_A + \tfrac{1}{12}\mathrm{ad}_B^2 + \tfrac{1}{12}t^2\,\mathrm{ad}_A^2 . \tag{58}$$

Noting that $\mathrm{ad}_A(A) = [A, A] = 0$, it follows that to cubic order,

$$\begin{aligned} B + \int_0^1 g(\exp(t\,\mathrm{ad}_A)\exp(\mathrm{ad}_B))(A)\,dt &= B + A - \tfrac{1}{2}[B, A] - \tfrac{1}{12}\big[A, [B, A]\big] + \tfrac{1}{12}\big[B, [B, A]\big] \\ &= A + B + \tfrac{1}{2}[A, B] + \tfrac{1}{12}\big[A, [A, B]\big] + \tfrac{1}{12}\big[B, [B, A]\big] , \end{aligned} \tag{59}$$

which confirms the result of eq. (44).

References:

The proofs of Theorems 1, 2 and 4 can be found in section 5.1 of *Symmetry Groups and Their Applications*, by Willard Miller Jr. (Academic Press, New York, 1972). The proof of Theorem 3 is based on results given in section 6.5 of *Positive Definite Matrices*, by Rajendra Bhatia (Princeton University Press, Princeton, NJ, 2007). Bhatia notes that eq. (29) has been attributed variously to Duhamel, Dyson, Feynman and Schwinger. See also R.M. Wilcox, J. Math. Phys. **8**, 962 (1967). Theorem 3 is also quoted in eq. (5.75) of *Weak Interactions and Modern Particle Theory*, by Howard Georgi (Dover Publications, Mineola, NY, 2009) [although the proof of this result is relegated to an exercise].

The proof of Theorem 2 using the results of Theorem 3 is based on my own analysis, although I would not be surprised to find this proof elsewhere in the literature. Finally, a nice discussion of the SL(2,$\mathbb{R}$) matrix that cannot be written as a single exponential can be found in section 3.4 of *Matrix Groups: An Introduction to Lie Group Theory*, by Andrew Baker (Springer-Verlag, London, UK, 2002), and in section 10.5(b) of *Group Theory in Physics*, Volume 2, by J.F. Cornwell (Academic Press, London, UK, 1984).