Mathematics Notes
Note 92

Summer 1998

Lie-Algebraic Representations of Product Integrals of Variable Matrices

Stanly Steinberg
Professor of Mathematics
Department of Mathematics and Statistics
University of New Mexico
Albuquerque NM 87131-1141 USA
stanly@math.unm.edu
http://math.unm.edu/~stanly/

Carl E. Baum
Air Force Reserach Laboratory DEHP
3550 Aberdeen Avenue SE Kirtland AFB, NM 87117-5776

Abstract

Lie algebraic ideas are used to find useful representations of product integrals of variable matrix functions. These representations are then used to construct explicit solutions to variable coefficient differential equations that have applications to transmission-line and wave-launcher problems.

Contents

List of Tables

1 Introduction

We will use some ideas from the theory of Lie algebras and groups to find representations of solutions of systems of first order ordinary differential equations. The system will be written in the form

$$\frac{d\vec{y}}{dt} = A(t)\,\vec{y} + \vec{f}(t)\,, \quad \vec{y}(0) = \vec{y}_0\,, \tag{1.1}$$

where $t \in [0, T]$ with $T > 0$ and $A(t)$ is a given sufficiently-smooth n by n matrix-valued function of $t \in [0, T]$, $\vec{f}(t)$ is a given sufficiently-smooth n-vector-valued function of $t \in [0, T]$, $\vec{y}_0$ is some given n-vector, and $n \geq 1$. It is well-known that such problems have a smooth solution $\vec{y}(t)$ for $t \in [0, T]$.

The most important representation theorem for such problems says that the solution $\vec{y}(t)$ of (1.1) can be written in the form

$$\vec{y}(t) = G(t, 0)\,\vec{y}_0 + \int_0^t G(t, \tau)\,d\tau\,, \quad 0 \leq t \leq T\,, \tag{1.2}$$

where $G(t, \tau)$ is a smooth n by n matrix-valued solution of

$$\frac{d}{dt}G(t, \tau) = A(t)\,G(t, \tau)\,, \quad G(\tau, \tau) = I\,,\,, \quad 0 \leq \tau \leq t \leq T\,, \tag{1.3}$$

where I is the n by n identity matrix. The function G is called the *Green's function, propagator*, or *matrizant*. These results can be found in many textbooks on ordinary-differential equations. The book by [6] by Dollard and Friedman reviews this material and additionally is the primary reference for some material we will use later.

In the constant coefficient case, that is, when $A(t) = A$ is independent of t, then

$$G(t, \tau) = e^{A\,(t-\tau)}\,, \tag{1.4}$$

where for any matrix M

$$e^M = \sum_0^\infty \frac{M^n}{n!}\,. \tag{1.5}$$

This series converges absolutely and uniformly in the entries of M, and can thus be differentiated term by term. So setting $M = A\,t$ gives

$$e^{A\,0} = I\,, \quad \frac{d}{dt}e^{A\,t} = A\,e^{A\,t}\,, \tag{1.6}$$

which implies that (1.4) is in fact the Green's function. If the matrix function $A(t)$ is not constant then Green's function G can still be represented in an exponential form using the product integral as described in Dollard and Friedman:

$$G(t, \tau) = \prod_\tau^t e^{ds\,A(s)}\,, \tag{1.7}$$

where the product integral can be defined as

$$\prod_{\tau}^{t} e^{ds\,A(s)} = \lim_{h\to 0} \prod_{i=0}^{N-1} e^{h\,A(\tau+i\,h)}\,, \tag{1.8}$$

where $h = (t-\tau)/N$, $A(t)$ is a continuous function of t, and the product in 1.8 means *left* dot multiplication by successive term (increasing i).

The main point we want to take from this introductory section is that the solution of systems of ordinary differential equations can be written in terms of exponentials of matrices, even in the variable coefficient case, where product integrals must be used. The theory and formulas presented here are quite useful, but our goals is to produce much more useful formulas using the notions of Lie algebras and groups.

This project resulted from the authors realizing that there was a close connection between some of the second authors work on transmission lines and wave launchers [2, 3, 4] and the first authors research on the applications of Lie methods to differential equations [9, 10, 11, 12, 13].

2 Lie Algebras and Lie Groups

The text [14] has a useful introduction to Lie algebra and groups, and Chapters 2 and 8 of the book [1] edited by Baum and Kritikos have nice overviews of Lie algebras and groups and their connection to Electromagnetics. However, we will include here a complete discussion of all of the facts that we need. We will restrict our attention to Lie algebras and groups of n by n matrices with entries from either the complex field or the real field. A *Lie algebra* of matrices is a non-empty set of matrices that is closed under addition, scalar multiplication and commutation. The commutator of two matrices is defined by

$$[A,B] = A\,B - B\,A\,. \tag{2.9}$$

A *Lie group* of matrices is is a non-empty set of matrices that is closed in multiplication and matrix inversion along with an additional property that will be described later.

The set of all n by n matrices is a Lie algebra called $gl(n)$ which stands for the *general linear* algebra. The general linear algebra $gl(n)$ is a linear space with addition being the standard addition of entries in the matrix and scalar multiplication being multiplication of all entries by an element of the field. The matrices with non-zero determinant form a Lie group called the *general linear* group. These examples will be our main algebras and groups with all other examples being subsets of these.

In applications to differential equations, the matrices $A(t)$ in (1.1) will be taken from a Lie algebra, while the Green's function, propagator, or matrizant used in (1.3) will be in a Lie group.

If $A(t) = A$ is constant, then the solutions of the system of ordinary differential equations (1.1) can be written in terms of the Green's function $G(t,\tau) = G(t-\tau)$ where

$$G(t) = e^{A\,t}\,. \tag{2.10}$$

The Lie group associated with a given Lie algebra is supposed to represent the set of all possible Green's functions generated by elements of the Lie algebra. More precisely, the Lie group associated with a Lie algebra is the set of all possible finite products of exponentials of elements from the Lie algebra.

To understand Lie algebras and groups better, we need a few facts about exponential of matrices. Recall that the exponential is defined by

$$e^{A\,t} = \sum_{n=0}^{\infty} \frac{A^n\,t^n}{n!}\,, \tag{2.11}$$

which is a uniformly and absolutely convergent series for all t, and consequently the entries in the $\exp(A\,t)$ are all infinitely differentiable. It is an easy to check that

$$\begin{aligned} e^{A\,0} &= I\,, \\ e^{A\,t}\,e^{-A\,t} &= I \\ \left(e^{A}\right)^{\dagger} &= e^{A^{\dagger}} \end{aligned} \tag{2.12}$$

$$
\begin{aligned}
e^{A(t_1+t_2)} &= e^{A t_1} e^{A t_2} = e^{A t_2} e^{A t_1}, \\
\frac{d}{dt} e^{A t} &= A e^{A t} = e^{A t} A,
\end{aligned}
$$

where $A^\dagger$ is the complex conjugate transpose of the matrix A. In particular, the second identity implies that the matrix $\exp(A t)$ is invertible with inverse $\exp(-A t)$.

An important tool for evaluating exponentials is the formula for how the exponential changes under a similarity transformation:

$$
e^{S^{-1} A S t} = S^{-1} e^{A t} S, \quad \det(S) \neq 0, \tag{2.13}
$$

which is easy to check from the series definition of the exponential.

It is easy to evaluate exponentials if the matrix can be put into a simple block form. For example, if A is block matrix of the form

$$
A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \tag{2.14}
$$

then

$$
A^n = \begin{pmatrix} A_1^n & 0 \\ 0 & A_2^n \end{pmatrix} \tag{2.15}
$$

and then the series definition of the exponential gives

$$
e^{A t} = \begin{pmatrix} e^{A_1 t} & 0 \\ 0 & e^{A_2 t} \end{pmatrix} . \tag{2.16}
$$

In particular, if A is a diagonal matrix, $A = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then

$$
e^{A t} = \operatorname{diag}\left(e^{\lambda_1 t}, e^{\lambda_2 t}, \cdots, e^{\lambda_n t}\right) . \tag{2.17}
$$

For any matrix there is a similarity transformation that reduces the matrix to Jordan form, that is, the matrix is block diagonal with Jordan blocks on the diagonal. So to compute the exponential of any matrix, we only need to know how to compute the exponentials of the Jordan blocks. Unfortunately, in general, it is difficult to find the Jordan form if a matrix contains symbolic parameters or floating-point numbers. Still, this is a useful tool for small matrices. The simplest Jordan block is 2 by 2:

$$
J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \tag{2.18}
$$

and, using the power series definition, its exponential is easily computed to be

$$
e^{J t} = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} . \tag{2.19}
$$

Another way to find this result is to convert the computation of the exponential to a system of differential equations. So assume that

$$e^{Jt} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.20}$$

where $a = a(t)$, $b = b(t)$, $c = c(t)$, $d = d(t)$. Differentiating gives

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tag{2.21}$$

So a, b, c, and d satisfy the differential equations

$$a' = \lambda\, a + c\,, \quad b' = \lambda\, b + d\,, \quad c' = \lambda\, c\,, \quad d' = \lambda\, d\,, \tag{2.22}$$

and the fact that the exponential is the identity at $t = 0$ givens the initial conditions

$$a(0) = 1\,, \quad b(0) = 0\,, \quad c(0) = 0\,, \quad d(0) = 1\,. \tag{2.23}$$

The solutions are easily seen to be

$$a(t) = e^{\lambda t}\,, \quad b(t) = t\, e^{\lambda t}\,, \quad c(t) = 0\,, \quad d(t) = e^{\lambda t}\,, \tag{2.24}$$

which agrees with the power-series computation.

Let us now return to the problem of identifying the Lie group generated by exponentials of elements from the general linear algebra $gl(n)$. It is well known [7] that the trace $\mathrm{tr}(A)$ of a matrix A and the determinant $\det(A)$ are related by:

Lemma:

$$\det(e^A) = e^{\mathrm{tr}(A)}\,. \tag{2.25}$$

Proof: Even though well-known, this is so important for us that we give a proof. It is also well known [7] that the determinant and trace of a matrix are invariant under similarity transforms, so if $\det(S) \neq 0$, then

$$\det(S^{-1}\, A\, S) = \det(A)\,, \quad \mathrm{tr}(S^{-1}\, A\, S) = \mathrm{tr}(A)\,. \tag{2.26}$$

For matrices in Jordan form (diagonal form, or even triangular form) it is simple to see that

$$\det(e^J) = e^{\mathrm{tr}(J)}\,. \tag{2.27}$$

Any matrix can be put into Jordan form, so for any A there is an S so that $J = S^{-1}\, A\, S$ and then

$$\det\left(e^A\right) = \det\left(S^{-1}\, e^A\, S\right) = \det\left(e^{S^{-1} A\, S}\right) = \det\left(e^J\right) = e^{\mathrm{tr}(J)} = e^{\mathrm{tr}(S\, J\, S^{-1})} = e^{\mathrm{tr}(A)}\,. \tag{2.28}$$

When the matrix has complex entries, the trace is an exponential of a complex number which cannot be zero, so the matrix is invertible. In fact, we already know $(\exp(A\,t) = \exp(-A\,t)$ (see (2.12)). When the entries of a matrix are real, then the trace is real, and because the exponential function is always positive, the exponential of a matrix always has a positive determinant and thus is invertible. So the group $gl(n)$ is the set of matrices with positive determinant.

We digress for a moment to consider the more difficult question: can every matrix M with positive determinant be written as an exponential? Clearly we should take the logarithm of M. If $|x| < 1$, then the base e logarithm of $1 + x$ can be expressed as a power series:

$$\ln(1+x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}. \tag{2.29}$$

The size of a matrix can be specified by giving a norm on all matrices. We will use the norm that is the square root of the sum of the squares of the entries of the matrix, called the *root span* which can be conveniently written as

$$||A|| = \mathrm{rsp}(A) = \sqrt{\mathrm{tr}\,(A\,A^{\dagger})}\,, \tag{2.30}$$

where $A^{\dagger}$ is the complex conjugate transpose of the matrix A. So if $||M|| < 1$, then we define

$$\ln(I+M) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{(M)^n}{n}. \tag{2.31}$$

An elementary power series calculation shows that for x sufficiently small x

$$e^{\ln(1+x)} = 1 + x\,, \quad ln(e^x) = x\,, \tag{2.32}$$

so that for M sufficiently small

$$e^{\ln(1+M)} = 1 + M\,, \quad \ln(e^M) = M\,, \tag{2.33}$$

Because any matrix near the identity matrix can be written as $I + M$ where M is small, we conclude that all matrices with non-zero determinant that are sufficiently near the identity can be written as the exponential of another matrix, in general with complex entries. This is definitely not true for all matrices with non-zero determinant. Because the logarithm has a limited radius of convergence and is generally hard to work with, the use of the logarithm to define the Lie algebra given a Lie group is not such a good idea.

To study a the Lie algebra associated with a Lie group, it is better to use differentiation, in particular, the logarithmic derivative. Let $G(t)$ be a function from some interval to the group and then define the elements of the Lie algebra to be the *product derivative*

$$A = D_t\,G(t)|_{t=a} = \left\{\frac{d}{dt}G(t)\right\}\,G^{-1}(t)\bigg|_{t=a}. \tag{2.34}$$

for all a in the interval and all possible functions $G(t)$. Actually one can restrict the interval to contain zero, the functions to have $G(0) = I$ and then

$$A = \frac{d}{dt}G(t)\Big|_{t=0} \tag{2.35}$$

are sufficient to give all elements in the algebra. Now this is nice as we see that if A is in a Lie algebra, then $G(t) = \exp(A\,t)$ is in the Lie group and then

$$\frac{d}{dt}G(t)\Big|_{t=0} = \frac{d}{dt}e^{A\,t}\Big|_{t=0} = A\,e^{A\,t}\Big|_{t=0} = A\,, \tag{2.36}$$

or the logarithmic derivative gives

$$\frac{d}{dt}G(t)\,G^{-1}(t) = \frac{d}{dt}e^{A\,t}e^{-A\,t} = A\,e^{A\,t}e^{-A\,t} = A\,, \tag{2.37}$$

so things are nicely consistent. The additional assumption in the definition of a Lie group mentioned earlier can be stated as: if G belongs to the Lie group, then there is a function $G(t)$ with $G(t)$ in the group and $G(0) = I$ and $G(1) = G$. This actually defines a connected Lie Group, for example, for matrices with real entries, the natural Lie group consists of the matrices with positive determinant (not just non-zero determinant).

So given a lie group the Lie algebra can be found by computing derivatives at the origin. Given a Lie algebra, the Lie group can be found by computing exponentials of the elements of the Lie algebra.

Now all of this will give us some more examples of Lie algebras and groups which we have summarized in Table 2.1. The general linear algebras and groups have already been discussed. Because the trace is a linear operation it is clear that the traceless matrices are preserved under scalar multiplication and addition. If $A = (a_{i,j})$ and $B = (b_{i,j})$ then

$$\mathrm{tr}(A\,B) = \sum_{i,j=1}^{n} a_{i,j}\,b_{j,i} = \mathrm{tr}(B\,A)\,, \tag{2.38}$$

so for all matrices

$$\mathrm{tr}([A, B]) = \mathrm{tr}(A\,B - B\,A) = \mathrm{tr}(A\,B) - \mathrm{tr}(B\,A) = 0\,. \tag{2.39}$$

In particular, the set of traceless matrices is closed under both linear operations and commutation, and so they form a lie algebra. Because the exponential of a traceless matrix must have determinant one due to formula (2.25), the Lie group associated with the traceless matrices are the matrices of determinant one.

Next we will show that the group of unitary matrices, that is, matrices that satisfy

$$M^{\dagger}M = I \tag{2.40}$$

algebra		group	
$gl(n)$	all A	$GL(n)$	$\det(M) \neq 0$
$sl(n)$	$\text{tr}A = 0$	$SL(n)$	$\det(M) = 1$
$u(n)$	$A^\dagger = -A$	$U(n)$	$M^\dagger M = I$

Table 2.1: Lie Algebras and Group

have the associated Lie algebra of skew symmetric matrices, that is, matrices that satisfy

$$A^\dagger = -A\,. \tag{2.41}$$

To show that the group associated with the algebra is correct, assume that $M = \exp(A\,t)$ with $A^\dagger = -A$. Then (2.12) gives

$$M^\dagger M = e^{A^\dagger t} e^{A\,t} = e^{-A\,t} e^{A\,t} = I\,, \tag{2.42}$$

so the exponential of a skew adjoint matrix is unitary. To show that the algebra associated with the group is correct, suppose that $M(t)$ is a matrix function satisfying

$$M^\dagger(t)\,M(t) = I\,, \tag{2.43}$$

so

$$M^\dagger(t) = M^{-1}(t)\,, \tag{2.44}$$

and then also

$$M(t)\,M^\dagger(t) = I\,, \tag{2.45}$$

Now set

$$A = M'(t)\,M^{-1}(t) = M'(t)\,M^\dagger(t) \tag{2.46}$$

and then differentiate (2.43) to get

$$M^{\dagger\prime}(t)\,M(t) + M^\dagger(t)\,M'(t) = 0\,, \tag{2.47}$$

Multiplying on the left by M and on the right by $M^\dagger$ gives

$$M(t)\,M^{\dagger\prime}(t) + M'(t)\,M^\dagger(t) = 0\,, \tag{2.48}$$

or

$$A^\dagger + A = 0\,. \tag{2.49}$$

So, A is skew adjoint, and the Lie algebra associated with the unitary matrices consists of the skew symmetric matrices and vice versa. Similar arguments can be used to find the connection between various Lie groups and Lie algebras. These results are summarized in Table 2.1. For more examples, see [14] or any text on Lie algebras and groups.

3 Lie Theory and Ordinary Differential Equations

In this section, we show how to use Lie theory to solve systems of ordinary differential equations. We begin with the derivation of two important formulas and then describe the "solution" algorithm.

3.1 Important Lie Formulas

To apply Lie theory to differential equations, we need a few special formulas. The first is
Lemma:

$$\frac{d}{dt}e^{f(t)\,A} = f'(t)\,A\,e^{f(t)\,A}\,, \tag{3.50}$$

where $f(t)$ is any smooth function and A is any constant square matrix.

This is easily checked using the power series definition (2.11) of the exponential.

The second formula is for the *adjoint action* of the Lie group on the Lie algebra, which appears repeatedly when formulas are simplified or rearranged. First, for any two square matrices A and B of the same size, define

$$[A,\diamond]^0\,B = B\,, \quad [A,\diamond]^n = [A,[A,\diamond]^{n-1}]\,, \tag{3.51}$$

so that

$$\begin{aligned}
[A,\diamond]^1\,B &= [A,B]\,,\\
[A,\diamond]^2\,B &= [A,[A,B]]\,,\\
[A,\diamond]^3\,B &= [A,[A,[A,B]]]\,,
\end{aligned}$$

and so forth. Now the adjoint action of A on B is given by

$$e^{[A,\diamond]}\,B = \sum_{n=0}^{\infty}\frac{[A,\diamond]^n\,B}{n!}\,. \tag{3.52}$$

Lemma:

$$e^{A}B\,e^{-A} = e^{[A,\diamond]}\,B\,. \tag{3.53}$$

Proof: This is really a formula for rearranging power series, and can be seen to be true by comparing the power series of both sides of the equation. However, an ordinary differential equation proof is more illuminating and points out the power of this "trick". Let

$$F(t) = e^{At}B\,e^{-At} - e^{t[A,\diamond]}\,B\,. \tag{3.54}$$

Then

$$\begin{aligned}
F'(t) &= A\,e^{At}B\,e^{-At} - e^{At}B\,e^{-At}A - [A,\diamond]\,e^{t[A,\diamond]}\,B\\
&= [A, e^{At}B\,e^{-At}] - [A, e^{t[A,\diamond]}\,B]\\
&= [A, F(t)]\,.
\end{aligned} \tag{3.55}$$

So

$$F'(t) = [A, F(t)]\,, \quad F(0) = 0\,, \tag{3.56}$$

which implies that $F(t) \equiv 0$.

Note that if A and B commute $[A, B] = 0$, then

$$e^{At} B\, e^{-At} = e^{t[A,\circ]}\, B = B\,. \tag{3.57}$$

3.2 Solution Algorithm

Recall that the system of ordinary differential equations (1.1) is determined by the matrix function $A(t)$. Because the $A(t)$ are n by n matrices, they certainly belong to the Lie algebra $GL(n)$. The first step in the solution process is to find the smallest Lie algebra $\mathcal{A}$ of matrices into which all of the matrices $A(t)$ fall.

The second step is to choose a basis for the Lie algebra $\mathcal{A}$, say

$$\mathcal{A} = \mathrm{span}(A_1, A_2, \cdots, A_k)\,, \quad 1 \le k \le n^2\,, \tag{3.58}$$

where here *span* means all real or complex linear combinations of the matrix bases elements A_i, that is

$$A = \sum_{i=1}^{k} a_i(t)\, A_i\,, \tag{3.59}$$

where $a_i(t)$ are either real or complex valued functions of t.

Now the previous discussion implies that for any real numbers a_i, $1 \le i \le k$, the matrices

$$M(a_1, a_2, \cdots, a_k) = e^{a_1 A_1 + a_2 A_2 + \cdots + a_k A_k}\,. \tag{3.60}$$

are in the Lie group and all of the matrices near the identity in the Lie group can be written this way. The numbers a_i are called the coordinates on the group near the identity. This is the typical mathematical approach to generating Lie groups from Lie algebras. However there are better ways to represent a Lie group if we are interested in finding simple formulas for the representation, which is:

$$M(a_1, a_2, \cdots, a_k) = e^{a_1 A_1}\, e^{a_2 A_2} \cdots e^{a_k A_k}\,. \tag{3.61}$$

So we have exchanged an exponential of a sum for a product of exponentials. Again it is clear that G is in the Lie group, but is also true that all elements of the group near the identity can be represented in this way. There are many possible variant of this representation. Also note that Lie algebras have many bases and the basis impacts the simplicity of the resulting formulas. One good idea is to choose as many elements of the basis as possible that commute with each other, that is, choose part of the basis as a basis for the largest commuting sub-algebra.

For any computation, a table of commutators

$$[A_i, A_j]\,, \quad 1 \le i, j \le k \tag{3.62}$$

will be needed, so the third step is to compute this table.

Recall that the solution of the differential equations is given in terms of G which is the Green's function, propagator, or matrizant, which is the solution of the system of differential equations (1.3):

$$\frac{d}{dt}G(t,\tau) = A(t)\, G(t,\tau)\,, \quad G(\tau,\tau) = I\,,\,, \quad 0 \leq \tau \leq t \leq T\,. \tag{3.63}$$

So Lie theory tells us that the Green's function can be written in the form

$$G(t,\tau) = e^{g_1(t,\tau)\,A_1}\, e^{g_2(t,\tau)\,A_2} \ldots e^{g_k(t,\tau)\,A_k}\,. \tag{3.64}$$

We call this the *fundamental* representation. Note that it depends strongly on the basis and on the ordering of the basis.

First note that because we must have $G(\tau,\tau) = I$, we need

$$g_i(\tau,\tau) = 0\,, \quad 1 \leq i \leq k\,. \tag{3.65}$$

Next we want to plug the representation (3.64) into the differential equation in (1.3), so we need to differentiate the representation. Applying the product rule and the derivative of exponentials formula (3.50) gives

$$\begin{aligned} \frac{d}{dt}G(t,\tau) \;=\;& g_1'(t,\tau)\,A_1\, e^{g_1(t,\tau)\,A_1}\, e^{g_2(t,\tau)\,A_2}\, e^{g_3(t,\tau)\,A_1} \ldots \\ & + e^{g_1(t,\tau)\,A_1}\, g_2'(t,\tau)\,A_2\, e^{g_2(t,\tau)\,A_2}\, e^{g_3(t,\tau)\,A_1} \ldots \\ & + e^{g_1(t,\tau)\,A_1}\, e^{g_2(t,\tau)\,A_2}\, g_3'(t,\tau)\,A_3\, e^{g_3(t,\tau)\,A_1} \ldots \\ & + \cdots \end{aligned} \tag{3.66}$$

Now we can plug in the derivative of G into the differential equation and then multiply the resulting equation on the *right* by the inverse of G to get the logarithmic, product, or multiplicative derivatives (see Dollard and Friedman [6], Section 1.3) on the left-hand side of the equation. The inverse of G can be written in the form

$$G^{-1}(t,\tau) = e^{-g_k(t,\tau)\,A_k} \ldots e^{-g_2(t,\tau)\,A_2}\, e^{-g_1(t,\tau)\,A_1}\,. \tag{3.67}$$

We note that the left-hand side of the resulting equation is a logarithmic derivative:

$$\frac{d}{dt}G(t,\tau)\, G^{-1}(t,\tau) = A(t)\,, \tag{3.68}$$

while the right-hand side has many exponentials that cancel giving the *determining* equation

$$\begin{aligned} A(t) \;=\;& g_1'(t,\tau)\,A_1 \\ & + g_2'(t,\tau)\, e^{g_1(t,\tau)\,A_1}\, A_2\, e^{-g_1(t,\tau)\,A_1} \\ & + g_3'(t,\tau)\, e^{g_1(t,\tau)\,A_1}\, e^{g_2(t,\tau)\,A_2}\, A_3\, e^{-g_2(t,\tau)\,A_2}\, e^{-g_1(t,\tau)\,A_1} \\ & + \cdots \end{aligned} \tag{3.69}$$

This equation along with the initial conditions (3.65) determine the g_i functions.

It is always illuminating to look at any formula when all of the basis elements commute, $[A_i, A_j] - 0$, $1 \leq i, j \leq k$. In this case the determining equations are

$$A(t) = a_1(t)\, A_1 + a_2(t)\, A_2 + \cdots + a_k(t)\, A_k = g_1' A_1 + g_2' A_2 + \cdots + g_k' A_k\,. \tag{3.70}$$

So

$$g_i'(t,\tau) = a(t)\,, \quad g_i(\tau,\tau) = 0\,, \quad 1 \leq i \leq k\,, \tag{3.71}$$

or

$$g_i(t,\tau) = \int_\tau^t a(s)\, ds\,, \quad 1 \leq i \leq k\,. \tag{3.72}$$

In particular, if the a_i are constant, $a_i(t) = a_i$, then

$$g_i(t,\tau) = a_i\,(t-\tau)\,, \quad 1 \leq i \leq k\,. \tag{3.73}$$

Finally, the matrizant has the form (3.64) where the order of the factors doesn't matter. Baum has already made significant use of this special representation, see [4], Appendix C.

The right-hand side of the determining equation can be evaluated using the adjoint-action formula (3.53), but to do this we need a table of the adjoint actions:

$$F_{i,j}(t) = e^{A_i t}\, A_j\, e^{-A_i t}\,, \quad 1 \leq i, j \leq k\,, \tag{3.74}$$

These adjoint actions can be evaluated using either the series in (3.53) or by solving the initial-value problems

$$\frac{d}{dt} F_{i,j}(t) = [A_i, F_{i,j}(t)]\,, \quad F_{i,j}(0) = A_j\,. \tag{3.75}$$

We now have enough theory to construct the functions g_i in the fundamental representation (3.64) of the Green's function. To finish, we need a table of the exponentials of the basis elements

$$G_i(t) = e^{A_i t}\,. \tag{3.76}$$

Again these can be evaluated by either using the series definition of the exponential or the fact that the G_i satisfy the initial-value problems

$$\frac{d}{dt} G_i(t) = A_i\, G_i(t)\,, \quad G_i(0) = I\,. \tag{3.77}$$

The resulting form of the determining equation is the equality of two expression which are linear combinations of the basis elements which says that the coefficient of the basis elements must be equal, which yields k (the number of basis elements) equations in k unknowns g_i. These equations can be solved for g_i' which yields k nonlinear differential equations. Now this seems like a bad deal. We started with n^2 linear variable-coefficient differential equations and end up with $k \leq n^2$ non-linear constant coefficient differential equations (which are difficult to solve even in the simplest cases). However, we have done this in such a way that if we choose the g_i as arbitrary functions, then we produce an $A(t)$ which gives a system that is exactly solvable with the solution of the homogeneous equations given by a known Green's function $G(t, 0)$. As we will see, this can be used to great advantage in studying physical systems.

4 Examples

In this section we will look at some one, two and three dimensional examples. Note that many of the computations were done using a computer algebra system. None of the computations are difficult, just that there are quite a few elementary computations. The computer algebra programs are available from S. Steinberg.

4.1 One-Dimensional Example

One dimensional examples are scalar equations whose Green's function satisfies

$$\frac{d}{dt}G(t,\tau) = A(t)\,G(t,\tau)\,, \quad G(\tau,\tau) = 1\,. \tag{4.78}$$

The solution of this initial value problem is

$$G(t,\tau) = e^{\int_\tau^t a(s)\,ds}\,. \tag{4.79}$$

One view of our job is to find formulas like this for higher-dimensional examples.

4.2 Two-Dimensional Examples

We now do a complete analysis of systems of two ordinary differential equations, that is, we assume that

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}, \tag{4.80}$$

and that $A(t)$ is a real matrix, but similar results are true for complex matrices.

4.2.1 The Obvious Basis

The first step is to choose a basis for the 2 by 2 matrices. An obvious simple basis is

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{4.81}$$

The matrix A is represented in this basis as

$$A(t) = a(t)\,A_1 + d(t)\,A_2 + b(t)\,A_3 + c(t)\,A_4\,. \tag{4.82}$$

The formula for the Green's function involves the exponential of the basis elements, so we will need these:

$$e^{t\,A_1} = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}, \quad e^{t\,A_2} = \begin{pmatrix} 1 & 0 \\ 0 & e^t \end{pmatrix}, \quad e^{t\,A_3} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad e^{t\,A_4} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}. \tag{4.83}$$

These formulas can be easily evaluated in two different ways: one is to compute a few terms of the power series definition and then identify the resulting series; the other is to convert the evaluation to an initial-value problem. For example, the power-series method gives

$$e^{tA_1} = \begin{pmatrix} 1+t+t^2/2+t^3/6+t^4/24+t^5/120+\cdots & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}, \tag{4.84}$$

while if we set

$$X(t) = e^{tA_1} = \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix}, \tag{4.85}$$

then

$$X'(t) = A_1 X(t), \quad X(0) = I. \tag{4.86}$$

or

$$\begin{pmatrix} r'(t) & s'(t) \\ u'(t) & v'(t) \end{pmatrix} = \begin{pmatrix} r(t) & s(t) \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} r(0) & s(0) \\ u(0) & v(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.87}$$

So $r(t) = e^t$, $s(t) = 0$, $u(t) = 0$ and $v(t) = 1$, as was found from the power series method.

The commutator table for the basis, with entries $[A_i, A_j]$, is given by

A_i \ A_j	A_1	A_2	A_3	A_4
A_1	0	0	A_3	$-A_4$
A_2	0	0	$-A_3$	A_4
A_3	$-A_3$	A_3	0	$A_1 - A_2$
A_4	A_4	$-A_4$	$-A_1 + A_2$	0

(4.88)

which is fairly simple. Next we need a table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra:

A_i \ A_j	A_1	A_2	A_3	A_4
A_1	A_1	A_2	$e^t A_3$	$e^{-t} A_4$
A_2	A_1	A_2	$e^{-t} A_3$	$e^t A_4$
A_3	$A_1 - t A_3$	$A_2 + t A_3$	A_3	$t A_1 - t A_2 - t^2 A_3 + A_4$
A_4	$A_1 + t A_4$	$A_2 - t A_4$	$-t A_1 + t A_2 + A_3 - t^2 A_4$	A_4

(4.89)

This was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas but can also be done by converting the evaluation to a initial-value problem for a system of ordinary differential equations.

Now we represent the Green's function of the the differential equation in the form (3.64):

$$G(t,\tau) = e^{g_1(t,\tau) A_1} e^{g_2(t,\tau) A_2} e^{g_3(t,\tau) A_3} e^{g_4(t,\tau) A_4}, \tag{4.90}$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.83) give the Green's function as a product of four matrices:

$$G(t,\tau) = \begin{pmatrix} e^{g_1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{g_2} \end{pmatrix} \begin{pmatrix} 1 & g_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ g_4 & 1 \end{pmatrix}. \tag{4.91}$$

The determining equation (3.69), which are computed by taking the logarithmic derivative of the Green's function and then using the table of adjoint actions, is

$$A(t) = (g_1' + g_3\, g_4')\, A_1 + (g_2' - g_3\, g_4')\, A_2 + (e^{g_1-g_2}\, g_3' - e^{g_1-g_2}\, g_3^2\, g_4')\, A_3 + e^{-g_1+g_2}\, g_4'\, A_4\,. \quad (4.92)$$

Next we solve equations (4.92) with $A(t)$ given by (4.80) for g_i' to obtain the following nonlinear system of differential equations for the g_i:

$$\begin{aligned} g_1' &= a - c\, e^{g_1-g_2}\, g_3\,, \\ g_2' &= d + c\, e^{g_1-g_2}\, g_3\,, \\ g_3' &= b\, e^{-g_1+g_2} + c\, e^{g_1-g_2}\, g_3^2\,, \\ g_4' &= c\, e^{g_1-g_2}\,. \end{aligned} \quad (4.93)$$

Even when $A(t)$ is constant, that is, a, b, c, and d are constants, the previous ordinary differential equations are seriously nonlinear and difficult to solve. However, we can create many solvable examples by choosing g_i as some function of t and τ. If we only consider the initial value problem without source terms and choose

$$g_1(t) = \alpha\, t\,, \quad g_2(t) = \beta\, t\,, \quad g_3(t) = \gamma\, t\,, \quad g_4(t) = \delta\, t\,, \quad (4.94)$$

then the determining equations (4.92) become

$$A(t) = \begin{pmatrix} \alpha + \delta\,\gamma\, t & e^{(\alpha-\beta)\,t}\,\gamma - \delta\, e^{(\alpha-\beta)\,t}\,\gamma^2\, t^2 \\ \delta\, e^{(-\alpha+\beta)\,t} & \beta - \delta\,\gamma\, t \end{pmatrix}. \quad (4.95)$$

If we assume the ordinary differential equation doesn't have a forcing term, then we only need $G(t) = G(t,0)$ which is given by (4.91):

$$G(t) = \begin{pmatrix} e^{\alpha t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{\beta t} \end{pmatrix} \begin{pmatrix} 1 & \gamma t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta t & 1 \end{pmatrix} = \begin{pmatrix} e^{\alpha t}\,(1 + \delta\,\gamma\, t^2) & e^{\alpha t}\,\gamma\, t \\ \delta\, e^{\beta t}\, t & e^{\beta t} \end{pmatrix}, \quad (4.96)$$

In the applications that interest us, it is common for $A(t)$ to have trace zero or be skew symmetric (which implies trace zero). As none of our basis elements are skew symmetric and only two elements, not three, have trace zero, these situations are hard to analyze. The next basis we choose will correct this.

4.2.2 Second Basis

We now choose a basis with three trace zero matrices and one skew-symmetric matrix:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (4.97)$$

The matrix A given in (4.80) is represented in this basis as

$$A(t) = \frac{a(t) + d(t)}{2}\, A_1 + \frac{a(t) - d(t)}{2}\, A_2 + \frac{b(t) + c(t)}{2}\, A_3 + \frac{b(t) - c(t)}{2}\, A_4\,. \quad (4.98)$$

The exponentials of the basis elements are

$$e^{tA_1} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}, \quad e^{tA_2} = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix},$$

$$e^{tA_3} = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}, \quad e^{tA_4} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}. \tag{4.99}$$

These forms seem more relevant to the study of physical problems than those given by the simple basis used in the previous section.

The commutator table with entries $[A_i, A_j]$ is given by

A_i \ A_j	A_1	A_2	A_3	A_4
A_1	0	0	0	0
A_2	0	0	$2A_4$	$2A_3$
A_3	0	$-2A_4$	0	$-2A_2$
A_4	0	$-2A_3$	$2A_2$	0

(4.100)

which is even simpler. Next we need a table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra:

A_i \ A_j	A_1	A_2	A_3	A_4
A_1	A_1	A_2	A_3	A_4
A_2	A_1	A_2	$\cosh(2t)A_3 + \sinh(2t)A_4$	$\cosh(2t)A_4 + \sinh(2t)A_3$
A_3	A_1	$\cosh(2t)A_2 - \sinh(2t)A_4$	A_3	$\cosh(2t)A_4 - \sinh(2t)A_2$
A_4	A_1	$\cos(2t)A_2 - \sin(2t)A_3$	$\cos(2t)A_3 + \sin(2t)A_2$	A_4

(4.101)

As before this was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas.

We now illustrate how to use an ordinary differential equation trick to evaluate one of the previous adjoint actions. If

$$X(t) = e^{A_2 t} A_3 e^{-A_2 t} = e^{[A_2,\diamond]} A_3\,, \tag{4.102}$$

then

$$X'(t) = [A_2, X(t)]\,, \quad X(0) = A_3\,,. \tag{4.103}$$

If we set

$$X(t) = \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix} \tag{4.104}$$

then we need

$$\begin{pmatrix} r'(t) & s'(t) \\ u'(t) & v'(t) \end{pmatrix} = \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix} \right] = \begin{pmatrix} 0 & 2s(t) \\ 2u(t) & 0 \end{pmatrix}, \tag{4.105}$$

and

$$\begin{pmatrix} r(0) & s(0) \\ u(0) & v(0) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.106}$$

This gives

$$X(t) = \begin{pmatrix} 0 & e^{2t} \\ e^{-2t} & 0 \end{pmatrix} = e^{2t}\frac{A_3 + A_4}{2} + e^{-2t}\frac{A_3 - A_4}{2} = \cosh(2t)\, A_3 + \sinh(2t)\, A_4. \tag{4.107}$$

Of course, this checks with the power-series solution.

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t,\tau) = e^{g_1(t,\tau)\, A_1}\, e^{g_2(t,\tau)\, A_2}\, e^{g_3(t,\tau)\, A_3}\, e^{g_4(t,\tau)\, A_4}, \tag{4.108}$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.99) give the Green's function as a product of four matrices:

$$G = \begin{pmatrix} e^{g_1} & 0 \\ 0 & e^{g_1} \end{pmatrix} \begin{pmatrix} e^{g_2} & 0 \\ 0 & e^{-g_2} \end{pmatrix} \begin{pmatrix} \cosh(g_3) & \sinh(g_3) \\ \sinh(g_3) & \cosh(g_3) \end{pmatrix} \begin{pmatrix} \cos(g_4) & \sin(g_4) \\ -\sin(g_4) & \cos(g_4) \end{pmatrix}. \tag{4.109}$$

The determining equations (3.69) with $A(t)$ given by (4.80) have the form

$$\begin{aligned} a &= g_1' + g_2' - \sinh(2\, g_3)\, g_4', \\ b &= \cosh(2\, g_2)\, g_3' + \sinh(2\, g_2)\, g_3' + \cosh(2\, g_2)\, \cosh(2\, g_3)\, g_4' + \cosh(2\, g_3)\, \sinh(2\, g_2)\, g_4', \\ c &= \cosh(2\, g_2)\, g_3' - \sinh(2\, g_2)\, g_3' - \cosh(2\, g_2)\, \cosh(2\, g_3)\, g_4' + \cosh(2\, g_3)\, \sinh(2\, g_2)\, g_4', \\ d &= g_1' - g_2' + \sinh(2 * g_3) * g_4'. \end{aligned} \tag{4.110}$$

We solve equations (4.110) for g_i' to obtain the following nonlinear system of differential equations for the g_i:

$$\begin{aligned} g_1' &= \frac{a+d}{2}, \\ g_2' &= \frac{a-d}{2} + \left(\frac{b-c}{2}\cosh(2\, g_2) - \frac{b+c}{2}\sinh(2\, g_2)\right)\tanh(2\, g_3), \\ g_3' &= \frac{b+c}{2}\cosh(2\, g_2) - \frac{b-c}{2}\sinh(2\, g_2), \\ g_4' &= \left(\frac{b-c}{2}\cosh(2\, g_2) - \frac{b+c}{2}\sinh(2\, g_2)\right)\operatorname{sech}(2\, g_3). \end{aligned} \tag{4.111}$$

As before, even when $A(t)$ is constant these equations are seriously nonlinear and difficult to solve. But again, we can create many solvable examples by choosing the g_i functions appropriately. Again, for an example, we choose

$$g_1(t) = \alpha t, \quad g_2(t) = \beta t, \quad g_3(t) = \gamma t, \quad g_4(t) = \delta t, \tag{4.112}$$

and then the determining equations become

$$
\begin{aligned}
a(t) &= \alpha+\beta-\delta\,\sinh(2\gamma t)\,,\\
b(t) &= \gamma\cosh(2\beta t)+\delta\cosh(2\beta t)\cosh(2\gamma t)+\gamma\sinh(2\beta t)+\delta\cosh(2\gamma t)\sinh(2\beta t)\,,\\
c(t) &= \gamma\cosh(2\beta t)-\delta\cosh(2\beta t)\cosh(2\gamma t)-\gamma\sinh(2\beta t)+\delta\cosh(2\gamma t)\sinh(2\beta t)\,,\\
d(t) &= \alpha-\beta+\delta\,\sinh(2\gamma t)\,.
\end{aligned}
\tag{4.113}
$$

If we assume the ordinary differential equation doesn't have a forcing term, then we only need $G(t)=G(t,0)$ which is given by (4.109):

$$
G(t)=\begin{pmatrix} e^{\alpha t} & 0 \\ 0 & e^{\alpha t}\end{pmatrix}\begin{pmatrix} e^{\beta t} & 0 \\ 0 & e^{-\beta t}\end{pmatrix}\begin{pmatrix} \cosh(\gamma t) & \sinh(\gamma t) \\ \sinh(\gamma t) & \cosh(\gamma t)\end{pmatrix}\begin{pmatrix} \cos(\delta t) & \sin(\delta t) \\ -\sin(\delta t) & \cos(\delta t)\end{pmatrix},
\tag{4.114}
$$

or

$$
\begin{aligned}
g_{1,1} &= e^{\alpha t+\beta t}\left(\cos(\delta t)\cosh(\gamma t)-\sin(\delta t)\sinh(\gamma t)\right),\\
g_{1,2} &= e^{\alpha t+\beta t}\left(\cosh(\gamma t)\sin(\delta t)+\cos(\delta t)\sinh(\gamma t)\right),\\
g_{2,1} &= -e^{\alpha t-\beta t}\left(\cosh(\gamma t)\sin(\delta t)-\cos(\delta t)\sinh(\gamma t)\right),\\
g_{2,2} &= e^{\alpha t-\beta t}\left(\cos(\delta t)\cosh(\gamma t)+\sin(\delta t)\sinh(\gamma t)\right).
\end{aligned}
\tag{4.115}
$$

In the applications that interest us, it is common for $A(t)$ to have trace zero, that is, $d=-a$, so that $g_1'=0$ and then $g_1(t,\tau)\equiv 0$. Another way of saying the same thing is that A_2, A_3 and A_4 form a basis for the traceless matrices. In any case, this doesn't make much of a simplification.

Another important case is where $A(t)$ is skew symmetric, that is, $a=d=0$ and $c=-b$:

$$
\begin{aligned}
g_1' &= 0\,,\\
g_2' &= b\cosh(2g_2)\tanh(2g_3)\,,\\
g_3' &= -b\sinh(2g_2)\,,\\
g_4' &= b\cosh(2g_2)\,\mathrm{sech}(2g_3)\,.
\end{aligned}
\tag{4.116}
$$

Although this is an impressive system of equations, its solution can easily be seen to be

$$
g_1(t,\tau)=0\,,\quad g_2(t,\tau)=0\,,\quad g_3(t,\tau)=0\,,\quad g_4(t,\tau)=\int_\tau^t b(s)\,ds\,.
\tag{4.117}
$$

This result is far easier to see starting from the fact that the Hermitian 2 by 2 matrices form a one dimensional space with A_4 being a basis.

4.2.3 The Pauli Spin Basis

We now look at 2 by 2 skew-symmetric matrices and choose the Pauli Spin matrices for a basis:

$$
A_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad A_2=\begin{pmatrix}0&-\jmath\\\jmath&0\end{pmatrix},\quad A_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.
\tag{4.118}
$$

Then any skew-symmetric matrix can be represent by

$$A(t) = a(t)\,A_1 + b(t)\,A_2 + c(t)\,A_3 = \begin{pmatrix} c(t) & a(t) - \jmath\, b(t) \\ a(t) + \jmath\, b(t) & -c(t) \end{pmatrix} \tag{4.119}$$

The exponentials of the basis elements are

$$\begin{aligned} e^{t A_1} &= \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}, \\ e^{t A_2} &= \begin{pmatrix} \cosh(t) & -\jmath \sinh(t) \\ \jmath \sinh(t) & \cosh(t) \end{pmatrix}, \\ e^{t A_3} &= \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}. \end{aligned} \tag{4.120}$$

The commutator table with entries $[A_i, A_j]$ is given by

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	0	$2\jmath A_3$	$-2\jmath A_2$
A_2	$-2\jmath A_3$	0	$2\jmath A_1$
A_3	$2\jmath A_2$	$-2\jmath A_1$	0

(4.121)

Next we need a table of the adjoint actions $e^{A_i t} A_j\, e^{-A_i t}$ of the Lie group on the Lie algebra:

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	A_1	$A_2 \cosh(2t) + \jmath A_3 \sinh(2t)$	$A_3 \cosh(2t) - \jmath A_2 \sinh(2t)$
A_2	$A_1 \cosh(2t) - \jmath A_3 \sinh(2t)$	A_2	$A_3 \cosh(2t) + \jmath A_1 \sinh(2t)$
A_3	$A_1 \cosh(2t) + \jmath A_2 \sinh(2t)$	$A_2 \cosh(2t) - \jmath A_1 \sinh(2t)$	A_3

(4.122)

As before this was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas.

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t, \tau) = e^{g_1(t,\tau) A_1}\, e^{g_2(t,\tau) A_2}\, e^{g_3(t,\tau) A_3}, \tag{4.123}$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.120) give the Green's function as a product of three matrices:

$$G = \begin{pmatrix} \cosh(g_1) & \sinh(g_1) \\ \sinh(g_1) & \cosh(g_1) \end{pmatrix} \begin{pmatrix} \cosh(g_2) & -\jmath \sinh(g_2) \\ \jmath \sinh(g_2) & \cosh(g_2) \end{pmatrix} \begin{pmatrix} e^{g_3} & 0 \\ 0 & e^{-g_3} \end{pmatrix}. \tag{4.124}$$

The determining equations (3.69) with $A(t)$ given by (4.119) have the form

$$\begin{aligned} a(t) &= g_1' + \jmath \sinh(2\, g_2(t))\, g_3', \\ b(t) &= \cosh(2\, g_1(t))\, g_2' - \jmath \cosh(2\, g_2(t)) \sinh(2\, g_1(t))\, g_3', \\ c(t) &= \jmath \sinh(2\, g_1(t))\, g_2' + \cosh(2\, g_1(t)) \cosh(2\, g_2(t))\, g_3'. \end{aligned} \tag{4.125}$$

We solve equations (4.125) for g_i' to obtain the following nonlinear system of differential equations for the g_i:

$$
\begin{aligned}
g_1'(t) &= a\cosh(2\,g_1(t))^2 - \jmath\, c\cosh(2\,g_1(t))\,\tanh(2\,g_2(t)) \\
&\quad - \sinh(2\,g_1(t))\,\left(a\sinh(2\,g_1(t)) + b\tanh(2\,g_2(t))\right) \\
g_2'(t) &= b\cosh(2\,g_1(t)) + \jmath\, c\sinh(2\,g_1(t)) \\
g_3'(t) &= \operatorname{sech}(2\,g_2(t))\,\left(c\cosh(2\,g_1(t)) - \jmath\, b\sinh(2\,g_1(t))\right)
\end{aligned}
\tag{4.126}
$$

4.3 Three-Dimensional Example

The full three-dimensional case involves nine parameters and thus there isn't much use in tabulating all of this. However, the skew-symmetric 3 by 3 matrices are only three dimensional. A convenient basis is

$$
A_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{4.127}
$$

Then any skew-symmetric matrix $A(t)$ can be represented as

$$
A(t) = a(t)\,A_1 + b(t)\,A_2 + c(t)\,A_3 = \begin{pmatrix} 0 & a(t) & c(t) \\ -a(t) & 0 & b(t) \\ -c(t) & -b(t) & 0 \end{pmatrix}. \tag{4.128}
$$

The exponentials of the basis elements are

$$
\begin{aligned}
e^{tA_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & \sin(t) \\ 0 & -\sin(t) & \cos(t) \end{pmatrix}, \\
e^{tA_2} &= \begin{pmatrix} \cos(t) & 0 & -\sin(t) \\ 0 & 1 & 0 \\ \sin(t) & 0 & \cos(t) \end{pmatrix}, \\
e^{tA_3} &= \begin{pmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
\tag{4.129}
$$

These matrices describe a rotation about each of the axes.

The commutator table with entries $[A_i, A_j]$ is given by

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	0	A_3	$-A_2$
A_2	$-A_3$	0	A_1
A_3	A_2	$-A_1$	0

(4.130)

which is well known for the rotations. The table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra are:

A_i \ A_j	A_1	A_2	A_3
A_1	A_1	$A_2 \cos(t) + A_3 \sin(t)$	$A_3 \cos(t) - A_2 \sin(t)$
A_2	$A_1 \cos(t) - A_3 \sin(t)$	A_2	$A_3 \cos(t) + A_1 \sin(t)$
A_3	$A_1 \cos(t) + A_2 \sin(t)$	$A_2 \cos(t) - A_1 \sin(t)$	A_3

(4.131)

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t,\tau) = e^{g_1(t,\tau) A_1}\, e^{g_2(t,\tau) A_2}\, e^{g_3(t,\tau) A_3}, \tag{4.132}$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.129) give the Green's function as a product of three matrices:

$$\begin{aligned} G(t,\tau) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(g_1(t,\tau)) & \sin(g_1(t,\tau)) \\ 0 & -\sin(g_1(t,\tau)) & \cos(g_1(t,\tau)) \end{pmatrix} \\ &\quad \begin{pmatrix} \cos(g_2) & 0 & -\sin(g_2) \\ 0 & 1 & 0 \\ \sin(g_2) & 0 & \cos(g_2) \end{pmatrix} \\ &\quad \begin{pmatrix} \cos(g_3) & \sin(g_3) & 0 \\ -\sin(g_3) & \cos(g_3) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{4.133}$$

The determining equations (3.69) can be evaluated using the adjoint action table to give

$$\begin{aligned} A(t) &= (g_1' + \sin(g_2) * g_3')\, A_1 \\ &\quad + (\cos(g_1) * g_2' - \cos(g_2) * \sin(g_1) * g_3')\, A_2 \\ &\quad + (\sin(g_1) * g_2' + \cos(g_1) * \cos(g_2) * g_3')\, A_3 \end{aligned} \tag{4.134}$$

With $A(t)$ given by (4.128), these equations can be solved for g_i' to obtain the following nonlinear system of differential equations for the g_i:

$$\begin{aligned} g_1' &= a - \tan(g_2)\, (c \cos(g_1) + b \sin(g_1)), \\ g_2' &= b \cos(g_1) + c \sin(g_1), \\ g_3' &= \sec(g_2)\, (c \cos(g_1) - b \sin(g_1)). \end{aligned} \tag{4.135}$$

The nonlinear system of equations is not easy to solve, but as before, we can create a large family of solvable examples by choosing the g_i functions. If we look only at the initial value problem (no source terms), and if we choose the a, b, and c that determine $A(t)$ in (4.134) as

$$\begin{aligned} a(t) &= \alpha + \gamma \sin(\beta t), \\ b(t) &= \beta \cos(\alpha t) - \gamma \cos(\beta t) \sin(\alpha t), \\ c(t) &= \gamma \cos(\alpha t) \cos(\beta t) + \beta \sin(\alpha t), \end{aligned} \tag{4.136}$$

then the greens function $G(t) = G(t,0)$ is given by the product of three simple matrices (4.133):

$$
\begin{aligned}
G(t) \;=\; & \begin{pmatrix} \cos(\alpha\, t) & \sin(\alpha\, t) & 0 \\ -\sin(\alpha\, t) & \cos(\alpha\, t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
& \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta\, t) & \sin(\beta\, t) \\ 0 & -\sin(\beta\, t) & \cos(\beta\, t) \end{pmatrix} \\
& \begin{pmatrix} \cos(\gamma\, t) & 0 & \sin(\gamma\, t) \\ 0 & 1 & 0 \\ -\sin(\gamma\, t) & 0 & \cos(\gamma\, t) \end{pmatrix} .
\end{aligned}
\tag{4.137}
$$

So, in summary, we can use Lie theory to convert systems of ordinary differential equations with variable coefficients into systems of nonlinear equations that are difficult to solve. However, the resulting formulas allow us to create a large number of solvable examples (along with the solutions).

5 Product Integrals and Lie Theory

We now turn to finding more useful representations of product integrals, so as in Section 3.2 we suppose that $A(t)$ is given continuous matrix-valued function where the matrices belong to a Lie algebra $\mathcal{A}$ that has a basis A_i, $1 \le i \le k$ so that

$$A = \sum_{i=1}^{k} a_i(t)\, A_i\,. \tag{5.138}$$

Recall that the Green's function can be written as a product integral

$$G(t,\tau) = \prod_{\tau}^{t} e^{A(s)\,ds}\,, \tag{5.139}$$

as in Equation (1.7).

The product integral is really a time ordered integral. So If $a \le b$, then the product integral can be defined by

$$\prod_{a}^{b} e^{A(s)\,ds} = \lim_{h\to 0} \prod_{i=0}^{N-1} e^{h\,A(a+i\,h)}\,, \tag{5.140}$$

where $h = (b-a)/N$. The order of the factors in this formula are critical:

$$\prod_{i=0}^{N-1} e^{h\,A(i\,h)} = e^{A(a+(N-1)\,h))}\, e^{A(a+(N-2)\,h)} \cdots e^{A(a)\,h}\,. \tag{5.141}$$

The other order of factors is obtained by defining

$$\prod_{b}^{a} e^{A(s)\,ds} = \lim_{h\to 0} \prod_{i=0}^{N-1} e^{h\,A(a+i\,h)}\,, \tag{5.142}$$

where now $h = (a-b)/N$, that is $h \le 0$ and

$$\prod_{i=0}^{N-1} e^{h\,A(i\,h)} = e^{A(a)\,h} \cdots e^{A(a+(N-2)\,h)}\, e^{A(a+(N-1)\,h))}\,. \tag{5.143}$$

So we have that

$$\prod_{b}^{a} e^{A(s)\,ds} \prod_{a}^{b} e^{A(s)\,ds} = \prod_{a}^{b} e^{A(s)\,ds} \prod_{b}^{a} e^{A(s)\,ds} = I\,, \tag{5.144}$$

no matter what the ordering of a and b are, because the factors is the definition of the products integrals exactly cancel. In particular, this implies that the product integral is always invertible and we have a simple formula for the inverse. More generally, we have for all a, b, and c that

$$\prod_{a}^{b} e^{A(s)\,ds} = \prod_{c}^{b} e^{A(s)\,ds} \prod_{a}^{c} e^{A(s)\,ds}\,. \tag{5.145}$$

The way the upper and lower indices work precisely captures the time ordering of the product integral.

Now the first thing to observe is that if A is a constant, then the approximations to the product integral can be evaluated explicitly by writing the product of exponential in the definition as an exponential of a sum to give

$$\prod_a^b e^{A(s)\,ds} = e^{A\,(b-a)}\,. \tag{5.146}$$

Unfortunately, this is not the representation used in our version of Lie theory. So let us try to find $a_i(t)$ such that

$$e^{A\,t} = \prod_{i=1}^{k} e^{a_i(t)\,A_i}\,. \tag{5.147}$$

However, this is just a simple variant of the problem we had in computing the determining equations (3.69) have seen before, so we take the logarithmic derivatives to get

$$\begin{aligned} A \;=\; & a_1'(t)\,A_1 \\ & +a_2'(t)\,e^{a_1(t)\,A_1}\,A_2\,e^{-a_1(t)\,A_1} \\ & +a_3'(t)\,e^{a_1(t)\,A_1}\,e^{a_2(t)\,A_2}\,A_3\,e^{-a_2(t)\,A_2}\,e^{-a_1(t)\,A_1} \\ & +\cdots \end{aligned} \tag{5.148}$$

The stumbling block here is the solution of the severely nonlinear equations for the a_i.

In the case that $A(t)$ does depend on t, the product integral does not give us an explicit representation in terms of exponentials, but there are formulas that do give nice insight into the structure of the solutions of ordinary differential equations. First, the product integral is explicitly designed so that

$$\frac{d}{dt}\prod_0^t e^{A(s)\,ds} = A(t)\prod_0^t e^{A(s)\,ds}\,, \quad \prod_0^0 e^{A(s)\,ds} = I\,. \tag{5.149}$$

We can gain some insight into product integrals by proving this. So if

$$M(t) = \prod_0^t e^{A(s)\,ds}\,, \tag{5.150}$$

then

$$\frac{dM}{dt}(t) = \lim_{\Delta t\to 0}\frac{M(t+\Delta t)-M(t)}{\Delta t}\,. \tag{5.151}$$

Using Formula (5.145) which depends critically on the ordering of the products in the product integral, we obtain

$$\prod_0^{t+\Delta t} e^{A(s)\,ds} = \prod_t^{t+\Delta t} e^{A(s)\,ds}\prod_0^t e^{A(s)\,ds}\,, \tag{5.152}$$

so that

$$\frac{dM}{dt}(t) = \lim_{\Delta t \to 0} \frac{\prod_t^{t+\Delta t} e^{A(s)\,ds} - I}{\Delta t} M(t) . \tag{5.153}$$

But for small Δt

$$\prod_t^{t+\Delta t} e^{A(s)\,ds} \approx I + \Delta t\, A , \tag{5.154}$$

so the limit evaluates to

$$\frac{dM}{dt}(t) = A(t)\, M(t) , \tag{5.155}$$

which is what was to be proved.

Now if we introduce the logarithmic, product, or multiplicative derivative [6] as

$$D_t M(t) = M'(t)\, M^{-1}(t) , \tag{5.156}$$

then formula (5.149) can be written as

$$D_t \prod_0^t e^{A(s)\,ds} = A(t) , \tag{5.157}$$

which says the logarithm derivative of the product integral is the identity, just like the Fundamental Theorem of Calculus says that the derivative of the integral is the identity. The analog of the other half of the Fundamental Theorem is

$$\prod_0^t e^{D_s A(s)\,ds} = A(t)\, A^{-1}(0) . \tag{5.158}$$

We must assume that $A(t)$ is nonsingular for the logarithmic derivative to make sense. The proof of this fact is simple:

$$A'(t) = A'(t)\, A^{-1}(t)\, A(t) = D_t A'(t)\, A(t) , \tag{5.159}$$

and the solution of this differential equation is given by

$$A(t) = \prod_0^t e^{D_s A(s)\,ds}\, A(0) , \tag{5.160}$$

which establishes the result.

Next, one can apply the formula in (2.25) to the definition of the product integral to obtain

Lemma:

$$\det \prod_a^b e^{A(s)\,ds} = e^{\int_a^b \mathrm{tr} A(s)\,ds} , \tag{5.161}$$

which shows that the determinant of a product integral is never zero. This completes the basic properties of the product integral.

Because of the probable non-commutivity of the matrices, there is one more important result, called the *sum rule*, that gives the analog of the scalar formula $\exp a + b = \exp(a)\ \exp(b)$ and was used by Baum [5]. If

$$M(t) = \prod_0^t e^{A(s)\,ds}, \tag{5.162}$$

then

$$\prod_0^t e^{(A(s)+B(s))\,ds} = \prod_0^t e^{A(s)\,ds} \prod_0^t e^{M^{-1}(s)\,B(s)\,M(s)\,ds}. \tag{5.163}$$

The proof of this is to differentiate both sides of the equation to see that these are the same and noting that both sides are the identity at $t = 0$.

6 Computing With Exponential

One of the reasons that Lie techniques are a powerful tool for computation is that they encourage the use of exponential identities; some of these important formulas are derived here. The techniques introduced in this section are fundamental to the derivation of all exponential identities and are based on the development in Steinberg [10]. The derivations in this section are based on formal calculations for general noncommuting symbols typically denoted by A, B, C, and so forth, but the reader may consider these matrices. Thus it is important to remember that

$$A\,B = B\,A \tag{6.164}$$

is *not* true for all A and B. However, some of the noncommuting variables may commute under multiplication, so any identities that are derived must reduce to standard identities when the variables commute. It is relatively easy to check some of the identities, or to find counter-examples to incorrect identities, by using 2 by 2 matrices. Lower case Roman variables, a, b, c and so forth will stand for scalars and thus $a\,b = b\,a$.

Here is a simple but important identity:

$$(A+B)^2 = A^2 + A\,B + B\,A + B^2\,. \tag{6.165}$$

Another way of writing this is

$$(A+B)^2 = (A^2 + 2A\,B + B^2) + B\,A - A\,B\,. \tag{6.166}$$

If we define the *commutator* of A and B by

$$[A, B] = AB - BA\,, \tag{6.167}$$

then

$$(A+B)^2 = (A^2 + 2A\,B + B^2) - [A, B]. \tag{6.168}$$

The previous formula records an identity as a standard identity for commuting variables plus a deviation that involves only *commutators* of the variables. Thus it is trivial to check that the identity is true for commuting variables as the commutator is then zero. We will always put identities in such a form.

The commutator satisfies an important set of identities:

Skew-Symmetry of the Commutator:

$$[A, B] = -[B, A]\,; \tag{6.169}$$

Linearity of the Commutator:

$$[A, b\,B + c\,C] = b\,[A, B] + c\,[A, C]\,; \tag{6.170}$$

Product Rule for the Commutator:

$$[A, B\,C] = [A, B]\,C + B\,[A, C]\,; \tag{6.171}$$

Jacobi Identity for the Commutator:

$$[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0. \tag{6.172}$$

The proofs of these identities are simple computations. These identities have an irritating consequence: expressions involving commutators do not have a unique form. In fact, as the number of commutators in a term increases, so do the possible representations of the expression.

As with matrices, the exponential of a non-commuting variable is defined by the power series

$$e^A=\sum_{k=0}^{\infty}\frac{A^k}{k!}. \tag{6.173}$$

Again, we make no assumptions on the convergence of the series; consequently all computations in this section are formal. Here formal means that two series are equal if they are equal term by term. Calculations that involve such series are tedious, so we have implemented a computer algebra program to do some of these calculations. One important identity which is true for noncommuting variables is

$$e^Ae^{-A}=I=e^{-A}e^A. \tag{6.174}$$

This identity is obvious; it only involves one noncommuting variable, so the fact that the variable does not commute with other noncommuting variables plays no role here. Thus the identity can be checked using the power series expansion for commuting variables. This identity shows that the multiplicative inverse B^{-1} of $B=e^A$ is $B^{-1}=e^{-A}$.

Is it true that

$$e^Ae^B=e^Be^A? \tag{6.175}$$

This can be checked using the power series expansion:

$$\begin{aligned} e^Ae^B &= (I+A+\frac{A^2}{2}+\cdots)(I+B+\frac{B^2}{2}+\cdots) \\ &= I+(A+B)+(\frac{1}{2}A^2+AB+\frac{1}{2}B^2)+\cdots. \end{aligned} \tag{6.176}$$

Interchanging A and B gives

$$e^Be^A=I+(B+A)+(\frac{1}{2}B^2+BA+\frac{1}{2}A^2)+\cdots \tag{6.177}$$

and then

$$e^Ae^B-e^Be^A=AB-BA+\cdots=[A,B]+\cdots. \tag{6.178}$$

Thus, because A and B do not commute, the exponentials do not commute.

Another way to see the same thing and also produce a useful formula is to multiply (6.175) by $\exp(-A)\exp(-B)$, and ask whether it is true that

$$e^{-A}e^{-B}e^Ae^B=I? \tag{6.179}$$

This can be checked by expanding the left hand side of the equation:

$$\begin{aligned} e^{-A}e^{-B}e^{A}e^{B} &= (I - A + \frac{A^2}{2} - \cdots) \\ &\quad (I - B + \frac{B^2}{2} - \cdots) \\ &\quad (I + A + \frac{A^2}{2} + \cdots) \\ &\quad (I + B + \frac{B^2}{2} + \cdots) \\ &= I + AB - BA + \cdots \\ &= I + [A, B] + \cdots . \end{aligned} \tag{6.180}$$

Thus, through quadratic expressions in A and B,

$$e^{-A}e^{-B}e^{A}e^{B} \approx e^{[A,B]}, \tag{6.181}$$

or

$$e^{A}e^{B} \approx e^{B}e^{A}e^{[A,B]}. \tag{6.182}$$

Note that any time we find part of a power series for an exponential, as in $I + [A, B] + \cdots$, we replace the power series by the exponential. If more terms in the series are computed there are more correction. The higher order calculations are very lengthy.

Another important question is is the exponential identity

$$e^{A+B} = e^{A}e^{B} \tag{6.183}$$

true? As above, the power series of each side of the above expression do not agree, so this identity cannot be true for noncommuting variables. Another way to see this, and also obtain part of an important formula, is to do a power series computation:

$$e^{-B}e^{-A}e^{A+B} \approx I - \frac{[A, B]}{2} + \cdots . \tag{6.184}$$

Thus through quadratic expressions in A and B,

$$e^{-B}e^{-A}e^{A+B} \approx e^{-[A,B]/2}, \tag{6.185}$$

or

$$e^{A+B} \approx e^{A}e^{B}e^{-[A,B]/2}. \tag{6.186}$$

Again, computing the higher-order corrections to this formula is tedious. The use of power series to compute such exponential identities is rather inefficient, so we now turn to another method.

As we have seen before, a powerful technique for deriving exponential identities involves differentiating noncommuting exponentials. Our first result is the same as (3.50) and shows that if the exponent depends in a simple way on t, then the t derivative of the exponential is natural. When the exponent depends in an arbitrary way on t, then the t derivative is more complicated. Let us begin with a simple case.

Proposition. If $a(t)$ is a scalar function of t and $a'(t) = da(t)/dt$, then

$$\frac{d}{dt}e^{a(t)A} = a'(t)Ae^{a(t)A}. \tag{6.187}$$

Proof. We have not required the exponential series to be convergent, so this is simply a formal statement; the series must be equal term by term. Of course, if the series converge, then the equality holds for the sum of the series. The proof is a computation:

$$\begin{aligned}
\frac{d}{dt}e^{a(t)A} &= \frac{d}{dt}\sum_{k=0}^{\infty}\frac{a^k(t)A^k}{k!} \\
&= \sum_{k=0}^{\infty}\frac{d}{dt}\frac{a^k(t)A^k}{k!} \\
&= \sum_{k=1}^{\infty}\frac{a'(t)a^k-1(t)A^k}{k-1!} \\
&= a'(t)\sum_{j=0}^{\infty}\frac{a^j(t)A^{j+1}}{j!} \\
&= a'(t)Ae^{a(t)A}.
\end{aligned} \tag{6.188}$$

Now let us turn to the more interesting case where the exponent $A(t)$ is an arbitrary function of t. First it is important to know that the usual formula for differentiating exponentials is not correct. If

$$A'(t) = \frac{d}{dt}A(t), \tag{6.189}$$

then

$$\frac{d}{dt}e^{A(t)} \neq A'(t)e^{A(t)}. \tag{6.190}$$

This can be easily be shown using 2 by 2 matrices. We now describe the correct result.

Theorem:

$$\begin{aligned}
\frac{d}{dt}e^{A(t)} &= \int_0^1 e^{\tau A(t)}A'(t)e^{-\tau A(t)}d\tau\, e^{A(t)} \\
&= e^{A(t)}\int_0^1 e^{-\tau A(t)}A'(t)e^{\tau A(t)}d\tau.
\end{aligned} \tag{6.191}$$

Proof. The exponential is defined by a power series:

$$e^{A(t)} = \sum_{n=0}^{\infty}\frac{A^n(t)}{n!}, \tag{6.192}$$

and thus

$$\frac{d}{dt}e^{A(t)} = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{m=0}^{n-1} A^m(t)\, A'(t)\, A^{n-m-1}(t)\,. \tag{6.193}$$

Set $A = A(t)$, interchange the order of summation, and do a little algebra to obtain

$$\begin{aligned} \frac{d}{dt}e^{A} &= \sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty} \frac{1}{n!} A^m\, A'\, A^{n-m-1} \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{(m+k+1)!} A^m\, A'\, A^k \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{m!k!}{(m+k+1)!} \frac{A^m}{m!} A' \frac{A^k}{k!}\,. \end{aligned} \tag{6.194}$$

However,

$$\frac{m!k!}{(m+k+1)!} = \int_0^1 \tau^m (1-\tau)^k \, d\tau\,. \tag{6.195}$$

If this is inserted into the previous expression and the series is summed, then

$$\frac{d}{dt}e^{A(t)} = \int_0^1 e^{\tau A(t)}\, A'(t)\, e^{-\tau A(t)}\, d\tau\, e^{A(t)}\,. \tag{6.196}$$

Changing the variable of integration for τ to $1-\tau$ gives

$$\frac{d}{dt}e^{A(t)} = e^{A(t)} \int_0^1 e^{-\tau A(t)}\, A'(t)\, e^{\tau A(t)}\, d\tau\,. \tag{6.197}$$

Another important tool that is useful in its own right, and essential for deriving the exponential identities that are needed, involves the *adjoint action* that we met before in the matrix context. If A is a given noncommuting variable, the adjoint action, Ad_A, generated by A is given by

$$\mathrm{Ad}_A\, B = [A, B]\,. \tag{6.198}$$

We will use a notation similar to the notation used for Lie derivatives:

$$[A, \diamond]\, B = \mathrm{Ad}_A\, B = [A, B]\,. \tag{6.199}$$

Now the formulas (6.169), (6.170), (6.171), (6.172) can immediately converted to identities about the adjoint operator. For example (6.169) gives

$$[A, \diamond]\, B = [A, B] = -[B, A] = -[B, \diamond]\, A\,. \tag{6.200}$$

So the adjoint operator satisfies the following important identities:
Skew-Symmetry of the Adjoint Operator:

$$[A, \diamond]\, B = -[B, \diamond]\, A\,; \tag{6.201}$$

Linearity of the Adjoint Operator:

$$[A,\diamond](bB+cC)=b[A,\diamond]B+c[A,\diamond]C\,; \qquad (6.202)$$

Derivation Property of the Adjoint Operator:

$$[A,\diamond](BC)=([A,\diamond]B)\,C+B\,([A,\diamond]C)\,; \qquad (6.203)$$

Bracket Property of the Adjoint Operator:

$$[A,\diamond][B,C]=[[A,\diamond]B,C]+[B,[A,\diamond]C]\,. \qquad (6.204)$$

The proofs of these identities follow immediately from the definition of the commutator.
As in (3.51) define the powers of the adjoint by

$$[A,\diamond]^0 B=B\,, \quad [A,\diamond]^n=[A,[A,\diamond]^{n-1}]\,,, \qquad (6.205)$$

so that

$$\begin{aligned}
[A,\diamond]^1 B &= [A,B]\,,\\
[A,\diamond]^2 B &= [A,[A,B]]\,,\\
[A,\diamond]^3 B &= [A,[A,[A,B]]]\,,
\end{aligned}$$

and so forth. Then adjoint operator can also be exponentiated:

$$e^{[A,\diamond]}=\sum_{i=0}^{\infty}\frac{1}{k!}[A,\diamond]^k\,. \qquad (6.206)$$

The case where the adjoint operator is multiplied by the scalar t will be used frequently:

$$e^{t[A,\diamond]}=\sum_{i=0}^{\infty}\frac{t^k}{k!}[A,\diamond]^k\,. \qquad (6.207)$$

The exponential of the adjoint operator is another operator:

$$\begin{aligned}
e^{t[A,\diamond]}B &= B+t[A,\diamond]B+\frac{t^2}{2}[A,\diamond]^2B+\frac{t^3}{3!}[A,\diamond]^3B+\cdots\\
&= B+t[A,B]+\frac{t^2}{2}[A,[A,B]]+\frac{t^3}{3!}[A,[A,[A,B]]]+\cdots\,. \qquad (6.208)
\end{aligned}$$

The next result, which we also saw in the matrix context, will be used many times in the following development.

Theorem:

$$e^A\,B\,e^{-A}=e^{[A,\diamond]}B\,. \qquad (6.209)$$

Proof: Define

$$F(t)=e^{tA}\,B\,e^{-tA}\,. \qquad (6.210)$$

Then $F(0) = B$ and

$$F'(t) = \frac{d}{dt}F(t) = A\,e^{tA}\,B\,e^{-tA} - e^{tA}\,B\,e^{-tA}A = [A, F(t)] = [A, \diamond]\,F(t)\,. \tag{6.211}$$

Next, define

$$G(t) = e^{t[A,\diamond]}B\,. \tag{6.212}$$

Then $G(0) = B$ and

$$G'(t) = [A, \diamond]G(t)\,. \tag{6.213}$$

Consequently, $F(t)$ and $G(t)$ satisfy the same initial value problem for a first order ordinary differential equation. Such functions are unique, so $F(t) = G(t)$ and this gives the theorem. This can also be proved by computing the power series expansions of $F(t)$ and $G(t)$. The differential equations gives a recursion for the coefficients in the power series and then a simple induction gives the result.

The exponential of an adjoint operator satisfies a set of identities analogous to the identities satisfied by exponential of a noncommuting variable:

Linearity of the Exponential of an Adjoint Operator:

$$e^{[A,\diamond]}(b\,B + c\,C) = b\,e^{[A,\diamond]}\,B + c\,e^{[A,\diamond]}\,C\,; \tag{6.214}$$

Product Preservation by the Exponential of an Adjoint Operator:

$$e^{[A,\diamond]}\,(B\,C) = (e^{[A,\diamond]}B)(e^{[A,\diamond]}C)\,; \tag{6.215}$$

Commutator Preservation by the Exponential of an Adjoint Operator:

$$e^{[A,\diamond]}[B, C] = [e^{[A,\diamond]}B, e^{[A,\diamond]}C]\,. \tag{6.216}$$

Proofs: The linearity is easy. The product preservation follows from the previous theorem:

$$e^{[A,\diamond]}(B\,C) = e^{A}\,(B\,C)\,e^{-A} = e^{A}\,B\,e^{-A}\,e^{A}\,C\,e^{-A} = (e^{[A,\diamond]}B)\,(e^{[A,\diamond]}C)\,. \tag{6.217}$$

The commutator is defined in terms of products, so the commutator preservation follows immediately from the product preservation.

Near the end of this section we will need a similarity property for exponentials of adjoint operators. To describe this property it is necessary to know how to apply any function to a noncommuting variable. So let

$$f(z) = \sum_{k=0}^{\infty} a_k z^k\,. \tag{6.218}$$

So far only $f(z) = \exp(z)$ has been considered. Later several functions defined in terms of the exponential will be used. In any case, $f(A)$ is given by the formal series

$$f(A) = \sum_{k=0}^{\infty} a_k A^k\,. \tag{6.219}$$

The following property is termed similarity rather than composition because it is also analogous to the notion of similarity transformations for matrices:
Similarity Property of the Exponential of an Adjoint Operator:

$$e^{[A,\diamond]} f(B) = f(e^{[A,\diamond]} B) \tag{6.220}$$

Proof: This is a computation:

$$e^{[A,\diamond]} f(B) = e^{[A,\diamond]} \sum_{k=0}^{\infty} a_k B^k = \sum_{k=0}^{\infty} a_k e^{[A,\diamond]} B^k = \sum_{k=0}^{\infty} a_k (e^{[A,\diamond]} B)^k = f(e^{[A,\diamond]} B) . \tag{6.221}$$

The previous two theorems can be combined to give a nice form for the derivative of the exponential of a general t-dependent quantity.
Corollary:

$$\frac{d}{dt} e^{A(t)} = \frac{e^{[A(t),\diamond]} - 1}{[A(t),\diamond]} A'(t)\, e^{A(t)} , \tag{6.222}$$

$$\frac{d}{dt} e^{A(t)} = e^{A(t)} \frac{1 - e^{-[A(t),\diamond]}}{[A(t),\diamond]} A'(t) . \tag{6.223}$$

Proof: Note that we are applying Formula (6.218) using the analytic function

$$f(z) = \frac{e^{-z} - 1}{z} \tag{6.224}$$

with z replaced by $[A(t),\diamond]$. Formula (6.195) for the derivative is

$$\frac{d}{dt} e^{A(t)} = \int_0^1 e^{\tau A(t)}\, A'(t)\, e^{-\tau A(t)}\, d\tau\, e^{A(t)} . \tag{6.225}$$

Formula (6.209) in the previous theorem allows this to be rewritten as

$$\frac{d}{dt} e^{A(t)} = \int_0^1 e^{\tau [A(t),\diamond]}\, d\tau\, A'(t)\, e^{A(t)} . \tag{6.226}$$

The anti-derivative with respect to τ is simple:

$$\int_0^1 e^{\tau [A(t),\diamond]}\, d\tau = \frac{e^{\tau [A(t),\diamond]} - I}{[A(t),\diamond]} , \tag{6.227}$$

which gives the first part of the theorem, and a similar calculation gives the second part of the theorem.

The results of the next two theorems are two important exponential identities that are frequently called the Baker-Campbell-Hausdorff (BCH) formulas (or identities). The proofs of these identities are base on computing logarithmic derivatives and then using some of the previous results. The logarithmic derivative of

$$F(t) = e^{tA} e^{tB} , \tag{6.228}$$

which is frequently used, has two possible forms:

$$F'(t)F^{-1}(t)\,, \quad F^{-1}(t)F'(t)\,. \tag{6.229}$$

The derivative is given by the product rule;

$$F'(t) = A\,e^{tA}\,e^{tB} + e^{tA}\,B\,e^{tB} = A\,e^{tA}\,e^{tB} + e^{tA}\,B\,e^{-tA}\,e^{tA}\,e^{tB}\,. \tag{6.230}$$

Multiplying on the right by the inverse,

$$F^{-1}(t) = e^{-tB}\,e^{-tA}\,, \tag{6.231}$$

gives

$$F'(t)\,F^{-1}(t) = A + e^{tA}\,B\,e^{-tA} = A\, + e^{t[A,\diamond]}B\,. \tag{6.232}$$

Now we can give the formula for the exponential of a sum:
Theorem:

$$e^{A+B} = e^A e^B e^{C_2} e^{C_3} e^{C_4} \ldots \tag{6.233}$$

and

$$e^{A+B} = \cdots e^{-C_4} e^{C_3} e^{-C_2} e^B e^A\,, \tag{6.234}$$

where C_k is a linear combination of k-fold commutators of A and B. In particular,

$$\begin{aligned} C_2 &= -\frac{1}{2}[A,B]\,, \\ C_3 &= \frac{1}{6}[A,[A,B]] + \frac{1}{3}[B,[A,B]]\,. \end{aligned} \tag{6.235}$$

Proof. We first prove that the second identity follows from the first. Note that $C_k = C_k(A,B)$, and because C_k is a k-fold commutator,

$$C_k(-A,-B) = (-1)^k C_k(A,B). \tag{6.236}$$

Now replacing A and B by $-A$ and $-B$ in the first identity gives

$$e^{-A-B} = e^{-A} e^{-B} e^{C_2} e^{-C_3} e^{C_4} \ldots\,. \tag{6.237}$$

Taking the inverse of both sides of this identity gives the second formula in the theorem.

Next, let

$$F(t) = e^{-tA} e^{-tB} e^{t(A+B)}\,. \tag{6.238}$$

Then

$$\begin{aligned} F'(t)F^{-1}(t) &= -B - e^{-tB}\,A\,e^{tB} + e^{-tB} e^{-tA}(A+B) e^{tA} e^{tB}\,, \\ &= -B - e^{-t[B,\diamond]}A + e^{-t[B,\diamond]} e^{-t[A,\diamond]}(A+B)\,. \end{aligned} \tag{6.239}$$

Next define

$$G(t) = e^{t^2 C_2} e^{t^3 C_3} e^{t^4 C_4} \ldots\,. \tag{6.240}$$

Then

$$G'(t)G^{-1}(t) = 2\,t\,C_2 + 3\,t^2\,e^{t^2[C_2,\circ]}C_3 + 4\,t^3\,e^{t^2[C_2,\circ]}e^{t^3[C_3,\circ]}C_4 + \cdots . \tag{6.241}$$

Expanding the exponentials of the adjoint operators yields:

$$\begin{aligned} F'(t)F^{-1}(t) &= -B - A + t\,[B,A] - \frac{t^2}{2}\,[B,[B,A]] + \cdots \\ &\quad +(A+B) - t\,[A,B] + \frac{t^2}{2}\,[A,[A,B]] + \cdots \\ &\quad -t\,[B,A] + t^2\,[B,[A,B]] + \cdots \\ &\quad +\frac{t^2}{2}\,[B,[B,A]] + \cdots \\ &= -t\,[A,B] + \frac{t^2}{2}\,([A,[A,B]] + 2\,[B,[A,B]]) + \cdots . \\ G'(t)G^{-1}(t) &= 2\,t\,C_2 + 3\,t^2\,C_3 + \cdots . \end{aligned} \tag{6.242}$$

Collecting coefficients of powers of t gives the formulas for C_2 and C_3. Any number of the C_k may be computed by carrying the calculations though the k-th power of t. An inspection of the computation shows that the coefficients of k-th power of t must involve only k-fold commutators of A and B.

This theorem shows how to break up an exponential of a sum into a product of exponentials. The next theorem is a converse; it shows how to combine a product of exponentials into an exponential of a sum.

Theorem:

$$e^A e^B = e^{A+B+D_2+D_3+\cdots}\,, \tag{6.243}$$

where D_k is a linear combination of k-fold commutators of A and B. In particular,

$$\begin{aligned} D_2 &= \frac{1}{2}[A,B]\,, \\ D_3 &= \frac{1}{12}\,[A,[A,B]] - \frac{1}{12}\,[B,[A,B]]\,. \end{aligned} \tag{6.244}$$

Proof: Set

$$F(t) = e^{tA}e^{tB}\,, \tag{6.245}$$

$$D(t) = t(A+B) + t^2 D_2 + t^3 D_3 + \cdots\,, \tag{6.246}$$

and

$$G(t) = e^{D(t)}\,. \tag{6.247}$$

Then

$$\begin{aligned} F'(t)F^{-1}(t) &= A + e^{t[A,\circ]}B \\ &= A + B + t[A,B] + \frac{t^2}{2}[A,[A,B]] + \cdots \end{aligned} \tag{6.248}$$

and

$$
\begin{aligned}
G'(t)G^{-1}(t) &= \int_0^1 e^{\tau[D,\diamond]} D' \, d\tau \\
&= \int_0^1 (D' + \tau[D, D'] + \cdots) \, d\tau \\
&= D' + \frac{1}{2}[D, D'] + \cdots .
\end{aligned} \tag{6.249}
$$

Now

$$
D' = (A + B) + 2\,t\,D_2 + 3\,t^2\,D_3 + \cdots , \tag{6.250}
$$

and

$$
[D, D'] = t^2\,[A + B, D_2] + \cdots . \tag{6.251}
$$

Thus

$$
G'(t)G^{-1}(t) = (A + B) + 2\,t\,D_2 + t^2\,(3\,D_3 + \frac{1}{2}\,[A + B, D_2]) + \cdots . \tag{6.252}
$$

Comparing coefficients of powers of t gives the values gives the formulas for D_2 and D_3.

The sum in the exponent in the previous theorem can be rearranged, and then some of the terms can be summed in closed form.

Theorem:

$$
e^A e^B = e^{C(A,B)} \tag{6.253}
$$

where

$$
C(A, t\,B) = \sum_{k=0}^{\infty} t^k\,C_k(A, B) . \tag{6.254}
$$

Note that C_k contains all terms that contain exactly k factors equal to B. In particular,

$$
\begin{aligned}
C_0(A, B) &= A , \\
C_1(A, B) &= \frac{[A, \diamond]}{1 - e^{-[A,\diamond]}} B \\
&= B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] + \cdots , \\
C_2(A, B) &= -\frac{1}{2} \frac{[A, \diamond]}{1 - e^{-[A,\diamond]}} \int_0^1 [S_1(\tau), S_2(\tau)] \, d\tau \\
&= -\frac{1}{12}\,[B, [A, B]] + \cdots .
\end{aligned} \tag{6.255}
$$

where

$$
\begin{aligned}
S_1(\tau) &= \frac{1 - e^{-\tau[A,\diamond]}}{[A, \diamond]} C_1(A, B) , \\
S_2(\tau) &= e^{-\tau[A,\diamond]}\,C_1(A, B) .
\end{aligned} \tag{6.256}
$$

Proof: It is clear $C(A,B)$ can be rearranged into the given form. Setting $t=0$ gives $C_0(A,B)=A$. Next, set

$$F(t)=e^{A}e^{tB}\,,\quad G(t)=e^{C(A,tB)}\,. \tag{6.257}$$

Then

$$F^{-1}(t)F'(t)=B\,, \tag{6.258}$$

and

$$G^{-1}(t)G'(t)=\int_0^1 e^{-\tau C}\,C'\,e^{\tau C}\,d\tau=\frac{1-e^{-[C,\diamond]}}{[C,\diamond]}\,C'\,. \tag{6.259}$$

Combining these two equations gives:

$$B=\int_0^1 e^{-\tau C}\,C'\,e^{\tau C}\,d\tau=\frac{1-e^{-[C,\diamond]}}{[C,\diamond]}\,C'\,. \tag{6.260}$$

Setting $t=0$ gives:

$$B=\frac{1-e^{-[A,\diamond]}}{[A,\diamond]}\,C_1(A,B)\,, \tag{6.261}$$

or

$$C_1(A,B)=\frac{[A,\diamond]}{1-e^{-[A,\diamond]}}B\,. \tag{6.262}$$

Recall that if

$$f(z)=\frac{z}{1-e^{-z}}=1+\frac{z}{2}+\frac{z^2}{12}-\frac{z^4}{720}+\cdots\,, \tag{6.263}$$

then

$$f([A,\diamond])=\frac{[A,\diamond]}{1-e^{-[A,\diamond]}}=1+\frac{[A,\diamond]}{2}+\frac{[A,\diamond]^2}{12}-\frac{[A,\diamond]^4}{720}+\cdots\,. \tag{6.264}$$

This formula gives the expanded form for C_1.

The computation of C_2 begins by differentiating the equation

$$B=\int_0^1 e^{-\tau C}\,C'\,e^{\tau C}\,d\tau \tag{6.265}$$

and setting

$$\begin{aligned} R_1(\tau) &= \frac{e^{-\tau[C,\diamond]}-1}{[C,\diamond]}\,C'\,,\\ R_2(\tau) &= e^{-\tau C}\,C'\,e^{\tau C}=e^{-\tau[C,\diamond]}\,C'\,, \end{aligned} \tag{6.266}$$

to get

$$0=\int_0^1 \left(e^{-\tau[C,\diamond]}\,C''+R_1(\tau)\,R_2(\tau)-R_2(\tau)\,R_1(\tau)\right)\,d\tau\,. \tag{6.267}$$

Integrating the left-most term gives

$$\frac{1-e^{-[C,\diamond]}}{[C,\diamond]}\,C''=-\int_0^1[R_1(\tau),R_2(\tau)]\,d\tau \tag{6.268}$$

or

$$C'' = -\frac{[C,\diamond]}{1-e^{-[C,\diamond]}}\int_0^1 [R_1(\tau), R_2(\tau)]\,d\tau\,. \tag{6.269}$$

Evaluating the previous formulas at $t = 0$ gives $C'' = 2\,C_2$, and

$$C_2 = -\frac{1}{2}\frac{[A,\diamond]}{1-e^{-[A,\diamond]}}\int_0^1 [S_1(\tau), S_2(\tau)]\,d\tau, \tag{6.270}$$

where

$$\begin{aligned} S_1(\tau) &= \frac{e^{-\tau[A,\diamond]}-1}{[A,\diamond]}C_1\,, \\ S_2(\tau) &= e^{-\tau[A,\diamond]}C_1\,. \end{aligned} \tag{6.271}$$

These quantities are now computed through third order commutators:

$$\begin{aligned} C_1 &= B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] + \cdots, \\ S_1 &= -\tau B + \left(\frac{\tau^2}{2} - \frac{\tau}{2}\right)[A,B] + \left(-\frac{\tau^3}{6} + \frac{\tau^2}{4} - \frac{\tau}{12}\right)[A,[A,B]] + \cdots, \\ S_2 &= B + \left(-\tau + \frac{1}{2}\right)[A,B] + \left(\frac{\tau^2}{2} - \frac{\tau}{2} + \frac{1}{12}\right)[A,[A,B]] + \cdots, \\ [S_1,S_2] &= \frac{1}{6}[B,[A,B]] + \cdots, \\ \int_0^1 [S_1,S_2]\,d\tau &= -\frac{1}{12}[B,[A,B]] + \cdots. \end{aligned} \tag{6.272}$$

We do not see any method of writing the next term in a simple form.

References

[1] C.E. Baum and H.N. Kritikos, *Electromagnetic Symmetry* (in the Electromagnetics Library series), Taylor and Francis, 1995.

[2] C.E. Baum, *Coupled Transmission-Line Model of Periodic Array of Wave Launchers*, Sensor and Simulation Notes 313, Air Force Weapons Laboratory, December 1988.

[3] C.E. Baum, *Canonical Examples for High-Frequency Propagation on Unit Cell of Wave-Launcher Array*, Sensor and Simulation Notes 317, Air Force Weapons Laboratory, April 1989.

[4] C.E. Baum, *High-Frequency Propagation on Nonuniform Multiconductor Transmission Lines in Uniform Media*, Interaction Notes 463, Air Force Weapons Laboratory, March 1988.

[5] C.E. Baum, *Nonuniform Multiconductor Transmission Lines*, Interaction Notes 516, Air Force Phillips Laboratory, February 1986.

[6] J.D. Dollard and C.N. Friedman, *Product Integration*, Encyclopedia of Mathematics and Its Applications, (F.E. Browder, ed.), Addison-Wesley Publishing Company, 1979.

[7] F.R. Gantmacher, *The Theory of Matrices*, Chelsea, 1959.

[8] D.V. Giri, *A Family of Canonical Examples for High Frequency Propagation on Unit Cell of Wave-Launcher Array*, Sensor and Simulation Notes 318, Air Force Weapons Laboratory, June 1989.

[9] L. Hlavaty, B. Wolf and S. Steinberg, *Riccati Equations and Lie Series*, Journal of Mathematical Analysis and Applications, **104** (1984), 246-263.

[10] S. Steinberg, *Lie Series, Lie Transformations and Their Applications*, Lie Methods in Optics, (J. Sáchez Mondragón and K. B. Wolf, eds.), Lecture Notes in Physics, Springer Verlag, 1985, 45-103.

[11] S. Steinberg, *Factored Product Expansions of Solutions of Nonlinear Differential Equations*, SIAM journal on Mathematical Analysis, **15** (1984), 108-115.

[12] S. Steinberg, *Lie Series and Nonlinear Ordinary Equations*, Journal of Mathematical Analysis and Applications, **101** (1984), 39-63.

[13] S. Steinberg, *Applications of the Lie Algebraic Formulas of Baker, Campbell, Hausdorff and Zassenhaus to the Explicit Solutions of Partial Differential Equations*, Journal of Differential Equations, **26** (1977), 404-434.

[14] B.G. Wybourne, *Classical Groups for Physicists* Wiley, 1974.