FIRST ORDER LINEAR ORDINARY DIFFERENTIAL EQUATIONS IN ASSOCIATIVE ALGEBRAS

GORDON ERLEBACHER & GARRRET E. SOBCZYK

ABSTRACT. In this paper, we study the linear differential equation

$$\frac{dx}{dt} = \sum_{i=1}^{n} a_i(t) x b_i(t) + f(t)$$

in an associative but non-commutative algebra $\mathcal{A}$, where the $b_i(t)$ form a set of commuting $\mathcal{A}$-valued functions expressed in a time-independent spectral basis consisting of mutually annihilating idempotents and nilpotents. Explicit new closed solutions are derived, and examples are presented to illustrate the theory.

1. INTRODUCTION

One of the basic differential equations has the form

$$\frac{dx}{dt} = a(t)x + f(t), \tag{1.1}$$

which is linear and of first order. A method of solution is to write the fundamental solution (also called integrating factor) as $\mu_a(t) = e^{\int_0^t a(t')dt'}$ so that the equation can be rewritten equivalently as

$$\frac{d(\mu_a^{-1}x)}{dt} = \mu_a^{-1}f(t),$$

with $\mu_a(0) = 1$. For $x_0 = x(0)$, this method yields the unique solution

$$x(t) = \mu_a(t)x_0 + \mu_a(t)\int_0^t \mu_a^{-1}(s)f(s)ds. \tag{1.2}$$

The main objective of this paper is to study the differential equation

$$\frac{dx}{dt} = \sum_{i=1}^{n} a_i(t) x b_i(t) + f(t) \tag{1.3}$$

with initial condition $x(0) = x_0$ in the general setting of an associative but non-commutative algebra $\mathcal{A}$. Here the $b_i(t)$ form a set of commuting $\mathcal{A}$-valued functions

expressed in a time-independent spectral basis consisting of mutually annihilating idempotents and nilpotents. Thus, our algebra $\mathcal{A}$ could be the matrix algebra $\mathcal{M}_n$ of real or complex $n \times n$ matrices, or a Clifford geometric algebra such as the quaternion algebra H.

Whereas some of the special cases of this differential equation are well-known, for example, for $\mathcal{A} \equiv \mathcal{M}_n$, (when $b_1 = \mathbf{1}$, $b_i = 0$, $i > 1$ and $a = \alpha(t)A$ for a constant matrix A, [6, p.189,193]), we believe that the more general forms of this equation are new and have potentially many applications. Different forms of (1.3) are used in control theory for motion planning for autonomous vehicles [11], in problems of vibration, in column structures under periodic axial loading [12] or the equation of motion for the classical Euler top in n-dimensional space [8]. The generalization and extension of our methods into other areas also appears promising. Possible areas where our methods might apply include the calculation of dichotomies for impulsive equations [9], and in generalizations of the noncommutative operational calculus developed in [3].

To motivate the utility of our approach, we consider the solution of the 2×2 matrix differential equation

$$\frac{dX}{dt} = AXB \tag{1.4}$$

with

$$A = \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

and $\lambda_1 \neq \lambda_2$. For reference, A has eigenvalues 1 and 2.

A standard approach to solving (1.4) is to first multiply it out,

$$\begin{aligned}
\frac{d}{dt} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} &= A \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \\
&= \begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} \\ \lambda_1 x_{21} & \lambda_2 x_{22} \end{pmatrix} \\
&= \begin{pmatrix} 4\lambda_1 x_{11} - 2\lambda_1 x_{21} & 4\lambda_2 x_{12} - 2\lambda_2 x_{22} \\ 3\lambda_1 x_{11} - 1\lambda_1 x_{21} & 3\lambda_2 x_{12} - 1\lambda_2 x_{22} \end{pmatrix}.
\end{aligned}$$

It is not possible to transform this result into an equivalent 2×2 linear equation of the form

$$\frac{dX}{dt} = CX$$

where $C \in \mathcal{M}_2$. On the other hand, equating each matrix element separately, we have four equations in four unknowns. We find that

$$\frac{d}{dt} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 4\lambda_1 x_{11} - 2\lambda_1 x_{21} \\ 4\lambda_2 x_{12} - 2\lambda_2 x_{22} \\ 3\lambda_1 x_{11} - 1\lambda_1 x_{21} \\ 3\lambda_2 x_{12} - 1\lambda_2 x_{22} \end{pmatrix},$$

or in matrix/vector form,

$$\frac{d}{dt} \vec{X} = C\vec{X}$$

where

$$C = \begin{pmatrix} 4\lambda_1 & 0 & -2\lambda_1 & 0 \\ 0 & 4\lambda_2 & 0 & -2\lambda_2 \\ 3\lambda_1 & 0 & -\lambda_1 & 0 \\ 0 & 3\lambda_2 & 0 & \lambda_2 \end{pmatrix} = B^T \otimes A. \tag{1.5}$$

and $\vec{X}$ is a column vector formed from the four elements of X. We have used $\otimes$ to denote the Kronecker tensor product [6, p.256]. See also [1, p. 123], [2], and [4, p.245].

The solution to this equation is well-known,

$$\vec{X} = e^{Ct} \vec{X(0)} \tag{1.6}$$

Although we have an explicit solution, it came at the price of transforming a 2×2 system into a 4×1 system of equations. Although this has not changed the number of variables, it has had the unfortunate consequence of taking an intrinsic object X and separating out its individual components. It would be better to find an expression for the 2×2 matrix X directly as a function of the intrinsic (or invariant) algebraic properties of the matrices A and B, rather than on representations that depend on the choice of a coordinate system used to construct A and B. Furthermore, when elements of more general algebras are used *in lieu* of matrices, working with the equivalent matrix representation of these elements (when they exist) often results in a loss of the geometric and algebraic information used to construct the original differential equation.

Based on the techniques described in this paper, we can immediately write down the solution

$$X(t) = \left(e^{\lambda_1 t} p_{A,1} + e^{2\lambda_1 t} p_{A,2}\right) X(0) p_{B,1} + \left(e^{\lambda_2 t} p_{A,1} + e^{2\lambda_2 t} p_{A,2}\right) X(0) p_{B,2}, \tag{1.7}$$

where we have used the *spectral decompositions*

$$A = p_{A,1} + 2p_{A,2}, \quad B = \lambda_1 p_{B,1} + \lambda_2 p_{B,2}$$

of the matrices A and B in terms of their spectral bases

$$p_{A,1} = 2I_2 - A, \quad p_{A,2} = A - I_2$$

and

$$p_{B,1} = \frac{B - \lambda_2 I_2}{\lambda_1 - \lambda_2}, \quad p_{B,2} = \frac{B - \lambda_1 I_2}{\lambda_2 - \lambda_1},$$

and where I_m is the $m \times m$ identity matrix. Although the solutions (1.6) and (1.7) are equivalent, the latter expression shows the dependence of the solution on intrinsic properties of the matrices A and B. In many cases, matrices that appear in differential equations have geometric or physical significance; therefore, solutions can be more easily interpreted if expressed in terms of these matrices and their properties.

The solutions we seek are expressed by exploiting certain algebraic properties of invariants constructed from the equation elements. Well-known examples of invariants of a matrix are its eigenvalues, its trace, and its determinant. In many cases, matrix elements depend on the basis used to define it. A rotation of the basis functions changes the elements of the matrix, but not the invariants. If the solution to a general linear differential equation is basis-independent, it becomes easier to analyze. For example, a discussion of stability involves eigenvalues. Rather than

study the stability of the 4×4 matrix C, it is more sensible to relate the stability directly to the eigenvalues of the 2×2 matrices A and B.

2. Spectral basis

A *spectral basis* in the associative algebra $\mathcal{A}$ is an ordered set

$$S = \begin{pmatrix} p_1 & q_1 & \ldots & q_1^{m_1-1} & \ldots & p_r & q_r & \ldots & q_r^{m_r-1} \end{pmatrix} \tag{2.1}$$

of idempotents and nilpotents of $\mathcal{A}$ that generate a commutative subalgebra $\mathcal{S}$ of $\mathcal{A}$ and satisfy the following *basic rules* [10]

- $p_1 + \cdots + p_r = 1$,
- $p_j p_i = p_i p_j = \delta_{ij} p_i$ for $i, j = 1, \ldots, r$,
- $q_i^{m_i} = 0$ but $q_i^{m_i-1} \neq 0$,
- $q_j p_i = p_i q_j = \delta_{ij} q_i$.

The spectral basis S is made up of the ordered *sub-blocks*

$$S_i = \begin{pmatrix} p_i & q_i & \ldots & q_i^{m_i-1} \end{pmatrix}$$

so that

$$S = \begin{pmatrix} S_1 & \ldots & S_r \end{pmatrix}$$

In writing the last equality, we are expressing the spectral basis in terms of the ordered sub-blocks S_i. We stress that the sets are *ordered*, because we will be treating them as row matrices of polynomials on which matrix operations are defined.

Note that

$$\operatorname{span}(S) = \operatorname{span}(S_1) \oplus \cdots \oplus \operatorname{span}(S_r)$$

where $\operatorname{span}(S_i)$ is the subspace of $\mathcal{A}$ spanned by the elements of S_i. The *dimension* of the spectral basis S is $m = \sum_{i=1}^{r} m_i$, and determines a commutative subalgebra $\mathcal{S}$ of $\mathcal{A}$. Of course, any element $b \in \mathcal{S}$ is a linear combination of the elements in S,

$$b = \sum_{i=1}^{r} \sum_{j=0}^{m_j-1} \beta_{i,j} q_i^j, \quad \beta_{i,j} \in \mathbb{C} \tag{2.2}$$

where we have adopted the convention that $q_i^0 = p_i$. In terms of the block decomposition of S,

$$bS^T = \begin{pmatrix} b_1 & \ldots & b_r \end{pmatrix} S^T$$

where $b_i = p_i b = \begin{pmatrix} \beta_{i,0} & \ldots & \beta_{i,m_i-1} \end{pmatrix} S_i^T$. If a spectral algebra S has only one block $S = S_1 = \begin{pmatrix} 1 & q & \ldots & q^{m-1} \end{pmatrix}$, we say that $\mathcal{S}$ is a *simple* spectral algebra. A simple element of $\mathcal{A}$ has a simple spectral basis.

To take advantage of the direct sum structure, we note that

$$bS_i^T = B_i S_i^T = (\beta_{i,0} I_{m_i} + N_i) S_i^T \tag{2.3}$$

where the matrix B_i formed from scalar components of b is defined by the upper triangular Toeplitz matrix

$$B_i = \begin{pmatrix} \beta_{i,0} & \beta_{i,1} & \ldots & \beta_{i,m_i-1} \\ 0 & \beta_{i,0} & \ldots & \beta_{i,m_i-2} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \beta_{i,0} \end{pmatrix}, \tag{2.4}$$

and where N_i is a strictly upper triangular Toeplitz matrix. We can now write

$$bS^T = \begin{pmatrix} B_1 & \ldots & B_r \end{pmatrix} S = \sum_{i=1}^{r} (\beta_{i,0} I_{m_i} + N_i) S_i^T. \tag{2.5}$$

Pre-multiplying bS^T by $E_m = \begin{pmatrix} E_{m_1} & \ldots & E_{m_r} \end{pmatrix}$ recovers b:

$$b = E_m b S^T = \sum_{i=1}^{r} E_{m_i} B_i S_i^T \tag{2.6}$$

where $m = \sum_{i=1}^{r} m_i$, and $E_{m_i} = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix}_{m_i}$.

Using (2.5), any analytic function $f(b)$ is computable in a finite number of operations (addition and multiplication in $\mathcal{A}$). We find that

$$f(b)S^T = \begin{pmatrix} f(B_1) & \ldots & f(B_r) \end{pmatrix} S^T = \sum_{i=1}^{r} f(\beta_{i,0} I_{m_i} + N_i) S_i^T. \tag{2.7}$$

With $N_i^0 = I_{m_i}$, $f(b)$ has the Taylor series-like expansion

$$f(\beta_{i,0} I_{m_i} + N_i) = \sum_{j=0}^{m_i-1} f^{(j)}(\beta_{i,0}) N_i^j \tag{2.8}$$

around $\beta_{i,0}$, with *normalized derivatives*

$$f^{(j)}(\beta_{i,0}) = \frac{1}{j!} \frac{d^j}{dx^j} f(x)|_{x=\beta_{i,0}} \tag{2.9}$$

for $i = 1, \ldots, r$ and $j = 0, \ldots, m_i - 1$. We can then substitute (2.8) and (2.9) into (2.7) to get

$$f(b)S^T = \sum_{i=1}^{r} \sum_{j=0}^{m_i-1} f^{(j)}(\beta_{i,0}) N_i^j S_i^T. \tag{2.10}$$

Multiplying both sides of this equation on the left by the contraction operator E_m, we find

$$f(b) = \sum_{i=1}^{r} \sum_{j=0}^{m_i-1} f^{(j)}(\beta_{i,0}) E_{m_i} N_i^j S_i^T. \tag{2.11}$$

For example, consider the expansion of e^b in the spectral basis S. Applying (2.11), where $f(x) = e^x$, we immediately conclude that

$$e^b = \sum_{i=1}^{r} e^{\beta_{i,0}} \sum_{j=0}^{m_i-1} \frac{1}{j!} E_{m_i} N_i^j S_i^T. \tag{2.12}$$

We will have use for special cases of this formula in Section 5.

2.1. **Minimal Polynomial.** Every element $b \in \mathcal{A}$ generates a commutative subalgebra $\mathcal{A}_b$ of $\mathcal{A}$ consisting of linear combinations of powers of the element b. We can take as the *standard basis* of this algebra the ordered set $D = \begin{pmatrix} 1 & b & \ldots & b^{m-1} \end{pmatrix}$; thus any element $c \in \mathcal{A}_b$ can be expressed in the form

$$c = \sum_{i=0}^{m-1} \beta_i b^i = D\{c\}_D^T \tag{2.13}$$

where $\{c\}_D = \begin{pmatrix} \beta_0 & \cdots & \beta_{m-1} \end{pmatrix}_D$ are the complex scalar components of the element $c \in \mathcal{A}_b$ with respect to the basis D and $\beta_i \in \mathbb{C}$. When the context is unambiguous, we will drop the subscript indicating the basis.

The *minimal polynomial* of the element b is the unique monic polynomial

$$\psi(x) = \prod_{i=1}^{r} (x - x_i)^{m_i} \tag{2.14}$$

of minimal degree with the property that $\psi(b) = 0$. We will always assume that the distinct roots x_i of the minimal polynomial are complex numbers, which guarantees that the primitive factors of the minimal polynomial will all be of the form $(x - x_i)^{m_i}$. The minimal polynomial uniquely determines the *spectral decomposition* of the element b,

$$b = \sum_{i=1}^{r} (x_i + q_i) p_i$$

for distinct $x_i \in \mathbb{C}$, where $p_i = p_i(b)$ and $q_i = q_i(b)$, $i = 1, \ldots, r$, are polynomials in b of degree less than m, and where the algebra $\mathcal{A}_b = \mathcal{S}[10]$.

3. Kronecker Tensor Products

In this section, we review some of the properties of tensor products in preparation for some of the manipulations that will follow. Following [6, p.256], and [1, p.123], [4, pp. 239-297], the tensor product of two matrices $A \in \mathcal{M}_{m,n}(\mathcal{A})$ and $B \in \mathcal{M}_{p,q}(\mathcal{A})$ satisfy

-
$$(A \otimes B) = (a_{ij}) \otimes (b_{kl}) := ((a_{ij})b_{kl}) \tag{3.1}$$

  (Note that some authors multiply matrix B by the elements a_{ij} of A.)

-
$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \tag{3.2}$$

  assuming that the matrix multiplications AB and CD are well defined.

-
$$A \otimes B = (A \otimes I_B)(I_A \otimes B) \tag{3.3}$$

  where A and B are square matrices, and I_A, I_B are the identity matrices with the same dimensions as A and B respectively.
- Multiplication by an element $x \in \mathcal{A}$:
$$x(A \otimes B) = (xA) \otimes B$$
$$(A \otimes B)x = A \otimes (Bx)$$

- Associativity:
$$(A \otimes B) \otimes C = A \otimes (B \otimes C) \tag{3.4}$$

- Transposition:
$$(A \otimes B)^T = A^T \otimes B^T \tag{3.5}$$

Both of these matrices and their elements are non-commutative in general.

4. Linear Superposition

When working with linear operators, the principle of linear superposition is a powerful tool for problem simplification. We apply this principle to solve the linear differential equation

$$\frac{dx}{dt} = \sum_{i=1}^{n} a_i(t)xb_i(t) + f(t) \tag{4.1}$$

where the time-dependent b_i's form a commutative set that can be expanded in terms of a set of time-independent commutative spectral basis elements,

$$b_i(t) = S_i\{b_i\}^T = \sum_{j=1}^{r_i} S_{i,j}p_{i,j}\{b_{i,j}\}^T \tag{4.2}$$

for $b_{i,j} = p_{i,j}b_i \in \mathcal{A}$,

$$S_{i,j} = \begin{pmatrix} p_{i,j} & q_{i,j} & \dots & q_{i,j}^{m_{i,j}-1} \end{pmatrix},$$

and the time-dependent coefficients

$$\{b_{i,j}\} = \begin{pmatrix} \beta_{i,j,0}(t) & \dots & \beta_{i,j,m_j-1}(t) \end{pmatrix},$$

where $\beta_{i,j,k} \in \mathbb{C}$ for $i = 1, \dots, n$ and $j = 1, \dots, r_i$. The commutativity of b_i and b_k implies the commutativity of the subspaces S_i and S_k. Furthermore, $p_{i,j}p_{i,k} = p_{i,j}p_{i,k}\delta_{j,k}$.

To simplify the problem, we seek to replace (4.1) by the new equation

$$\frac{dy}{dt} = \sum_{i=1}^{n} c_i(t)yd_i(t) + g(t) \tag{4.3}$$

where the d_i are now simple elements of $\mathcal{A}$, and $c_i, g \in \mathcal{A}$. Using (4.2) and the commutativity of matrices of scalars with elements of $\mathcal{A}$, it follows that

$$b_i(t)p_{(j)} = \sum_{j=1}^{r_i} S_{i,j}p_{i,j}\{b_{i,j}\}^T p_{(j)} = p_{(j)}S_{i,j_i}\{b_{i,j_i}\}^T, \tag{4.4}$$

where $p_{(j)} = p_{1,j_1} \dots p_{n,j_n}$. The second equality is based on $p_{(j)}p_{i,j} = p_{(j)}\delta_{j,j_i}$.

Next, multiplying (4.1) to the right by $p_{(j)}$, and using (2.13), (4.2) and (4.4) gives

$$\frac{dxp_{(j)}}{dt} = \sum_{i=1}^{n} a_i(t)xp_{(j)}S_{i,j_i}\{b_{i,j_i}\}^T + fp_{(j)}$$

$$= \sum_{i=1}^{n} a_i(t)xp_{(j)}b_{i,j_i} + fp_{(j)} \tag{4.5}$$

Equation (4.5) is of type (4.3) with $y = xp_{(j)}$, $c_i = a_i$, $d_i = b_{i,j_i}$, and $g(t) = f(t)p_{(j)}$. Note that $a_i(t)$ need not have a simple spectral basis and remains a general time-dependent element of $\mathcal{A}$. $x_{(j)} = xp_{(j)}$ represents the solution to (4.1) projected onto the subspace spanned by $S_{(j)} = \otimes_{i=1}^{n} \text{span}(S_{i,j_i})$. The full solution is constructed by linear superposition of $r_1 r_2 \dots r_n$ elemental solutions $xp_{(j)}$:

$$x(t) = \sum_{(j)} xp_{(j)} = \sum_{j_1=1}^{r_1} \cdots \sum_{j_n=1}^{r_n} xp_{1,j_1} \dots p_{n,j_n} \tag{4.6}$$

Several comments are in order. Since each b_i is expanded in its own fixed basis, one cannot assume that $S_{i,k}$ and $S_{j,l}$ are orthogonal to each other when $i \neq j$. When orthogonality $S_{i,j}$ with respect to i is not satisfied, the solutions $xp_{(j)}$ are not mutually orthogonal, nor are they necessarily linearly independent.

Additional simplifications are possible if the a_i are also expanded in a time-independent spectral basis S_{ai}, independent from the spectral basis of the b_i. Let

$$a_i = S_{ai}\{a_i\}^T = \sum_{j=i}^{s_i} \{a_{i,j}\} p_{ai,j} S_{a_{i,j}}^T$$

in correspondence with (4.2), $S_{ai,j} = \begin{pmatrix} q_{ai,j}^0 & \cdots & q_{ai,j}^{n_i-1} \end{pmatrix}$, and
$p_{a(k)} = p_{a1,k_1} \ldots p_{an,k_n}$, along with the convention $q_{ai,j}^0 = p_{ai,j}$. Pre-multiply (4.5) by $p_{a(k)}$,

$$\frac{dp_{a(k)}xp_{(j)}}{dt} = \sum_{i=1}^{n} a_{i,k_i} p_{a(k)} xp_{(j)} b_{i,j_i} + p_{a(k)} fp_{(j)} \tag{4.7}$$

which is of the form (4.3) with $y = p_{a(k)}xp_{(j)}$, and where both $d_i = a_{i,k_i}$ and $c_i = b_{i,j_i}$ are simple elements of $\mathcal{A}$. The full solution to (4.1) is simply a linear superposition of elemental solutions $p_{a(k)}xp_{(j)}$:

$$x = \sum_{(j)} \sum_{(k)} p_{a(k)}xp_{(j)}$$

In the sections that follow, we consider only the solution of the differential equation (4.3).

5. Fundamental solution

The solution $x = \mu_a(t)x_0$ to the linear homogeneous equation

$$\frac{dx}{dt} = a(t)x$$

with initial conditions $x_0 = x(0)$, $a(t)$ and $x \in \mathcal{A}$, is the basic building block for all solutions to the equations considered in this paper. The *fundamental solution* $\mu_a(t)$ is formally assumed to satisfy the property

$$\frac{d\mu_a}{dt} = a(t)\mu_a \tag{5.1}$$

with the initial condition $\mu_a(0) = 1$.

The fundamental solution has been discussed by many authors in different contexts. In [12], the fundamental solution is studied under the condition that $a(t)$ is periodic. In [7], the fundamental solution determines the controllability and observability of the system. In [5], the author studies the fundamental solution in terms of what he calls "expansions that grow on trees." We show in this paper how, under certain conditions and assumptions, the fundamental solution leads to a family of solutions that can be expressed in closed form.

By convention, the identity element e of $\mathcal{A}$ is represented by 1, and we assume that the left and right inverses of $a \in \mathcal{A}$ (if they exist) are equal. Under these conditions, $a^{-1}a = aa^{-1} = e = 1$. A subscript is attached to the generalized integrating factor μ_a to reflect the $\mathcal{A}$-valued function $a(t)$ that it is associated with.

The equation satisfied by an assumed inverse μ_a^{-1} is determined by differentiating $\mu_a(t)\mu^{-1}(t) = 1$ with respect to t. One finds

$$\frac{d\mu_a^{-1}}{dt} = -\mu_a^{-1}a = (-\mu_a^{-1}a\mu_a)\mu_a^{-1}. \tag{5.2}$$

Using the definition of μ_a, we derive an implicit relation for μ_a^{-1}:

$$\mu_a^{-1}(t) = \mu_{-\mu_a^{-1}a\mu_a}(t).$$

In the event that $\mu_a(t)$ and $a(t)$ commute, μ_a and its inverse are simply related by $\mu_a^{-1}(t) = \mu_{-a}(t)$.

The fundamental solution $\mu_a(t)$ can be found directly through Picard iteration, generated by successive integrations of (5.1) with respect to t. We find

$$\mu_a(t) = 1 + \int_0^t a(s_1)ds_1 + \int_0^t \int_0^{s_1} a(s_1)a(s_2)ds_1ds_2$$
$$+ \int_0^t \int_0^{s_1} \int_0^{s_2} a(s_1)a(s_2)a(s_3)ds_1ds_2ds_3 + \ldots \tag{5.3}$$

expressed as an infinite series of iterated integrals. Similarly, successive integration of (5.2) leads to

$$\mu_a^{-1}(t) = 1 - \int_0^t a(s_1)ds_1 + \int_0^t \int_0^{s_1} a(s_2)a(s_1)ds_1ds_2$$
$$- \int_0^t \int_0^{s_1} \int_0^{s_2} a(s_3)a(s_2)a(s_1)ds_1ds_2ds_3 + \ldots \tag{5.4}$$

When a does not depend on time, a and μ_a commute and μ_a and μ_a^{-1} reduce to the exponential forms

$$\mu_a(t) = \sum_{j=0}^{\infty} \frac{a^j t^j}{j!} \equiv e^{at}, \tag{5.5}$$

$$\mu_a^{-1}(t) = \sum_{j=0}^{\infty} \frac{(-1)^j a^j t^j}{j!} \equiv e^{-at}. \tag{5.6}$$

If $a(t_1)$ and $a(t_2)$ commute for all t_1 and t_2,

$$\mu_a(t) = e^{\int_0^t a(s)ds} \tag{5.7}$$

and $\mu_a(t)a(s) = a(s)\mu_a(t)$. This condition is automatically satisfied for any time-dependent simple element of $\mathcal{A}$ whose spectral basis is independent of time.

We now wish to study the conditions under which $\mu_{a+b}(t)$ can be expressed in terms of μ_a and μ_b, for time dependent $\mathcal{A}$-valued functions $a(t), b(t)$. This is easily accomplished by expressing the solution to

$$\frac{dx}{dt} = (a+b)x \tag{5.8}$$

in two different ways. From the leftmost equality, we obtain

$$x = \mu_{a+b}x_0 \tag{5.9}$$

Alternatively, substituting $x = \mu_a z$ into (5.8), we find, with the help of (5.1) that

$$\frac{dx}{dt} = a\mu_a z + \mu_a \frac{dz}{dt} = (a+b)\mu_a z, \tag{5.10}$$

which simplifies to

$$\mu_a \frac{dz}{dt} = b\mu_a z \,. \tag{5.11}$$

Equation (5.11) has the solution $z = \mu_{\mu_a^{-1} b \mu_a} z_0$ which, in terms of $x = \mu_a z$, is

$$x = \mu_a \mu_{\mu_a^{-1} b \mu_a} x_0 \,. \tag{5.12}$$

Equating the two alternative expressions for $x(t)$ given by (5.9) and (5.12), we conclude, for general $a(t)$ and $b(t) \in \mathcal{A}$, that

$$\mu_{a+b} = \mu_a \mu_{\mu_a^{-1} b \mu_a}. \tag{5.13}$$

Explicit solutions for μ_a and μ_a^{-1} are difficult to derive from (5.3) and (5.4) for two reasons. First, the series has an infinite number of terms. Second, it is not possible, in general, to derive explicit closed formulas for the integrals of general time-dependent functions that take their values in $\mathcal{A}$. However, when $a(t)$ has an expansion in a constant spectral basis $S = \begin{pmatrix} S_1 & \dots & S_r \end{pmatrix}$, then μ_a is easily expressed as a finite series of integrals with integrands in $\mathbb{C}$.

Let

$$a(t) = \sum_{i=1}^{r} \{a_i\}_{S_i} S_i^T p_i = \sum_{i=1}^{r} E_{m_i}(\alpha_{i,0}(t) I_{m_i} + N_i(t)) S_i^T p_i,$$

be the spectral expansion (see (2.11)) of a, where $S_i = \begin{pmatrix} 1 & \dots & q_i^{m_i-1} \end{pmatrix} p_i$. From (2.12) it follows that

$$\mu_a = \sum_{i=1}^{r} e^{\int_0^t \alpha_{i,0}(s)ds} E_{m_i} e^{\int_0^t N_i(s)ds} S_i^T p_i.$$

Since the last integral on the right is the exponential function of a strictly upper triangular Toeplitz matrix, it can be written as the finite sum

$$e^{\int_0^t N_i(s)ds} = \sum_{j=0}^{m_i-1} \frac{1}{j!} \Big[\int_0^t N_i(s)ds \Big]^j.$$

6. SOLUTION TO $dx/dt = a(t)xb(t) + f(t)$

Let $a(t) \in \mathcal{A}$ be an arbitrary time-dependent function of t, and $b(t) \in S$ for the simple spectral basis $S = \begin{pmatrix} 1 & q & \dots & q^{m-1} \end{pmatrix}$. Thus,

$$b(t) = \sum_{j=0}^{m-1} \beta_j(t) q^j = \{b\} S^T \tag{6.1}$$

where $\beta_0(t), \dots, \beta_{m-1}(t)$ are the time-dependent scalar components of b in the spectral basis S. Multiplying (6.1) on the right by S^T, we find that $bS^T = BS^T$, where $B \in \mathcal{M}_m$ is given by (2.4) after removal of the i subscript.

We are concerned with deriving the general solution to the inhomogeneous equation

$$\frac{dx}{dt} = axb + f \tag{6.2}$$

with initial condition $x(0) = x_0$. Multiplying (6.2) on the right by S^T, we find that

$$\frac{dxS^T}{dt} = axbS^T + fS^T = BaxS^T + fS^T.$$

A major simplification occurs when B is decomposed into the sum of a diagonal matrix $\beta_0 I_m$ and a nilpotent matrix N, $B = \beta_0 I_m + N$. One must therefore solve

$$\frac{dxS^T}{dt} = (\mathcal{L} + \mathcal{K})xS^T + fS^T \tag{6.3}$$

where we have defined $\mathcal{L} = \beta_0 a$ and $\mathcal{K} = aN$ in anticipation of the next section.

We first compute the solution $x_H S^T$ to the homogeneous equation (setting $f = 0$ in (6.3)), followed by an application of the method of undetermined constants. From (5.12), the homogeneous solution is

$$x_H S^T = \mu_{\mathcal{L}} \mu_\Gamma x_0 S^T \tag{6.4}$$

where $\Gamma = \mu_{\mathcal{L}}^{-1} \mathcal{K} \mu_{\mathcal{L}} \in \mathcal{M}_m(\mathcal{A})$. The solution cannot be simplified further, because in general, $\mu_{\mathcal{L}} a \neq a \mu_{\mathcal{L}}$.

To apply the method of underdetermined constants, we replace x_0 by an unknown element $z(t) \in \mathcal{A}$ and substitute the modified homogeneous solution $\mu_{\mathcal{L}} \mu_\Gamma z(t) S^T$ back into (6.3), which leads to

$$\frac{dxS^T}{dt} = (\mathcal{L} + \mathcal{K})xS^T + \mu_{\mathcal{L}} \mu_\Gamma \frac{dz}{dt} S^T = (\mathcal{L} + \mathcal{K})xS^T + fS^T \tag{6.5}$$

or

$$\frac{dz}{dt} S^T = \mu_\Gamma^{-1} \mu_{\mathcal{L}}^{-1} f(t) S^T$$

with particular solution

$$zS^T = \int_0^t \mu_\Gamma^{-1}(s) \mu_{\mathcal{L}}^{-1}(s) f(s) S^T ds \tag{6.6}$$

Combining the homogeneous and particular solutions leads to the general solution

$$x(t)S^T = \mu_{\mathcal{L}}(t)\mu_\Gamma(t)x_0 S^T + \Big[\mu_{\mathcal{L}}(t)\mu_\Gamma(t) \int_0^t \mu_\Gamma^{-1}(s)\mu_{\mathcal{L}}^{-1}(s)f(s)ds \Big] S^T \tag{6.7}$$

The solution to (6.2) is found by pre-multiplication of (6.7) by E_m, with the result

$$x(t) = E_m \mu_{\mathcal{L}}(t)\mu_\Gamma(t)x_0 S^T + E_m \Big[\mu_{\mathcal{L}}(t)\mu_\Gamma(t) \int_0^t \mu_\Gamma^{-1}(s)\mu_{\mathcal{L}}^{-1}(s)f(s)ds \Big] S^T \tag{6.8}$$

Our method of solution brings about a major simplification. The scalar time-dependent matrix $N(t) \in \mathcal{M}_n(\mathbb{C})$ is a strictly upper triangular Toeplitz nilpotent matrix, while $\mu_{\mathcal{L}} \in \mathcal{A}$. Therefore, Γ is a strictly upper triangular Toeplitz nilpotent $m \times m$ matrix with elements in $\mathcal{A}$ and the property $\Gamma(t_1)\Gamma(t_2) = \Gamma(t_2)\Gamma(t_1)$ remains true for all t_1 and t_2. As a consequence, μ_Γ is a *finite* sum of m iterated integrals.

7. Solution to $dx/dt = \sum_i a_i x b_i + f$

We now extend the methodology of the previous two sections to solve the linear inhomogeneous equation

$$\frac{dx}{dt} = A(t)xB^T(t) + f(t) \tag{7.1}$$

in $\mathcal{A}$ with $x(0) = x_0$. $A = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix}$ and $B = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix}$ are row vectors of time-dependent elements in $\mathcal{A}$ that satisfy the commutativity relations $b_i(t)b_j(t) = b_j(t)b_i(t)$ for all time. As explained in Sections 2 and 4, we only consider simple spectral bases S_i and use the principle of linear superposition to reconstruct the full

solution. The $b_i(t)$ remain mutually commutative if they are expressed as a time-dependent linear combination of time-invariant commuting simple spectral bases S_i. Therefore, each b_i has an expansion of the form

$$b_i(t) = \sum_{k=0}^{m_i-1} b_{ik}(t) q_i^k \tag{7.2}$$

where q_i is nilpotent with multiplicity index m_i.

Recall from (2.3), that

$$b_i(t) S_i^T = B_i S_i^T = (b_{i,0} I_{m_i} + N_i) S_i^T \tag{7.3}$$

where B_i is an upper triangular Toepliz matrix (see (2.4)). We proceed by defining

$$S = S_1 \otimes \cdots \otimes S_n, \quad \text{and} \quad S^T = S_1^T \otimes \cdots \otimes S_n^T.$$

Note that $b_i S^T = \bar{B}_i S^T$, where we have defined

$$\bar{B}_i = I_{m_1} \otimes \cdots \otimes B_i \otimes \cdots \otimes I_{m_n}.$$

As a general notational device, the bar over a symbol indexed by i represents a tensor product of identity matrices, with the ith matrix replaced by the symbol, itself of dimension $m_i \times m_i$. Therefore, with the help of (3.2), we find that $xB^T S^T = x\bar{B}^T S^T = \bar{B}^T x S^T$, where

$$\bar{B} = \begin{pmatrix} \bar{B}_1 & \dots & \bar{B}_n \end{pmatrix}.$$

Multiplying (7.1) on the right by S^T, we obtain

$$\frac{dx}{dt} S^T = A x B^T S^T + f S^T = A \bar{B}^T x S^T + f S^T. \tag{7.4}$$

Following the strategy of the previous section, slot i in $\bar{B}_i$ is decomposed into the sum of a diagonal and a nilpotent matrix, so that

$$\bar{B}_i = b_{i,0} \bar{I}_{m_i} + \bar{N}_i.$$

The following definitions will make (7.4) formally identical to (6.5) in the previous section:

$$\bar{\mathcal{K}} = \sum_{i=1}^n a_i \bar{N}_i, \tag{7.5}$$

$$\bar{\mathcal{L}} = \sum_{i=1}^n b_{0,i} a_i \bar{I}_{m_i} = I_M \sum_{i=1}^n b_{0,i} a_i = I_M \mathcal{L}$$

where $M = m_1 m_2 \dots m_n$ and $\mathcal{L} = \sum_{i=1}^n b_{0,i} a_i$. With the preceding definitions, (7.4) becomes

$$\frac{dx S^T}{dt} = (\bar{\mathcal{L}} + \bar{\mathcal{K}}) x S^T + f S^T, \tag{7.6}$$

which has the homogeneous solution ($f = 0$)

$$x_H S^T = \mu_{\bar{\mathcal{L}}} \mu_{\bar{\Gamma}} x_0 S^T, \tag{7.7}$$

where $\bar{\Gamma} = \mu_{\bar{\mathcal{L}}}^{-1} \bar{\mathcal{K}} \mu_{\bar{\mathcal{L}}} \in \mathcal{M}_M(\mathcal{A})$. The general solution to (7.6) is immediately found to be

$$x(t) = E_M \mu_{\bar{\mathcal{L}}}(t) \mu_{\bar{\Gamma}}(t) x_0 S^T + E_M \mu_{\bar{\mathcal{L}}}(t) \mu_{\bar{\Gamma}}(t) \int_0^t \mu_{\bar{\Gamma}}^{-1}(s) \mu_{\bar{\mathcal{L}}}^{-1}(s) f(s) S^T ds. \tag{7.8}$$

Recall that Γ is defined by

$$\bar{\Gamma} = \sum_{i=1}^{n} \bar{\Gamma}_i,$$

where $\bar{\Gamma}_i = \mu_{\mathcal{L}}^{-1} a_i \bar{N}_i \mu_{\mathcal{L}} \in \mathcal{M}_M(\mathcal{A})$. Let us now compute an explicit representation for $\mu_{\bar{\Gamma}}$, defined recursively by

$$\mu_{\bar{\Gamma}} = 1 + \int_0^t \bar{\Gamma}(s)\mu_{\bar{\Gamma}}(s)ds = 1 + \sum_{i=1}^{n} \int_0^t \bar{\Gamma}_i(s)\mu_{\bar{\Gamma}}(s)ds$$

At the next iteration level,

$$\mu_{\bar{\Gamma}} = 1 + \sum_{i_1=1}^{n} \int_0^t ds_1 \bar{\Gamma}_{i_1}(s_1)\Big(1 + \sum_{i_2=1}^{n} \int_0^{s_1} ds_2 \bar{\Gamma}_{i_2}(s_2)\mu_{\bar{\Gamma}}(s_2)\Big) \qquad (7.9)$$

$$= 1 + \sum_{i_1=1}^{n} \int_0^t ds_1 \bar{\Gamma}_{i_1}(s_1) + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \int_0^t ds_1 \int_0^{s_1} ds_2\, \bar{\Gamma}_{i_1}(s_1)\bar{\Gamma}_{i_2}(s_2)\mu_{\bar{\Gamma}}(s_2)$$

$$(7.10)$$

The kth iterated integral takes the form

$$\mathcal{I}^{(k)} = \sum_{i_1} \cdots \sum_{i_k} \int_0^t ds_1 \int_0^{s_1} \cdots \int_0^{s_{k-1}} ds_k \prod_{j=1}^{k} \bar{\Gamma}_{i_j}(s_j), \qquad (7.11)$$

so that $\mu_{\bar{\Gamma}} = \sum_{k=0}^{\infty} \mathcal{I}^{(k)}$, where, by convention, $\mathcal{I}^{(0)} = 1$. A similar development yields $\mu_{\bar{\Gamma}}^{-1} = \sum_{k=0}^{\infty}(-1)^k \mathcal{J}^{(k)}$ with $\mathcal{J}^{(0)} = 1$, and

$$\mathcal{J}^{(k)} = \sum_{i_1} \cdots \sum_{i_n} \int_0^t ds_1 \int_0^{s_1} ds_2 \dots \int_0^{s_n} ds_{n-1} \prod_{j=k}^{1} \bar{\Gamma}_{i_j}(s_j),$$

8. Examples

Example 1. We work out in detail the case $n = 2$ and $m_i = 2$ with the elements a_i and b_i functions of time. The equation to solve is

$$\frac{dx}{dt} = a_1(t)xb_1(t) + a_2(t)xb_2(t), \qquad (8.1)$$

where $b_1 = \beta_{1,1,0}(t)p_{1,1} + \beta_{1,1,1}(t)q_{1,1} + \beta_{1,2,0}(t)p_{1,2} + \beta_{1,2,1}(t)q_{1,2}$, $b_2 = \beta_{2,1,0}(t)p_{2,1} + \beta_{2,1,1}(t)q_{2,1} + \beta_{2,2,0}(t)p_{2,2} + \beta_{2,2,1}(t)q_{2,2}$, and $x(0) = x_0$. As explained in Section 4, the principle of linear superposition allows us to solve first

$$\frac{dx}{dt} = a_1 x c_1 + a_2 x c_2$$

where

$$c_1 = \gamma_{1,0} + \gamma_{1,1}q_1, \quad c_2 = \gamma_{2,0} + \gamma_{2,1}q_2, \qquad (8.2)$$

are simple elements of $\mathcal{A}$. Define $\mathcal{L} = a_1\gamma_{1,0} + a_2\gamma_{2,0}$, the direct product

$$S^T = \begin{pmatrix} 1 \\ q_1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ q_2 \end{pmatrix}$$

of spectral bases, and the direct product

$$E_4 = E_2 \otimes E_2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

of row vectors. The solution is (see (7.8), (7.9),(7.11))

$$x(t) = E_4 \mu_{\mathcal{L}} \Big(1 + \sum_{i=1}^{2} \int_0^t ds_1 \, \bar{\Gamma}_i(s_1) + \sum_{i_1=1}^{2} \sum_{i_2=1}^{2} \int_0^t ds_1 \int_0^{s_1} ds_2 \, \bar{\Gamma}_{i_1}(s_1) \bar{\Gamma}_{i_2}(s_2) \Big) x_0 S^T$$

(8.3)

The first term of (8.3) is clearly $T_1 = E_4 \mu_{\mathcal{L}}(t) x_0 S^T = \mu_{\mathcal{L}}(t) x_0$. The second term in (8.3) is

$$T_2 = E_4 \mu_{\mathcal{L}}(t) \sum_{i=1}^{2} \int_0^t ds_1 \, \bar{\Gamma}_i(s_1) x_0 S^T$$

The matrix Γ_i has the form

$$\Gamma_i = \zeta_i \begin{pmatrix} 0 & \zeta_i \\ 0 & 0 \end{pmatrix} = \zeta_i J_2$$

(8.4)

where $\zeta_i = \mu_{\mathcal{L}}^{-1} a_i \gamma_{i,1} \mu_{\mathcal{L}} \in \mathcal{A}$ and J_2 is a nilpotent matrix of degree 2. It is straightforward to derive

$$E_4 \bar{\Gamma}_1 x_0 S^T = \zeta_1 (E_2 J_2 \otimes E_2 J_2)(x_0 S_1^T \otimes S_2^T) = \zeta_1 x_0 q_1$$

Similarly, $E_4 \bar{\Gamma}_1 x_0 S^T = \zeta_2 x_0 q_2$.

With these definitions,

$$T_2 = \mu_{\mathcal{L}}(t) \int_0^t ds_1 \, (\zeta_2(s_1) x_0 q_2 + \zeta_1(s_1) x_0 q_1)$$

The structure of the third term, T_3, of (8.3) is easily determined from (8.4) and the properties of tensor products. Recalling that $\Gamma_i(s_1)\Gamma_i(s_2) = 0$, we find

$$T_3 = \sum_{i_1=1}^{2} \sum_{i_2=1}^{2} \mu_{\mathcal{L}}(t) \int_0^t ds_1 \int_0^{s_1} ds_2 \, E_4 \bar{\Gamma}_{i_1}(s_1) \bar{\Gamma}_{i_2}(s_2) x_0 S^T$$

$$= \mu_{\mathcal{L}}(t) \int_0^t ds_1 \int_0^{s_1} ds_2 \, (\zeta_1(s_1)\zeta_2(s_2) + \zeta_2(s_1)\zeta_1(s_2)) x_0 q_1 q_2 \, .$$

Returning to the original variables, we obtain the solution to our simplified original equation (8.1), with b's defined by (8.2),

$$x(t) = \mu_{\mathcal{L}}(t) x_0 + \mu_{\mathcal{L}}(t) \int_0^t ds_1 \, \big([\mu_{\mathcal{L}}^{-1} a_2 \gamma_{2,1} \mu_{\mathcal{L}}]_{s_1} x_0 q_2 + [\mu_{\mathcal{L}}^{-1} a_1 \gamma_{1,1} \mu_{\mathcal{L}}]_{s_1} x_0 q_1 \big)$$

$$+ \mu_{\mathcal{L}}(t) \int_0^t ds_1 \int_0^{s_1} ds_2 \, \Big([\mu_{\mathcal{L}}^{-1} a_1 \gamma_{1,1} \mu_{\mathcal{L}}]_{s_1} [\mu_{\mathcal{L}}^{-1} a_2 \gamma_{2,1} \mu_{\mathcal{L}}]_{s_2}$$

$$+ [\mu_{\mathcal{L}}^{-1} a_2 \gamma_{2,1} \mu_{\mathcal{L}}]_{s_1} [\mu_{\mathcal{L}}^{-1} a_1 \gamma_{1,1} \mu_{\mathcal{L}}]_{s_2} \Big) x_0 q_1 q_2$$

(8.5)

Expressions within square brackets are evaluated at the value indicated by their subscript.

As explained in Section 4, in order to solve the original equation (8.1), it is first multiplied on the right by $p_{1,i} p_{2,j}$ to get the projected equations

$$\frac{dx_{ij}}{dt} = a_1 x_{ij}(\beta_{1,i,0} + \beta_{1,i,1} q_{1,i}) + a_2 x_{ij}(\beta_{2,j,0} + \beta_{2,j,1} q_{2,j}),$$

(8.6)

where $x_{ij} = xp_{1,i}p_{2,j}$. Each of these equations has a solution given by (8.5). Thus,

$$x_{ij}(t) = \mu_{\mathcal{L}_{ij}}(t)x_0 p_{1,i}p_{2,j} + \mu_{\mathcal{L}_{ij}}(t)\int_0^t ds_1 \left([\mu_{\mathcal{L}_{ij}}^{-1}a_2\beta_{2,j,1}\mu_{\mathcal{L}_{ij}}]_{s_1}x_0 p_{1,i}q_{2,j}\right.$$
$$+ [\mu_{\mathcal{L}_{ij}}^{-1}a_1\beta_{1,i,1}\mu_{\mathcal{L}_{ij}}]_{s_1}x_0 q_{1,i}p_{2,j}\Big)$$
$$+ \mu_{\mathcal{L}_{ij}}(t)\int_0^t ds_1 \int_0^{s_1} ds_2 \left([\mu_{\mathcal{L}_{ij}}^{-1}a_1\beta_{1,i,1}\mu_{\mathcal{L}_{ij}}]_{s_1}[\mu_{\mathcal{L}_{ij}}^{-1}a_2\beta_{2,j,1}\mu_{\mathcal{L}_{ij}}]_{s_2}\right.$$
$$+ [\mu_{\mathcal{L}_{ij}}^{-1}a_2\beta_{2,i,1}\mu_{\mathcal{L}_{ij}}]_{s_1}[\mu_{\mathcal{L}_{ij}}^{-1}a_1\beta_{1,j,1}\mu_{\mathcal{L}_{ij}}]_{s_2}\Big)x_0 q_{1,i}q_{2,j}$$

where $\mathcal{L}_{ij} = \beta_{1,i,0}a_1 + \beta_{2,j,0}a_2$. The full solution to (8.6) is

$$x(t) = \sum_{i=1}^2 \sum_{j=1}^2 x_{ij}(t). \tag{8.7}$$

In the special case when b_1 and b_2 are expressed in the same constant spectral basis $S = \{p_1, q_1, p_2, q_2\}$, the explicit full solution (8.7) simplifies to

$$\begin{aligned}
x(t) &= \sum_{i=1}^2 x_{ii}(t)\\
&= \mu_{\mathcal{L}_{11}}(t)x_0 p_1 + \mu_{\mathcal{L}_{22}}(t)x_0 p_2\\
&+ \mu_{\mathcal{L}_{11}}(t)\int_0^t ds_1 [\mu_{\mathcal{L}_{11}}^{-1}(a_2\beta_{2,1,1} + a_1\beta_{1,1,1})\mu_{\mathcal{L}_{11}}]_{s_1}x_0 q_1\\
&+ \mu_{\mathcal{L}_{22}}(t)\int_0^t ds_1 [\mu_{\mathcal{L}_{22}}^{-1}(a_2\beta_{2,2,1} + a_1\beta_{1,2,1})\mu_{\mathcal{L}_{22}}]_{s_1}x_0 q_2,
\end{aligned} \tag{8.8}$$

where $\mathcal{L}_{ii}(t) = b_{1,i,0}a_1 + b_{2,i,0}a_2$.

If a_i and $\beta_{i,j}$ are independent of time, (8.5) reduces to

$$\begin{aligned}
x(t) =& e^{\mathcal{L}t}x_0 + e^{\mathcal{L}t}\int_0^t ds\, e^{-\mathcal{L}s}(a_1\gamma_{1,1}e^{\mathcal{L}s}x_0 q_1 + e^{-\mathcal{L}s}a_2\gamma_{2,1}e^{\mathcal{L}s}x_0 q_2)\\
&+ e^{\mathcal{L}t}\int_0^t ds_1 \int_0^{s_1} ds_2 \left(e^{-\mathcal{L}s_1}\gamma_{1,1}a_1 e^{\mathcal{L}(s_1-s_2)}\gamma_{2,1}a_2 e^{\mathcal{L}s_2}\right.\\
&+ e^{-\mathcal{L}s_1}\gamma_{2,1}a_2 e^{\mathcal{L}(s_1-s_2)}\gamma_{1,1}a_1 e^{\mathcal{L}s_2}\Big)x_0 q_1 q_2,
\end{aligned} \tag{8.9}$$

where we have used $\mu_{\mathcal{L}} = e^{\mathcal{L}t} = e^{(a_1\gamma_{1,0}+a_2\gamma_{2,0})t}$ and $\mu_{\mathcal{L}}^{-1} = e^{-\mathcal{L}t} = e^{-(a_1\gamma_{1,0}+a_2\gamma_{2,0})t}$. When a_1 and a_2 commute, the solution (8.9) simplifies further,

$$x(t) = e^{\mathcal{L}t}x_0 + te^{\mathcal{L}t}(a_1\gamma_{1,1}x_0 q_1 + a_2\gamma_{2,1}x_0 q_2) + t^2 e^{\mathcal{L}t}\gamma_{1,1}\gamma_{2,1}a_1 a_2 x_0 q_1 q_2. \tag{8.10}$$

Finally, if $\gamma_{i,0} = 1$ and $\gamma_{i,1} = 0$, and a_1 and a_2 do not commute, we obtain

$$x(t) = e^{\mathcal{L}t}x_0 \tag{8.11}$$

where $\mathcal{L}(t) = (a_1 + a_2)t$, as the solution to

$$\frac{dx}{dt} = (a_1 + a_2)x$$

as expected.

Example 2. In this example, we consider the matrix differential equation

$$\frac{dX}{dt} = A_1 X B_1 + A_2 X B_2 \tag{8.12}$$

where X is a 3×3 matrix,

$$B_1 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

and $A_1 = B_1^T$, $A_2 = B_2^T$. It is easily checked that B_1 and B_2 commute, and so do A_1 and A_2. The matrices B_1 and B_2 can be written in the form

$$B_1 = -I_3 + Q_1, \quad B_2 = 2I_3 + Q_2$$

where the nilpotents Q_1 and Q_2 are uniquely determined and both have index of nilpotency 2. Note also that $Q_1 Q_2 = 0$. The solution to (8.12) is given by substituting these quantities into (8.10):

$$X(t) = e^{\mathcal{L}t}\Big(X(0) + t(A_1 X(0) Q_1 + A_2 X(0) Q_2)\Big). \tag{8.13}$$

Let $X(0) = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$, and note that

$$\mathcal{L} = -A_1 + 2A_2 = \begin{pmatrix} 5 & 0 & 0 \\ -1 & 5 & 0 \\ 2 & 0 & 5 \end{pmatrix} = 5I_3 + Q$$

where $Q = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$ and $Q^2 = 0$. From (2.12), it follows that $e^{\mathcal{L}t} = e^{5t}\,(I_3 + Qt)$ and the solution (8.13) takes the explicit form

$X(t) = e^{5t}\times$

$$\begin{pmatrix} c_{11} & c_{12} - c_{11}t & c_{13} + 2c_{11}t \\ c_{21} - c_{11}t & c_{22} + (c_{11} - c_{12} - c_{21})t + c_{11}t^2 & c_{23} + (-c_{13} + 2c_{21})t - 2c_{11}t^2 \\ c_{31} + 2c_{11}t & c_{32} + (2c_{12} - c_{31})t - 2c_{11}t^2 & c_{33} + (c_{11} + 2c_{13} + 2c_{31})t + 4c_{11}t^2 \end{pmatrix}$$

Example 3: Quaternions. We solve the differential equation $\frac{dx}{dt} = axb + f(t)$ where $a, b \in H_C$, the algebra of quaternions over the complex numbers. Any quaternion $a \in H_C$ can be written in the form

$$a = a_0 + \mathbf{a}$$

where $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ is the *complex vector* part of the quaternion, $a_k \in C$ are complex scalars, and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the basis elements of the associative quaternion algebra satisfying the famous defining relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

We solve the above equation under the assumptions that $a(t)$ is a general time-dependent quaternion-valued function, and $b(t)$ is a time-dependent quaternion-valued function of any of the three possible spectral types:

(I) $b(t) = b_0(t)$
(II) $b(t) = b_1(t)p_1 + b_2(t)p_2$
(III) $b(t) = b_0(t) + b_1(t)q$

where $b_0(t), b_1(t), b_2(t)$ are complex scalar valued functions of the real parameter t.

For type I, we find immediately that the solution is

$$x_H(t) = \mu_{b_0 a}(t) x_0,$$

for the homogeneous part, and

$$x(t) = \mu_{b_0 a}(t) \left[x_0 + \int_0^t \mu_{b_0 a}^{-1} f(s) ds \right]$$

for the full solution.

For type II, we find by superposition of the solutions of type I, that

$$x(t) = \mu_{b_1 a}(t) x_0 p_1 + \mu_{b_2 a}(t) x_0 p_2,$$

for the homogeneous part, and

$$x_H(t) = \sum_{i=1}^2 \mu_{b_i a}(t) \left[x_0 + \int_0^t \mu_{b_i a}^{-1}(s) f(s) ds \right] p_i$$

for the full solution.

For type III, we have, using the results from Section 6,

$$x_H(t) = E_2 \mu_{b_0 a} \mu_\Gamma(t) x_0 S^T,$$

for the homogeneous part, and

$$x(t) = E_2 \mu_{b_0 a} \mu_\Gamma(t) \left[x_0 + \int_0^t \mu_\Gamma^{-1}(s) \mu_{b_0 a}^{-1}(s) f(s) ds \right] S^T$$

for the full solution, $\Gamma = \gamma J_2$, $\gamma(t) = \mu_{b_0 a}^{-1} a \mu_{b_0 a} b_1 \in H_C$,

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is a nilpotent of index 2 and $S = \begin{pmatrix} 1 & q \end{pmatrix}$.

Next we derive the explicit representation of the solution for type III in terms of elementary operations in $\mathcal{A}$. Since $\Gamma^2 = 0$, we have

$$\mu_\Gamma(t) = \left(I_2 + J_2 \int_0^t ds\, \gamma(s) \right)$$

Similarly,

$$\mu_\Gamma^{-1}(t) = \left(I_2 - J_2 \int_0^t ds\, \gamma(s) \right) \tag{8.14}$$

The homogeneous solution can be written as

$$x_H(t) = \mu_{b_0 a}(t) \left(x_0 + \int_0^t ds\, \gamma(s) x_0 q \right)$$

while the non-homogenous part, $x_{NH} = x - x_H$ has a more complex form.

We derive the non-homogeneous part in stages:

$$x_{NH} = E_2 \mu_{b_0 a}(t) I_2 \left(1 + J_2 \int_0^t ds\, \gamma(s) \right) \left(\int_0^t ds\, \mu_\Gamma^{-1}(s) \mu_{b_0 a}^{-1}(s) f(s) \right) S^T$$

$$= E_2 \mu_{b_0 a}(t) I_2 \left(\int_0^t \mu_\Gamma^{-1}(s) \mu_{b_0 a}^{-1}(s) f(s) ds \right) S^T$$

$$+ E_2 J_2 \mu_{b_0 a}(t) \left(\int_0^t ds\, \gamma(s) \right) \left(\int_0^t ds\, I_2 \mu_{b_0 a}^{-1}(s) f(s) \right) S^T \tag{8.15}$$

where $J_2^2 = 0$ was used to eliminate some terms. Substituting (8.14) into (8.15), we find after simplification,

$$
x_{NH}(t) = \mu_{b_0 a}(t) \int_0^t ds\, \mu_{b_0 a}^{-1}(s) f(s)
$$
$$
- \mu_{b_0 a}(t) \int_0^t ds\, \mu_{b_0 a}^{-1}(s) f(s) \int_s^t ds_1\, \gamma(s_1) q \tag{8.16}
$$

REFERENCES

[1] A. Abian. *Linear Associative Algebras*. Pergamon Press Inc., 1971.
[2] P. Fernandes and B. Plateau. Efficient descriptor-vector multiplication in stochastic automata networks. *Journal of the ACM*, 45(3):381–414, May 1998.
[3] H. E. Heatherly and J. P. Huffman. Noncommutative operational calculus. *Electronic Journal of Differential Equations, http://ejde.math.unt.edu*, Conf. 02:11–18, 1999.
[4] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
[5] A. Iserles. Expansions that grow on trees. *Notices American Mathematics Society*, 49:430–440, April 2002.
[6] P. Lancaster. *Theory of Matrices*. Academic Press, 1969.
[7] H. Leiva and S. Siegmund. A necessary algebraic condition for controllability and observability of linear time-varying systems. *IEEE Trans. Automatic Control, (to appear)*, 2003.
[8] A. V. Mikhailov and V. V. Sokolov. Integrable ODEs on associated algebras. *Communications in Mathematical Physics*, 211:231–251, April 2000.
[9] Raúl Naulin. Exponential dichotomies for linear systems with impulsive effects. *Electronic Journal of Differential Equations, http://ejde.math.unt.edu*, Conf. 06, pp. 225–241, 2001.
[10] G. Sobczyk. The missing spectral basis in algebra and number theory. *The American Mathematical Monthly*, 108(4):336–346, 2001.
[11] H. Struemper. Nilpotent approximation and nilpotentization for under-actuated systems on matrix lie groups. In *Proceedings of the IEEE CDC '98*, pp. 16–18, 1998.
[12] X. Wang and J. K. Hale. On monodromy matrix computation. *Comput. Methods Appl. Mech. Engineering*, 190, pp. 2263–2275, 2001.

GORDON ERLEBACHER

DEPARTMENT OF MATHEMATICS, FLORIDA STATE UNIVERSITY, TALLAHASSEE, FL 32306, USA
E-mail address: `erlebach@math.fsu.edu`

GARRET E. SOBCZYK

UNIVERSIDAD DE LAS AMÉRICAS, DEPARTAMENTO DE FÍSICO-MATEMÁTICAS, APARTADO POSTAL #100, SANTA CATARINA MÁRTIR, 72820 CHOLULA, PUE., MÉXICO
E-mail address: `sobczyk@mail.udlap.mx`