university of groningen

faculty of mathematics
and natural sciences

The Magnus Expansion

Bachelor Project Mathematics

February 2017

Student: R.M. Hollander

First supervisor: dr. A.E. Sterk

Second supervisor: prof. dr. H.L. Trentelman

Abstract

For a homogeneous system of linear differential equations with a constant coefficient matrix, the fundamental matrix can be computed for example using the Jordan Canonical Form. However, when the coefficient matrix depends on a single variable t, this method does not always provide a correct solution. The fundamental matrix can be computed using a numerical method, for example the Picard iterative method, but using such method, one can lose important qualitative properties. Wilhelm Magnus provided a method to approximate the fundamental matrix, such that these qualitative properties are preserved. In this thesis, we will state Magnus' theorem and it's proof. We will compute the Magnus Expansion for some simple examples, and compare the solutions with the fundamental matrices obtained by applying Picard iteration.

Contents

1 Introduction

A differential equation is a mathematical equation that relates independent variables, functions, and derivatives of functions [1]. Differential equations play an important role in several disciplines such as physics, chemistry and engineering. Many fundamental laws of physics can be formulated as differential equations, for example the heat equation in thermodynamics, the Euler-Lagrange equation in classical mechanics and radioactive decay in nuclear physics.

An nth-order ordinary differential equation can be written in the general form

$$f(t, y(t), y'(t), ..., y^{(n)}(t)) = 0. \tag{1}$$

A solution is defined to be an n-times differentiable function $y(t)$ that satisfies equation (1). The differential equation is called *ordinary* because the function $y(t)$ depends on a single variable t. If the function depends on several independent variables, the equation is called a *partial differential equation*.

An nth-order differential equation is called explicit if it is of the form

$$y^n = f(t, y, y', ..., y^{(n-1)}),$$

otherwise it is called implicit [1].

A system of linear differential equations is of the form

$$y'(t) = A(t)y(t) + b(t).$$

When this system is homogeneous, the differential equation is given by

$$y'(t) = A(t)y(t), \tag{2}$$

where $y(t)$ is an $n \times 1$ vector of unknown functions and $A(t)$ is an $n \times n$ matrix called the coefficient matrix. This form of differential equations often occurs in for example systems theory and dynamical systems, where it is a state-space representation of a physical system without any input, with $A(t)$ denoting the 'state matrix' and $y(t)$ the 'state vector'.

Definition 1. The $n \times n$ matrix function $Y(t)$ is called a fundamental matrix of system (2) if and only if it is nonsingular and solves the matrix system $Y'(t) = A(t)Y(t)$.

1.1 Constant coefficient matrix

In the case $A(t) = A$ is a constant matrix, the general solution of the differential equation (2) with initial condition $y(0) = y_0$ is given by

$$y(t) = e^{At}y_0.$$

A fundamental matrix looks like

$$Y(t) = e^{At}, \tag{3}$$

and hence the matrix exponential e^{At} needs to be computed, which can be done for example with the Jordan Canonical Form of the matrix A.

Example 1. Consider the differential equation

$$y'(t) = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix} y(t),$$

with initial condition $y(0) = 1$.

The Jordan Canonical Form of the coefficient matrix A is given by

$$J = P^{-1}AP = \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Therefore, the matrix exponential is given by

$$e^{At} = Pe^{Jt}P^{-1} = P \begin{pmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{pmatrix} P^{-1} = \begin{pmatrix} e^{2t}(1-2t) & -4te^{2t} \\ te^{2t} & e^{2t}(1+2t) \end{pmatrix},$$

and hence the fundamental matrix is given by

$$Y(t) = \begin{pmatrix} e^{2t}(1-2t) & -4te^{2t} \\ te^{2t} & e^{2t}(1+2t) \end{pmatrix}.$$

1.2 Coefficient matrix dependent on variable t

Now consider the initial value problem

$$y'(t) = A(t)y(t), \qquad y(0) = y_0, \tag{4}$$

where $A(t)$ is a matrix that is entry-wise continuous in the variable t.

For $n = 1$, the fundamental matrix is given by

$$Y(t) = e^{(\int_0^t A(s)ds)}. \tag{5}$$

However, for $n > 1$, this fundamental matrix is not always a solution for the initial value problem. For example, if we take the coefficient matrix dependent on the variable t as follows

$$A(t) = \begin{pmatrix} 0 & 1 \\ -t & 0 \end{pmatrix},$$

and we approximate the matrix exponential using the expansion

$$e^{(\int\limits_0^t A(s)ds)} = \sum_{k=0}^{\infty} \frac{\left(\int_0^t A(s)ds\right)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} 0 & t \\ -\frac{1}{2}t^2 & 0 \end{pmatrix}^k,$$

we obtain the following formula for $\tilde{Y}(t)$ using equation (5)

$$\tilde{Y}(t) \approx \begin{pmatrix} 1 - \frac{1}{4}t^3 + \frac{1}{96}t^6 - \frac{1}{5760}t^9 - \cdots & t - \frac{1}{2}t^4 + \frac{1}{480}t^7 + \cdots \\ -\frac{1}{2}t^2 + \frac{1}{24}t^5 - \frac{1}{960}t^8 - \cdots & 1 - \frac{1}{4}t^3 + \frac{1}{96}t^6 - \frac{1}{5760}t^9 - \cdots \end{pmatrix}.$$

Assume that $\tilde{Y}(t)$ is a fundamental matrix of the initial value problem. Calculating the lefthandside of (4) by taking the derivative of $\tilde{Y}(t)$, gives us

$$\tilde{Y}'(t) \approx \begin{pmatrix} -\frac{3}{4}t^2 + \frac{1}{16}t^5 + \frac{1}{640}t^8 - \cdots & 1 - \frac{1}{3}t^3 + \frac{7}{480}t^6 + \cdots \\ -t + \frac{5}{24}t^4 - \frac{8}{960}t^7 - \cdots & -\frac{3}{4}t^2 + \frac{1}{16}t^5 + \frac{1}{640}t^8 - \cdots \end{pmatrix},$$

and calculating the righthandside, we obtain

$$A(t)\tilde{Y}(t) = \begin{pmatrix} -\frac{1}{2}t + \frac{1}{24}t^5 - \frac{1}{960}t^8 - \cdots & 1 - \frac{1}{4}t^3 + \frac{1}{96}t^6 - \frac{1}{5760}t^9 - \cdots \\ -t + \frac{1}{4}t^4 - \frac{1}{96}t^7 + \frac{1}{5760}t^10 + \cdots & -t^2 + \frac{1}{2}t^5 - \frac{1}{480}t^8 - \cdots \end{pmatrix}.$$

Although these matrices seem almost equivalent, the coefficients are different and therefore they are not exactly the same. However, this should be the case if the fundamental matrix $\tilde{Y}(t)$ satisfies equation (4). Hence our assumption is wrong and therefore $\tilde{Y}(t)$ is not a fundamental matrix for the initial value problem.

It is often necessary to approximate the fundamental matrix using a numerical method, since a closed form analytic solution may not always be possible to find [8]. A well-known method which can be used to approximate the solutions of a differential equation with coefficient matrix that is not constant, is the Picard iterative method.

1.3 Picard Iteration

Suppose we want to find the fundamental matrix satisfying the initial value problem

$$Y'(t) = A(t)Y(t), \qquad Y(0) = I.$$

Integrating $Y'(\tau)$ with respect to τ from 0 to t, and using the Fundamental Theorem of Calculus, we obtain

$$Y(t) - Y(0) = \int_0^t A(\tau)Y(\tau)d\tau,$$

which gives the fixed point equation for the differential equation

$$Y(t) = I + \int_0^t A(\tau)Y(\tau)d\tau.$$

Assuming that $A(t)$ is integrable and starting with the approximation $Y_0(t) = I$, we obtain the Picard iterative scheme

$$\begin{aligned}
Y_0(t) &= I \\
Y_1(t) &= I + \int_0^t A(\tau)Y_0(\tau)d\tau \\
&\vdots \\
Y_n(t) &= I + \int_0^t A(\tau)Y_{n-1}(\tau)d\tau,
\end{aligned}$$

which converges to the fundamental matrix $Y(t)$ satisfying the initial value problem.

Picard iteration as constructive method can be used to prove the general existence and uniqueness theorems [8]. The fundamental matrix approximated with the Picard iterative method is not in the form of an exponential, whereas for a constant coefficient matrix the fundamental is of exponential form (3). However, in 1954, Wilhelm Magnus provided another method to approximate the fundamental matrix of an initial value problem that is of exponential form, called the Magnus Expansion. Later in this article, the Picard iterative method will be used to make a comparison with the Magnus Expansion.

2 The Magnus Expansion

Wilhelm Magnus (1907-1990) was a German American mathematician who studied at the University of Frankfurt, where he received his doctorate in 1931. Some of the subjects he worked on were combinatorial group theory and Lie algebras [7].

Magnus' *On the Exponential Solution of Differential Equations for a Linear Operator* [5] was stimulated by a paper of K.O. Friedrichs, who encountered some purely algebraic problems in connection with the theory of linear operators in quantum mechanics [5]. Friedrichs considered the initial value problem given by (4), where $A(t)$ is the linear operator depending on a real variable t. In his article, Magnus provides a way to approximate the fundamental matrix of this system. His approach is to express the solution as the exponential of the $n \times n$ matrix function $\Omega(t)$,

$$Y(t) = \exp \Omega(t),$$

which is subsequently constructed as a series expansion, also called the *Magnus Expansion*,

$$\Omega(t) = \sum_{k=0}^{\infty} \Omega_k(t).$$

It is important to note that $\Omega(t)$ is not the same as $\int_0^t A(s)ds$, therefore the Magnus Expansion yields a solution different from (5). The expression for $\Omega(t)$ satisfies the condition that the partial sums of this series become Hermitian after multiplication by i if iA is a Hermitian operator [5]. The original theorem and proof given by Magnus was published in *On the Exponential Solution of Differential Equations for a Linear Operator* [5]. In this article, we follow the exposition by Blanes et al. [6]. For the reader's convenience, before stating the theorem, we will first give some mathematical expressions and notations that appear most frequently.

2.1 Preliminaries

The mathematical concept that the Magnus Expansion mostly concerns is a Lie group, or its associated Lie algebra. In *Introduction to Lie Algebras* [2], the following definition of a Lie algebra is defined:

Definition 2. Let F be a field. A *Lie algebra* over F is an F-vector space $\mathfrak{g}$, together with a map, the *Lie bracket*

$$[-,-] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}, \qquad (A, B) \longmapsto [A, B], \tag{6}$$

satisfying the following properties

(i) this map is bilinear;

(ii) $[A, A] = 0$ for all $A \in \mathfrak{g}$;

(iii) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$ for all $A, B, C \in \mathfrak{g}$.

The Lie bracket $[A, B]$ is referred to as the *commutator* of A and B, and condition (iii) is also called the Jacobi Identity. Since, by condition (i) the map is bilinear, we have

$$\begin{aligned} 0 &= [A + B, A + B] \\ &= [A, A] + [A, B] + [B, A,] + [B, B] \\ &= [A, B] + [B, A], \end{aligned}$$

and therefore condition (ii) can also be stated as follows

$$[A, B] = -[B, A].$$

Example 2. For the general linear group of invertible matrices $GL_n(\mathbb{R})$, the Lie bracket is defined as the commutator of matrices

$$[A, B] = AB - BA, \qquad A, B \in GL_n(\mathbb{R}).$$

Indeed, since

$$\begin{aligned} [\lambda A + \mu B, C] &= (\lambda A + \mu B)C - C(\lambda A + \mu B) \\ &= \lambda AC + \mu BC - \lambda CA - \mu CB \\ &= \lambda(AC - CA) + \mu(BC - CB) \\ &= \lambda[A, C] + \mu[B, C], \end{aligned}$$

and in the same way

$$[A, \lambda B + \mu C] = \lambda[A, B] + \mu[A, C],$$

this map is bilinear.

Furthermore, for any $A \in GL_n(\mathbb{R})$, we clearly have $[A, A] = AA - AA = 0$, and

$$\begin{aligned} [A, [B, C]] &= [A, BC - CB] \\ &= A(BC - CB) - (BC - CB)A \\ &= ABC - ACB - BCA + CBA. \end{aligned}$$

Likewise, we have

$$[B, [C, A]] = BCA - BAC - CAB + ACB,$$

and

$$[C, [A, B]] = CAB - CBA - ABC + BAC.$$

Adding these terms yields $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$, which proves that the commutator also satisfies condition (iii) of Definition 2. Therefore the commutator of matrices is a Lie bracket for the general group of invertible matrices.

Definition 3. Associated with any $A \in \mathfrak{g}$ we can define the linear adjoint operator

$$\mathrm{ad}_A : \mathfrak{g} \longrightarrow \mathfrak{g},$$

which acts according to

$$\mathrm{ad}_A(B) = [A, B], \qquad \mathrm{ad}_A^j(B) = [A, \mathrm{ad}_A^{j-1}(B)], \qquad \mathrm{ad}_A^0(B) = B, \qquad j \in \mathbb{N}, B \in \mathfrak{g}. \tag{7}$$

The exponential of this ad_A operator is given by

$$\mathrm{Ad}_A = \exp(\mathrm{ad}_A), \tag{8}$$

and acts according to

$$\mathrm{Ad}_A(B) = \exp(A)B\exp(-A) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathrm{ad}_A^k(B), \qquad B \in \mathfrak{g}. \tag{9}$$

Definition 4. A matrix norm is a non-negative real number $\|A\|$ associated with each matrix $A \in \mathbb{C}^{n\times n}$ satisfying

(i) $\|A\| \geq 0$ and $\|A\| = 0$ if and only if $A = O_n$;

(ii) $\|\alpha A\| = |\alpha|\|A\|$ for all scalars α;

(iii) $\|A + B\| \leq \|A\| + \|B\|$.

Sometimes also the sub-multiplicative property,

$$\|AB\| \leq \|A\|\|B\|, \tag{10}$$

is added, but not every norm satisfies this property [6].

Corollary 1. *If we consider a norm in a matrix Lie algebra $\mathfrak{g}$ satisfying the sub-multiplicative property* (10)*, it is clear that $\|[A, B]\| \leq 2\|A\|\|B\|$ and $\|ad_A\| \leq 2\|A\|$ for any matrix $A \in \mathfrak{g}$ [6].*

In Magnus' proposal, also the so-called Bernoulli numbers B_n occur.

The Bernoulli numbers are a sequence of rational numbers with many interesting arithmetic properties [3]. They are defined as coefficients in the following Taylor series

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}, \tag{11}$$

and it can easily be proven that $B_n \in \mathbb{Q}$ for all n. The first few nonzero Bernoulli numbers are $B_0 = 1, B_1 = \frac{1}{2}, B_2 = \frac{1}{6}, B_3 = \frac{1}{30}$. In general one has $B_{2m+1} = 0$ for $m \geq 1$ [6].

Lemma 1. *The radius of convergence of the series expansion given by* $\frac{t}{e^t-1}$ *is* 2π*.*

Proof. This Lemma can be proven by complex analysis. Since the denominator t of $\frac{t}{e^t-1}$ is an entire function, the only non-removable singularities will occur when the denominator is 0. This is the case for $k = \pm n\pi i$ for $n = 1, 2, \ldots$. The singularities closest to the origin, which is the center of the power series, are at $k = \pm 2\pi i$. The distance from the center to either one of those points is 2π, so the radius of convergence is 2π. □

2.2 Magnus' proposal and proof

Before stating Magnus' theorem, we will give two lemma's which will be used to prove the theorem. The lemma's and their proofs are from Blanes et al. [6]

Lemma 2. *The derivative of a matrix exponential can be written alternatively as*

(i) $\frac{d}{dt}\exp(\Omega(t)) = d\exp_{\Omega(t)}(\Omega'(t))\exp(\Omega(t))$

(ii) $\frac{d}{dt}\exp(\Omega(t)) = \exp(\Omega(t))\, d\exp_{-\Omega(t)}(\Omega'(t))$

(iii) $\frac{d}{dt}\exp(\Omega(t)) = \int_0^1 e^{\alpha\Omega(t)}\Omega'(t)e^{(1-\alpha)\Omega(t)}d\alpha$

where $d\exp_\Omega(C)$ *is defined by its (everywhere convergent) power series*

$$d\exp_\Omega(C) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\ \mathrm{ad}\,_\Omega^k(C) \equiv \frac{\exp(\mathrm{ad}_\Omega) - I}{\mathrm{ad}_\Omega}(C).$$

Proof. (i) Let $\Omega(t)$ be a matrix-valued differentiable function and define

$$Y(\alpha, t) \equiv \frac{\partial}{\partial t}\Big(\exp(\alpha\Omega(t))\Big)\exp(-\alpha\Omega(t)).$$

Then taking the derivative of $Y(\alpha, t)$ with respect to α yields

$$\begin{aligned}
\frac{\partial Y}{\partial \alpha} &= \frac{\partial}{\partial t}\Big(\Omega\exp(\alpha\Omega)\Big)\exp(-\alpha\Omega) + \frac{\partial}{\partial t}\Big(\exp(\alpha\Omega)\Big)(-\Omega)\exp(-\alpha\Omega) \\
&= \left(\Omega'\exp(\alpha\Omega) + \Omega\frac{\partial}{\partial t}\Big(\exp(\alpha\Omega)\Big)\right)\exp(-\alpha\Omega) - \frac{\partial}{\partial t}\Big(\exp(\alpha\Omega)\Big)\Omega\exp(-\alpha\Omega) \\
&= \Omega'\exp(\alpha\Omega)\exp(-\alpha\Omega) + \Omega\frac{\partial}{\partial t}\exp(\alpha\Omega)\exp(-\alpha\Omega) - \frac{\partial}{\partial t}\Big(\exp(\alpha\Omega)\Big)\exp(-\alpha\Omega) \\
&= \exp(\alpha\Omega)\Omega'\exp(-\alpha\Omega).
\end{aligned} \tag{12}$$

Combining this with equations (8) and (9) from Definition 3 , we obtain

$$\frac{\partial Y}{\partial \alpha} = \exp(\mathrm{ad}_{\alpha\Omega})(\Omega') = \sum_{k=0}^{\infty}\frac{\alpha^k}{k!}\mathrm{ad}_\Omega^k(\Omega'). \tag{13}$$

Furthermore, we observe that

$$\frac{d}{dt}(\exp\Omega)\exp(-\Omega) = Y(1,t), \tag{14}$$

and since $Y(0,t) = 0$, we have

$$Y(1,t) = \int_0^1 \frac{\partial}{\partial\alpha} Y(\alpha,t)d\alpha. \tag{15}$$

If we take equation (13), (14) and (15) together, we obtain

$$\frac{d}{dt}(\exp\Omega)\exp(-\Omega) = \int_0^1 \sum_{k=0}^{\infty} \frac{\alpha^k}{k!}\mathrm{ad}_\Omega^k(\Omega')d\alpha = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}_\Omega^k(\Omega'),$$

and dividing both sides by $\exp(-\Omega)$ yields

$$\frac{d}{dt}(\exp\Omega) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\mathrm{ad}_\Omega^k(\Omega')\exp(\Omega) = d\exp_\Omega(\Omega')\exp(\Omega),$$

which proves part (i).

(ii) Multiplying both sides of (i) with $\exp(-\Omega)$ we obtain

$$\begin{aligned}
\exp(-\Omega)\frac{d}{dt}\exp\Omega &= \exp(-\Omega)d\exp_\Omega(\Omega')\exp(\Omega) \\
&= \exp(\mathrm{ad}_{-\Omega})d\exp_\Omega(\Omega') \text{ (by equation (9) from Definition 3)} \\
&= \exp(\mathrm{ad}_{-\Omega})\frac{\exp(\mathrm{ad}_\Omega) - I}{\mathrm{ad}_\Omega}(\Omega') \\
&= -\frac{I - \exp(\mathrm{ad}_{-\Omega})}{\mathrm{ad}_\Omega}(\Omega') \\
&= \frac{I - \exp(\mathrm{ad}_{-\Omega})}{\mathrm{ad}_{-\Omega}}(\Omega') \\
&= d\exp_{-\Omega}(\Omega').
\end{aligned}$$

Then multiplying left and right with $\exp(\Omega)$ gives us

$$\frac{d}{dt}\exp(\Omega(t)) = \exp(\Omega(t))d\exp_{-\Omega(t)}(\Omega'(t)).$$

(iii) Using (14) and (15) and filling in $\frac{\partial}{\partial\alpha}Y(\alpha,t)$ obtained in (12), yields

$$\frac{d}{dt}(\exp\Omega)\exp(-\Omega) = \int_0^1 \frac{\partial}{\partial\alpha}Y(\alpha,t)d\alpha = \int_0^1 \exp(\alpha\Omega)\Omega'\exp(-\alpha\Omega)d\alpha.$$

Multiplying both sides with $\exp(\Omega)$ gives us

$$\begin{aligned}\frac{d}{dt}(\exp\Omega) &= \exp(\Omega)\int_0^1 \exp(\alpha\Omega)\Omega'\exp(-\alpha\Omega)d\alpha\\ &= \int_0^1 \exp(\alpha\Omega)\Omega'\exp(\Omega)\exp(-\alpha\Omega)d\alpha\\ &= \int_0^1 e^{\alpha\Omega}\Omega' e^{\Omega}e^{-\alpha\Omega}d\alpha\\ &= \int_0^1 e^{\alpha\Omega}\Omega' e^{(1-\alpha)\Omega}d\alpha,\end{aligned}$$

which proves (iii).

□

Lemma 3. *If the eigenvalues of the linear operator ad_Ω are different from $2m\pi i$ with $m \in \{\pm1, \pm2, \ldots\}$, then $d\exp_\Omega$ is invertible. Furthermore,*

$$d\exp_\Omega^{-1}(C) = \frac{\mathrm{ad}_\Omega}{e^{\mathrm{ad}_\Omega} - I}C = \sum_{k=0}^{\infty}\frac{B_k}{k!}\ \mathrm{ad}_\Omega^k(C),$$

and the convergence of the $d\exp_\Omega^{-1}$ expansion is assured if $\|\Omega\| < \pi$.

Proof. The eigenvalues of $d\exp_\Omega$ are of the form

$$\lambda = \sum_{k=0}^{\infty}\frac{\mu^k}{(k+1)!} = \frac{e^\mu - 1}{\mu},$$

where μ is an eigenvalue of ad_Ω. The values of λ are non-zero, therefore $d\exp_\Omega$ is invertible. Furthermore, we have

$$I = \frac{e^{\mathrm{ad}_\Omega} - I}{\mathrm{ad}_\Omega}(C)\cdot\frac{\mathrm{ad}_\Omega}{e^{\mathrm{ad}_\Omega} - I}(C) = \mathrm{d}\exp_\Omega(C)\cdot\mathrm{d}\exp_\Omega^{-1}(C).$$

Hence, using the Taylor series in (11), we obtain

$$\mathrm{d}\exp_\Omega^{-1}(C) = \frac{\mathrm{ad}_\Omega}{e^{\mathrm{ad}_\Omega} - I}(C) = \sum_{k=0}^{\infty}\frac{B_k}{k!}\mathrm{ad}_\Omega^k.$$

From Corollary 1, we have that $\|\mathrm{ad}_\Omega\| \leq 2\|\Omega\|$, where $\|\Omega\|$ is the norm in Lie algebra g satisfying the sub-multiplicative property (10). Hence, if $\|\Omega\| < \pi$, we have

$$\|\mathrm{ad}_\Omega\| \leq 2\|\Omega\| < 2\pi.$$

Therefore by Lemma 1, we are within the radius of convergence of

$$\frac{\mathrm{ad}_\Omega}{e^{\mathrm{ad}_\Omega} - I}(C),$$

thus the expansion $d\exp_\Omega^{-1}$ is convergent for $\|\Omega\| < \pi$. □

Using Lemma (2) and (3), in Blanes et al. the theorem of Magnus can be rewritten in the following way:

Theorem 1. *(Rewritten Magnus' Theorem) Let $A(t)$ be a known function of t in an associative ring R, and let $Y(t)$ be an unknown function satisfying*

$$\frac{dY}{dt} = AY, \qquad Y(0) = 1.$$

Then, if certain unspecified conditions of convergence are satisfied, $Y(t)$ can be written in the form

$$Y(t) = \exp \Omega(t),$$

with $\Omega(t)$ defined by

$$\Omega'(t) = d\exp_{\Omega}^{-1}(A(t)), \qquad \Omega(0) = 0, \tag{16}$$

and

$$d\exp_{\Omega}^{-1}(A) = \sum_{k=0}^{\infty} \frac{B_k}{k!}\ \mathrm{ad}_{\Omega}^{k}(A), \tag{17}$$

where B_k are the Bernoulli numbers. Integration of (16) leads to an infinite series for $\Omega(t)$, the first terms of which are

$$\Omega(t) = \int_0^t A(t_1)dt_1 - \frac{1}{2}\int_0^t \left[\int_0^{t_1} A(t_2)dt_2, A(t_1)\right] dt_1 + \cdots . \tag{18}$$

Proof. Taking the derivative of $Y(t) = \exp(\Omega(t))Y_0$ and using Lemma 2 (i) yields

$$\frac{dY}{dt} = \frac{d}{dt}\Big(\exp(\Omega(t))\Big)Y_0 \equiv d\exp_{\Omega}(\Omega')\exp(\Omega(t))Y_0.$$

Comparing this with $Y'(t) = A(t)Y(t)$ we obtain

$$A(t) = d\exp_{\Omega}(\Omega'),$$

and applying the inverse operator of $d\exp_{\Omega}$ from Lemma 3 yields

$$\Omega'(t) = d\exp_{\Omega}^{-1}(A(t)) = \sum_{k=0}^{\infty} \frac{B_k}{k!}\mathrm{ad}_{\Omega}^{k}(A(t))$$

which proves (16) and (17).

The proof of (18) is carried out by taking the numerical values of the Bernoulli numbers and writing out the first terms of $\Omega'(t)$ as

$$\Omega'(t) = A(t) - \frac{1}{2}[\Omega(t), A(t)] + \frac{1}{12}[\Omega(t), [\Omega(t), A(t)]] + \cdots .$$

By defining

$$\Omega_0(t) = 0, \qquad \Omega_1(t) = \int_0^t A(t_1) dt_1,$$

and applying Picard fixed point iteration, one gets

$$\Omega_n(t) = \int_0^t \left(A(t_1) - \frac{1}{2}[\Omega_{n-1}(t), A] + \frac{1}{12}[\Omega_{n-1}(t), [\Omega_{n-1}(t), A]] + \cdots \right) dt_1.$$

By putting $\Omega(t) = \lim_{n\to\infty}\Omega_n(t)$, we obtain equation (18).

$\square$

The first terms of the Magnus expansion are given by

$$\begin{aligned}
\Omega_1(t) &= \int_0^t A(t_1) dt_1 \\
\Omega_2(t) &= \frac{1}{2}\int_0^t \int_0^{t_1} \big[A(t_1), A(t_2)\big] dt_2 dt_1 \\
\Omega_3(t) &= \frac{1}{6}\int_0^t \int_0^{t_1} \int_0^{t_2} \Big(\big[A(t_1), [A(t_2), A(t_3)]\big] + \big[[A(t_1), A(t_2)], A(t_3)\big] \Big) dt_3 dt_2 dt_1 \\
\Omega_4(t) &= \frac{1}{12}\int_0^t \int_0^{t_1} \int_0^{t_2} \int_0^{t_3} \Bigg(\Big[\big[[A(t_1), A(t_2)], A(t_3)\big], A(t_4)\Big] + \Big[A(t_1), \big[[A(t_2), A(t_3)], A(t_4)\big]\Big] \\
&\quad + \Big[A(t_1), \big[A(t_2), [A(t_3), A(t_4)]\big]\Big] + \Big[A(t_2), \big[A(t_3), [A(t_4), A(t_1)]\big]\Big] \Bigg) dt_4 dt_3 dt_2 dt_1.
\end{aligned} \tag{19}$$

These first terms seem to be symmetric, they contain combinations of Lie brackets and integrals that can easily be computed. However, this symmetry is deceptive and the terms $\Omega_k(t)$ are increasingly complex, therefore higher terms of the expansion are complicated to compute.

3 Simple Examples

Again, consider the initial value problem given by

$$Y'(t) = A(t)Y(t), \qquad Y(0) = I. \tag{20}$$

In Example 1 in Section 1.1, we computed the fundamental matrix for (20) with the constant coefficient matrix

$$A(t) = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix}.$$

Using the Jordan Canonical Form, we saw that the fundamental matrix of this system is of the form

$$Y(t) = \begin{pmatrix} e^{2t}(1-2t) & -4te^{2t} \\ te^{2t} & e^{2t}(1+2t) \end{pmatrix}.$$

Now, we will apply the Magnus Expansion to this example. Computing the first term of the expansion, we obtain

$$\Omega_1(t) = \int_0^t \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix} dt_1 = \begin{pmatrix} 0 & -4t \\ t & 4t \end{pmatrix}.$$

Since the matrix A is constant, it is clear that $A(t_1)(At_2) = A(t_2)A(t_1)$. Hence the matrix is commutative, so

$$[A(t_1), A(t_2)] = A(t_1)(At_2) - A(t_2)A(t_1) = 0.$$

All Lie brackets appearing in the terms of the Magnus Expansion cancel out, so the matrix function $\Omega(t)$ only consists of $\Omega_1(t)$. Therefore the fundamental matrix is given by

$$Y(t) = \exp \Omega(t) = \exp \Omega_1(t) = \exp \begin{pmatrix} 0 & -4t \\ t & 4t \end{pmatrix} = \begin{pmatrix} e^{2t}(1-2t) & -4te^{2t} \\ te^{2t} & e^{2t}(1+2t) \end{pmatrix},$$

which is exactly the same as the fundamental matrix computed with the Jordan Canonical Form. In fact, for all systems with coefficient matrices that are commutative, the only nonzero term in the Magnus Expansion is $\Omega_1(t)$. In all these cases, the fundamental matrix is of the form

$$Y(t) = e^{\int_0^t A(t_1)dt_1}. \tag{21}$$

Now, consider the initial value problem (20) with coefficient matrices $A(t)$ dependend on the variable t. For the non-commutative examples we compute the fundamental matrix using Magnus' approach with the first four terms of the expansion. We also approximate the fundamental matrix using the Picard iterative method (as seen in Section 1.3), and we will make a comparison between the two methods.

3.1 The 2×2 rotation matrix

Suppose we take as coefficient matrix the 2×2 rotation matrix

$$A(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Computing the first term of the Magnus Expansion, we obtain

$$\Omega_1(t) = \int_0^t \begin{pmatrix} \cos(t_1) & -\sin(t_1) \\ \sin(t) & \cos(t) \end{pmatrix} dt_1 = \begin{pmatrix} \sin(t) & \cos(t) \\ -\cos(t) & \sin(t) \end{pmatrix}.$$

However, since

$$\begin{aligned} A(t_1)A(t_2) &= \begin{pmatrix} \cos(t_1) & -\sin(t_1) \\ \sin(t_1) & \cos(t_1) \end{pmatrix} \begin{pmatrix} \cos(t_2) & -\sin(t_2) \\ \sin(t_2) & \cos(t_2) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t_1)\cos(t_2) - \sin(t_1)\sin(t_2) & -\cos(t_1)\sin(t_2) - \sin(t_1)\cos(t_2) \\ \sin(t_1)\cos(t_2) + \cos(t_1)\sin(t_2) & -\sin(t_1)\sin(t_2) + \cos(t_1)\cos(t_2) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t_2)\cos(t_1) - \sin(t_2)\sin(t_1) & -\cos(t_2)\sin(t_1) - \sin(t_2)\cos(t_1) \\ \sin(t_2)\cos(t_1) + \cos(t_2)\sin(t_1) & -\sin(t_2)\sin(t_1) + \cos(t_2)\cos(t_1) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t_2) & -\sin(t_2) \\ \sin(t_2) & \cos(t_2) \end{pmatrix} \begin{pmatrix} \cos(t_1) & -\sin(t_1) \\ \sin(t_1) & \cos(t_1) \end{pmatrix} = A(t_2)A(t_1), \end{aligned}$$

the 2×2 rotation matrix is commutative and therefore the only nonzero term in the Magnus Expansion is $\Omega_1(t)$. Hence the fundamental matrix is of the form (21).

3.2 Simple 2×2 matrix dependent on t

Suppose we take the 2×2 coefficient matrix

$$A(t) = \begin{pmatrix} 2 & t \\ 0 & -1 \end{pmatrix}.$$

For $t_1 \neq t_2$, we have

$$\begin{aligned} A(t_1)A(t_2) = \begin{pmatrix} 2 & t_1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & t_2 \\ 0 & -1 \end{pmatrix} &= \begin{pmatrix} 4 & 2t_2 - t_1 \\ 0 & 1 \end{pmatrix} \\ &\neq \begin{pmatrix} 4 & 2t_1 - t_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & t_2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & t_1 \\ 0 & -1 \end{pmatrix} = A(t_2)A(t_1), \end{aligned}$$

so the coefficient matrix $A(t)$ is not commutative. Therefore the Lie brackets appearing in the terms of the Magnus Expansion will not necessarily cancel out.

Computing the first four terms of the Magnus Expansion, we obtain

$$\Omega_1(t) = \int_0^t \begin{pmatrix} 2 & t_1 \\ 0 & -1 \end{pmatrix} dt_1 = \begin{pmatrix} 2t & \frac{1}{2}t^2 \\ 0 & -t \end{pmatrix}$$

$$\Omega_2(t) = \frac{1}{2}\int_0^t \int_0^{t_1} \begin{pmatrix} 0 & t_2 - 3t_1 \\ 0 & 0 \end{pmatrix} dt_1 dt = \begin{pmatrix} 0 & -\frac{5}{12}t^3 \\ 0 & 0 \end{pmatrix}$$

$$\Omega_3(t) = \frac{1}{6}\int_0^t \int_0^{t_1} \int_0^{t_2} \begin{pmatrix} 0 & 9t_1 - 12t_2 + 3t_3 \\ 0 & 0 \end{pmatrix} dt_3 dt_2 dt_1 = \begin{pmatrix} 0 & \frac{1}{24}t^4 \\ 0 & 0 \end{pmatrix}$$

$$\Omega_4(t) = \frac{1}{12}\int_0^t \int_0^{t_1} \int_0^{t_2} \int_0^{t_3} \begin{pmatrix} 0 & -18t_1 + 36t_2 - 36t_3 - 18t_4 \\ 0 & 0 \end{pmatrix} dt_4 dt_3 dt_2 dt_1 = \begin{pmatrix} 0 & -\frac{1}{36}t^5 \\ 0 & 0 \end{pmatrix}.$$

Adding these terms gives us an approximation for the matrix function $\Omega(t)$ given by

$$\Omega(t) \approx \begin{pmatrix} 2t & \frac{1}{2}t^2 - \frac{5}{12}t^3 + \frac{1}{24}t^4 - \frac{1}{36}t^5 \\ 0 & -t \end{pmatrix}.$$

Using Mathematica [10], we compute the exponential of this approximation of $\Omega(t)$ to obtain an approximation for the fundamental matrix of the system. This fundamental matrix is given by

$$Y(t) = \exp\Omega(t) \approx \begin{pmatrix} 1 + 2t + 2t^2 + \frac{4t^3}{3} + \frac{2t^4}{3} + \cdots & \frac{t^2}{2} - \frac{t^3}{6} + \frac{t^4}{12} - \frac{t^5}{9} - \frac{23t^6}{288} - \cdots \\ 0 & 1 - t + \frac{t^2}{2} - \frac{t^3}{6} + \frac{t^4}{24} - \cdots \end{pmatrix}.$$

The fundamental matrix obtained by applying the Picard iterative method is given by

$$Y(t) \approx \begin{pmatrix} 1 + 2t + 2t^2 + \frac{4t^3}{3} + \frac{2t^4}{3} + \cdots & \frac{t^2}{2} + \frac{t^4}{8} + \frac{t^5}{60} + \frac{t^6}{80} - \cdots \\ 0 & 1 - t + \frac{t^2}{2} - \frac{t^3}{6} + \frac{t^4}{24} - \cdots \end{pmatrix}$$

We see that the fundamental matrices obtained by the different methods are, besides some coefficients and one different power, almost equivalent.

3.3 Airy's Equation

The Airy function $Ai(t)$, named after the British astronomer George Biddell Airy, and the related Bairy function $Bi(t)$ are linearly independent solutions to Airy's equation

$$\frac{d^2y}{dt^2} - ty = 0.$$

Suppose we take the coefficient matrix of the form

$$A(t) = \begin{pmatrix} 0 & 1 \\ -t & 0 \end{pmatrix},$$

thus we want to solve the 2×2 system of differential equations given by

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -t & 0 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

This matrix differential equation is equivalent to the second order differential equation

$$y_1''(t) - ty_1(t) = 0,$$

which is equal to Airy's equation. Hence solving the initial value problem given by (20) with coefficient matrix $A(t)$ is equivalent to solving the second order Airy's equation.

One can easily prove that the matrix $A(t)$ is not commutative. Computing the first four terms of the Magnus Expansion for Airy's equation, we obtain

$$\Omega_1(t) = \int_0^t \begin{pmatrix} 0 & 1 \\ -t_1 & 0 \end{pmatrix} dt_1 = \begin{pmatrix} 0 & t \\ -\frac{1}{2}t^2 & 0 \end{pmatrix}$$

$$\Omega_2(t) = \frac{1}{2} \int_0^t \int_0^{t_1} \begin{pmatrix} -t_2 + t_1 & 0 \\ 0 & -t_1 + t_2 \end{pmatrix} dt_2 dt_1 = \begin{pmatrix} \frac{1}{6}t^3 & 0 \\ 0 & -\frac{1}{6}t^3 \end{pmatrix}$$

$$\Omega_3(t) = \frac{1}{6} \int_0^t \int_0^{t_1} \int_0^{t_2} \begin{pmatrix} 0 & -4t_2 + 2t_3 + 2t_1 \\ -2t_1t_2 + 4t_1t_3 - 2t_2t_3 & 0 \end{pmatrix} dt_3 dt_2 dt_1$$

$$= \begin{pmatrix} 0 & 0 \\ -\frac{1}{120}t^5 & 0 \end{pmatrix}$$

$$\Omega_4(t) = \frac{1}{12} \int_0^t \cdots \int_0^{t_3} \begin{pmatrix} 4(t_1t_2 - t_1t_3 + t_2t_4 - t_3t_4) & 0 \\ 0 & 4(-t_1t_2 + t_1t_3 - t_2t_4 + t_3t_4) \end{pmatrix} dt_4 dt_3 dt_2 dt_1$$

$$= \begin{pmatrix} \frac{1}{360}t^6 & 0 \\ 0 & -\frac{1}{360}t^6 \end{pmatrix}.$$

By adding these matrices, we obtain an approximation for $\Omega(t)$ given by

$$\Omega(t) \approx \begin{pmatrix} \frac{1}{6}t^3 + \frac{1}{360}t^6 & t \\ -\frac{1}{2}t^2 - \frac{1}{120}t^5 & -\frac{1}{6}t^3 - \frac{1}{360}t^6 \end{pmatrix}.$$

Using Mathematica [10], we compute the exponential of this approximation of $\Omega(t)$ to obtain the fundamental matrix. This fundamental matrix is given by

$$Y(t) = \exp\Omega(t) \approx \begin{pmatrix} 1 - \frac{t^3}{12} - \frac{t^6}{720} + \frac{t^9}{1296} + \cdots & t - \frac{t^4}{12} + \frac{7t^7}{2160} + \cdots \\ -\frac{t^2}{2} + \frac{t^5}{30} - \frac{t^8}{1080} - \cdots & 1 - \frac{5t^3}{12} + \frac{t^6}{48} + \frac{t^9}{6480} - \cdots \end{pmatrix}.$$

The fundamental matrix obtained by applying the Picard iterative method is given by

$$Y(t) \approx \begin{pmatrix} 1 - \frac{t^3}{6} + \frac{t^6}{180} - \frac{t^9}{12960} + \cdots & t - \frac{t^4}{12} + \frac{t^7}{504} - \cdots \\ -\frac{t^2}{2} + \frac{t^5}{30} - \frac{t^8}{1440} + \cdots & 1 - \frac{t^3}{3} + \frac{t^6}{72} - \frac{t^9}{4536} + \cdots \end{pmatrix}.$$

Also in this example, we observe that the fundamental matrices computed with the different methods are almost identical up to some differences in the coefficients.

4 Comparison of Magnus Expansion and Picard Iteration

In the examples with the 2×2 rotation matrix and Airy's equation, we observed that the Magnus Expansion and the Picard iterative method nearly gave the same approximations to the fundamental matrix. The small differences may be due to the use of approximations of $\Omega(t)$. The approximation of the fundamental matrix is simpler using a numerical method, for example the Picard iterative method, than using the Magnus Expansion, since there are less computations needed. Furthermore, the terms $\Omega_k(t)$ are increasingly complex, therefore higher terms of the Magnus Expansion are complicated to compute. However, using a numerical method such as Picard iteration, one can lose important qualitative properties. This motivates the study of the Magnus Expansion.

If we suppose that the matrix $A(t)$ belongs to some Lie algebra $\mathfrak{g}$ for all t, then also any sum of multiple integrals of nested commutators belongs to that algebra $\mathfrak{g}$. Hence, since all the terms in the Magnus Expansion have a structure similar to the first terms of the expansion, $\Omega(t)$ and any approximation to it obtained by truncation of the Magnus Expansion, will also belong to that same Lie algebra [6]. This is important for physical applications, since many properties of evolution operators are related to the fact that they belong to a certain Lie group. For instance, in quantum mechanics, the time-evolution operator belongs to the unitary group, which is the group of complex matrices A of which its conjugate transpose is equal its inverse, that is $A^\dagger = A^{-1}$.

Consider the one-dimensional time-dependent Schrödinger equation ($\hbar = 1$) given by

$$i\hbar \frac{\partial U(t,t_0)}{\partial t} = H(t)U(t,t_0), \qquad U(t_0,t_0) = I. \tag{22}$$

Here $H(t)$ is the Hamiltonian of the system which is Hermitian[1]. The Schrödinger equation is still valid when $H(t)$ is time-dependent. The time-evolution operator $U(t,t_0)$ belongs to the unitary group. When one would use a more familiar approximation scheme, for example the time-dependent perturbation theory (TDPT), to compute the solution of (22), the unitary condition of $U(t,t_0)$ is lost [11]. However, the Magnus Expansion can be used to compute approximations to the Schrödinger equation (22) such that the evolution operator stays in the unitary group.

For the sake of simplicity, we redefine the Hamiltonian as follows

$$\bar{H}(t) = -iH(t)/\hbar. \tag{23}$$

Since $H(t)$ is Hermitian, the redefined Hamiltonian $\bar{H}(t)$ is anti-Hermitian[2]. Note that when two operators A and B are anti-Hermitian, then their commutator, given by

$$[A,B] = AB - BA,$$

[1] A matrix A is Hermitian if it is equal to its conjugate transpose, that is $A^\dagger = A$.

[2] A matrix A is anti-Hermitian, or skew-Hermitian, if its conjugate transpose is equal to its negative, that is $A^\dagger = -A$.

is also anti-Hermitian [11]. Using the redefined Hamiltonian (23), the Schrödinger equation can be rewritten as

$$\frac{\partial}{\partial t}U(t,t_0) = \bar{H}(t)U(t,t_0), \qquad U(t_0,t_0) = I. \tag{24}$$

According to Magnus' proposal, there exists a solution for this inital value problem of the form

$$U(t,t_0) = e^{\Omega(t,t_0)}. \tag{25}$$

Any unitary matrix is the exponential of an anti-Hermitian matrix [11], so to ensure the unitarity of $U(t,t_0)$, the matrix function $\Omega(t,t_0)$ has to be anti-Hermitian.

The function $\Omega(t,t_0)$ is constructed as the series expansion

$$\Omega(t,t_0) = \sum_{k=0}^{\infty} \Omega_k(t,t_0),$$

of which the lowest order terms are given by

$$\begin{aligned}
\Omega_1(t,t_0) &= \frac{1}{t}\int_0^t \bar{H}(t_1)dt_1 \\
\Omega_2(t,t_0) &= -\frac{1}{2t}\int_0^t\int_0^{t_1} [\bar{H}(t_1),\bar{H}(t_2)]dt_2dt_1 \\
\Omega_3(t,t_0) &= \frac{1}{6t}\int_0^t\int_0^{t_1}\int_0^{t_2} \Big([\bar{H}(t_1),[\bar{H}(t_2),\bar{H}(t_3)]] + [[\bar{H}(t_1),\bar{H}(t_2)],\bar{H}(t_3)]\Big)dt_3dt_2dt_1.
\end{aligned}$$

These terms contain a sum of multiple integrals of nested commutators. The commutator of two anti-Hermitian operators is anti-Hermitian, so the function $\Omega(t,t_0)$ obtained by adding the terms $\Omega_k(t,t_0)$ is also anti-Hermitian. This ensures that the time-evolution operator $U(t,t_0)$ given by (25) is a unitary matrix. Hence, using the Magnus Expansion, the solution to the Schrödinger equation stays in the unitary group, while this would not be the case using a more familiar approximation scheme as the TDPT.

Therefore, although the Magnus Expansion is more complicated than numerical methods such as Picard iteration, it is of importance in certain disciplines where the evolution operator belongs to some Lie group.

Acknowledgements

I would like to thank my first supervisor dr. A.E. Sterk for his guidance and support during the writing of my bachelor thesis. He was always available for questions and helped me when I got stuck on a problem.

References

[1] Wolfgang Walter. Ordinary Differential Equations. *Graduate Texts in Mathematics, Springer,* 1991

[2] K. Erdmann and M.J. Wildon. Introduction to Lie Algebras. *Springer Undergraduate Mathematics Series,* 2006

[3] Jordan Schettler. Bernoulli Numbers. *Department of Mathematics, The University of Arizona*

[4] R.A. Silverman. Complex Analysis with Applications. *Dover Publications,* 1974

[5] Wilhelm Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics,* VII:649-673, 1954

[6] S. Blanes, F. Casas, J.A. Oteo, J. Ros. The Magnus expansion and some of its applications. *Elsevier Physics Reports,* Volume 470: pages 151-238, 2008

[7] J.J. O'Connor and E.F. Robertson. http://www-history.mcs.st-andrews.ac.uk/Biographies/Magnus.html. *School of Mathematics and Statistics University of St Andrews, Scotland*, 1997

[8] Tim Carlson. Magnus' Expansion as an approximation tool for ODEs. *The University of Alaska Fairbanks,* 2005

[9] M.L. Abell, J.P. Braselton. Introductory Differential Equations. *Elsevier,* 2014

[10] Wolfram Research, Inc. Mathematica Version 11.0 *Wolfram Research, Inc,* 2016

[11] S. Blanes, F. Casas, J.A. Oteo, J. Ros. A pedagogical approach to the Magnus expansion. *European Journal of Physics,* Volume 31: pages 907-918, 2010

[12] MIT OpenCourseWare. Quantum Theory of Radiation Interactions, http://ocw.mit.edu, 2012