Department of Mathematical Sciences

University of South Africa

Pretoria

Ordinary
Differential Equations

Only Study guide for **MAT3706**

Author

Prof M Grobbelaar

Revised by

Dr SA van Aardt

© 2013 University of South Africa

All rights reserved

Printed and published by the
University of South Africa
Muckleneuk, Pretoria

Page-layout by the Department
Mathematical Sciences

MAT3706/1/2014

70111413

Contents

PREFACE		iv
CHAPTER 1	Linear Systems of Differential Equations Solution of Systems of Differential Equations by the Method of Elimination	1
CHAPTER 2	Eigenvalues and Eigenvectors and Systems of Linear Equations with Constant Coefficients	23
CHAPTER 3	Generalised Eigenvectors (Root Vectors) and Systems of Linear Differential Equations	40
CHAPTER 4	Fundamental Matrices Non–homogeneous Systems The Inequality of Gronwall	51
CHAPTER 5	Higher Order One–dimensional Equations as Systems of First Order Equations	73
CHAPTER 6	Analytic Matrices and Power Series Solutions of Systems of Differential Equations	80
CHAPTER 7	Nonlinear Systems Existence and Uniqueness Theorem for Linear Systems	93
CHAPTER 8	Qualitative Theory of Differential Equations Stability of Solutions of Linear Systems Linearization of Nonlinear Systems	101
APPENDIX A	Symmetric matrices	103
APPENDIX B	Refresher notes on Methods of Solution of One-dimensional Differential Equations	104

Preface

The module MAT3706 is a continuation of the module APM2611 in which one–dimensional differential equations are treated. Obviously you should have a good knowledge of the contents of this module before starting to study MAT3706.

As you will notice, Chapter 8 of the study guide refers you to Chapter 10 of the prescribed book and there are no additional notes on this work in the study guide. Many exercises and examples in Chapters 1 – 7 of the study guide also refer to the prescribed book. Therefore, although the study guide is to a great extent complete, *it is essential that you use the prescribed book together with the study guide.* (Since the same book is also prescribed for APM2611, you should already have this book.)

You will notice that the study guide contains a large number of exercises. Many of these were taken from the prescribed book while others come from Goldberg and Schwartz (see References below). Do as many as you find necessary to master the tutorial matter.

The prescribed book by Zill & Wright are simply referred to as Z&W in the study guide.

In compiling Chapters 1 to 7, we have, apart from the prescribed book, made use of the following books:

FINIZIO, N & LADAS, G, (1989), *Ordinary Differential Equations with Modern Applications*, Wadsworth Publishing Company, Belmont, California.

GOLDBERG, J L & SCHWARTZ, A J, (1972), *Systems of Ordinary Differential Equations: An Introduction*, Harper & Row, New York.

RICE BERNARD J & STRANGE JERRY D, (1989), *Differential Equations with Applications*, The Benjamin/Cummings Publishing Company, Inc, Redwood City, California.

RAINVILLE, ED, (1981), *Elementary Differential Equations*, Collier Macmillan Publishers, London.

TENENBAUM, M & POLLARD, H, (1964), *Ordinary Differential Equations*, Harper & Row, New York.

CHAPTER 1

Linear Systems of Differential Equations
Solution of Systems of Differential Equations by the Method of Elimination

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding systems of linear differential equations (DE's):

- polynomial differential operators;
- matrix notation of linear systems of DE's;
- degeneracy of systems;
- homogeneous and non–homogeneous systems;
- conditions for the existence of a solution of a system of DE's;
- conditions for the uniqueness of the solution of initial value problems;
- linear independency of solutions of systems of DE's.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- write a system of DE's in operator notation;
- determine whether a system is degenerate or not;
- solve a system of DE's using the operator method (method of elimination);
- solve a system of DE's using the method of triangularization;
- determine the linear independency of the solutions;
- apply the solution techniques for systems of DE's in solving relevant physical problems.

1.1 INTRODUCTION

Mathematical models of real–life situations often lead to linear systems of differential equations. Linear systems of differential equations arise for instance in the theory of electric circuits, in Economics and in Ecology, the branch of science in which the interaction between different species is investigated. Systems of differential equations have been used by, among others, L.F. Richardson[1] in as early as 1939 in devising a mathematical model of arms races, and by E. Ackerman[1] and colleagues in the study of a mathematical model for the detection of diabetes. As we shall see later on (see Chapter 5) systems of differential equations also appear when an n–th order linear differential equation is reduced to a linear system of differential equations.

1.2 DEFINITIONS AND BASIC CONCEPTS

Definition 1.1 *By a system of n first–order linear differential equations we mean a set of n simultaneous equations of the form*

$$\begin{aligned}
\dot{x}_1(t) &= a_{11}(t)\,x_1 + a_{12}(t)\,x_2 + \ldots + a_{1n}(t)\,x_n + Q_1(t) \\
\dot{x}_2(t) &= a_{21}(t)\,x_1 + a_{22}(t)\,x_2 + \ldots + a_{2n}(t)\,x_n + Q_2(t) \\
&\;\;\vdots \\
\dot{x}_n(t) &= a_{n1}(t)\,x_1 + a_{n2}(t)\,x_2 + \ldots + a_{nn}(t)\,x_n + Q_n(t)
\end{aligned} \tag{1.1}$$

where the functions $a_{ij},\ i,j = 1,2,\ldots,n$ and the functions $Q_k,\ k = 1,2,\ldots,n,$ are all given functions of t on some interval J. We assume that all the functions a_{ij} and all the functions Q_k are continuous on J.

The functions $a_{ij}(t)$ and $Q_k(t)$ are known functions, while $x_1(t),\ x_2(t),\ldots,x_n(t)$ are the unknowns. The functions $a_{ij}(t)$ are called the *coefficients* of the linear system (1.1). We say that the system (1.1) has *constant coefficients* when each a_{ij} is a constant function. The terms $Q_k(t)$ are called *forcing terms.* The system (1.1) is called *homogeneous* when $Q_k(t)$ is zero for all $k = 1,2,\ldots,n$ and *non–homogeneous* (*inhomogeneous*) if at least one $Q_k(t)$ is not zero.

Remark

The homogeneous system obtained by putting $Q_k(t) = 0$ for all $k = 1,2,\ldots,n,$

$$\begin{aligned}
\dot{x}_1(t) &= a_{11}(t)\,x_1 + a_{12}(t)\,x_2 + \ldots + a_{1n}(t)\,x_n \\
\dot{x}_2(t) &= a_{21}(t)\,x_1 + a_{22}(t)\,x_2 + \ldots + a_{2n}(t)\,x_n \\
&\;\;\vdots \\
\dot{x}_n(t) &= a_{n1}(t)\,x_1 + a_{n2}(t)\,x_2 + \ldots + a_{nn}(t)\,x_n
\end{aligned} \tag{1.2}$$

may be expressed in the form

$$\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X} \tag{1.3}$$

with $\mathbf{A}(t) = [a_{ij}(t)],\quad i,j = 1,2,\ldots,n,\quad \mathbf{X} = [x_1, x_2, \ldots, x_n]^T$, the transpose of $[x_1, x_2, \ldots, x_n]$ and $\dot{\mathbf{X}} = [\dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n]^T$. If all the functions a_{ij} are constants, (1.3) becomes $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$.

[1]See pp. 168–169 of *Ordinary Differential Equations with Modern Applications* by N. Finizio and G. Ladas, Wadsworth Publishing Company, Belmont, California, 1989, for full references.

It is clear that a solution of this system (see Definition 1.3) should be an n–dimensional *vector.* Use your knowledge of matrix multiplication to show that the product of the matrix $\mathbf{A}(t)$ and the column vector $\mathbf{X}$ actually produces the vector

$$\begin{bmatrix} a_{11}(t)\,x_1 + a_{12}(t)\,x_2 + \ldots + a_{1n}(t)\,x_n \\ a_{21}(t)\,x_1 + a_{22}(t)\,x_2 + \ldots + a_{2n}(t)\,x_n \\ \vdots \\ a_{n1}(t)\,x_1 + a_{n2}(t)\,x_2 + \ldots + a_{nn}(t)\,x_n \end{bmatrix}.$$

Definition 1.2 *A* **higher–order** *system with* **constant coefficients** *is a system of the form*

$$\begin{aligned} P_{11}(D)\,x_1 + P_{12}(D)\,x_2 + \ldots + P_{1n}(D)\,x_n &= h_1(t) \\ P_{21}(D)\,x_1 + P_{22}(D)\,x_2 + \ldots + P_{2n}(D)\,x_n &= h_2(t) \\ &\vdots \\ P_{n1}(D)\,x_1 + P_{n2}(D)\,x_2 + \ldots + P_{nn}(D)\,x_n &= h_n(t) \end{aligned} \tag{1.4}$$

where all the $P_{ij},\ i,j = 1,2,\ldots,n$ *are* **polynomial differential operators** *and all the functions* $h_i(t)$*,* $i = 1,2,\ldots,n$ *are defined on an interval* J*. The determinant*

$$\det \begin{bmatrix} P_{11}(D) & \cdots & P_{1n}(D) \\ P_{21}(D) & \cdots & P_{2n}(D) \\ & \vdots & \\ P_{n1}(D) & \cdots & P_{nn}(D) \end{bmatrix}$$

is called the **determinant of the system** (1.4)*. The system* (1.4) *is* **non–degenerate** *if its determinant is non–zero, otherwise it is* **degenerate**.

The determinant of a system of differential equations is of great importance, since it provides a method of determining the correct number of arbitrary constants in the solution of the system (see Theorem 1.5).

Definition 1.3 *A* **solution** *of the system* (1.1) *of differential equations is an n–dimensional vector function*

$$\left[x_1(t),\, x_2(t),\ldots,x_n(t)\right]^T,$$

of which each component is defined and differentiable on the interval J and which, when substituted into the equations in (1.1)*, satisfies the equations for all t in J.*

As in the case of ordinary differential equations, we have

Theorem 1.4 (**The Superposition Principle**) *Any linear combination of solutions of a system of differential equations is also a solution of the system.*

For example, each of the two vectors $\left[-2e^{2t},\, e^{2t}\right]^T$ and $\left[e^{3t}, -e^{3t}\right]^T$ is a solution of the system

$$\begin{aligned} \dot{x}_1 &= x_1 - 2x_2 \\ \dot{x}_2 &= x_1 + 4x_2. \end{aligned}$$

By the superposition principle, the linear combination $\left[-2c_1e^{2t}+c_2e^{3t}, c_1e^{2t}-c_2e^{3t}\right]^T$, with c_1, c_2 arbitrary constants, is also a solution of the system.

Before defining the concept of a general solution of a system of differential equations, we need the concept of linear independence of vector functions, which is obviously an extension of the concept of linearly independent functions.

Definition 1.5 *Suppose the vector functions*

$$\mathbf{X}_1(t)=\begin{bmatrix} x_{11}(t)\\ x_{21}(t)\\ \vdots\\ x_{n1}(t)\end{bmatrix},\ \mathbf{X}_2(t)=\begin{bmatrix} x_{12}(t)\\ x_{22}(t)\\ \vdots\\ x_{n2}(t)\end{bmatrix},\ldots,\ \mathbf{X}_n(t)=\begin{bmatrix} x_{1n}(t)\\ x_{2n}(t)\\ \vdots\\ x_{nn}(t)\end{bmatrix}$$

are solutions, defined on an interval J, of a system of n first order linear differential equations in n unknowns. We say that $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$ are **linearly independent** *if it follows from*

$$c_1\mathbf{X}_1+c_2\mathbf{X}_2+\ldots+c_n\mathbf{X}_n=\mathbf{0} \text{ for all } t \text{ in } J$$

that

$$c_1=c_2=\ldots=c_n=0.$$

The following theorem provides a criterion which can be used to check the linear dependence of solutions of a system of differential equations:

Theorem 1.6 *The n solutions*

$$\begin{bmatrix} x_{11}(t)\\ x_{21}(t)\\ \vdots\\ x_{n1}(t)\end{bmatrix},\ \begin{bmatrix} x_{12}(t)\\ x_{22}(t)\\ \vdots\\ x_{n2}(t)\end{bmatrix},\ldots,\begin{bmatrix} x_{1n}(t)\\ x_{2n}(t)\\ \vdots\\ x_{nn}(t)\end{bmatrix}$$

of the system (1.1) *are linearly independent on the interval J if and only if the* **Wronskian determinant**

$$\det\begin{bmatrix} x_{11}(t) & \ldots & x_{1n}(t)\\ x_{21}(t) & \ldots & x_{2n}(t)\\ & \vdots & \\ x_{n1}(t) & \ldots & x_{nn}(t)\end{bmatrix}$$

of the system (1.2) *is never zero on J.*

We state without proof the following basic existence theorems for linear systems of differential equations:

Theorem 1.7 (**Existence of a solution of a homogeneous system**) *There exists n* **linearly independent solutions** *of the system* (1.2).

Furthermore, if the n column vectors

$$\begin{bmatrix} x_{11}(t)\\ x_{21}(t)\\ \vdots\\ x_{n1}(t)\end{bmatrix},\ \begin{bmatrix} x_{12}(t)\\ x_{22}(t)\\ \vdots\\ x_{n2}(t)\end{bmatrix},\ldots,\begin{bmatrix} x_{1n}(t)\\ x_{2n}(t)\\ \vdots\\ x_{nn}(t)\end{bmatrix}$$

are linearly independent solutions of the linear system (1.2)*, a* **general solution** *is given by*

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = c_1 \begin{bmatrix} x_{11}(t) \\ x_{21}(t) \\ \vdots \\ x_{n1}(t) \end{bmatrix} + c_2 \begin{bmatrix} x_{12}(t) \\ x_{22}(t) \\ \vdots \\ x_{n2}(t) \end{bmatrix} + \ldots + c_n \begin{bmatrix} x_{1n}(t) \\ x_{2n}(t) \\ \vdots \\ x_{nn}(t) \end{bmatrix},$$

i.e.

$$\begin{aligned} x_1(t) &= c_1x_{11}(t) + c_2x_{12}(t) + \ldots + c_nx_{1n}(t) \\ x_2(t) &= c_1x_{21}(t) + c_2x_{22}(t) + \ldots + c_nx_{2n}(t) \\ &\vdots \\ x_n(t) &= c_1x_{n1}(t) + c_2x_{n2}(t) + \ldots + c_nx_{nn}(t), \end{aligned}$$

where $c_1, c_2, \ldots, c_n$ *are arbitrary constants.*

Theorem 1.8 (**Existence of a solution of a non–homogeneous system**) *If the* n *solutions*

$$\begin{bmatrix} x_{11}(t) \\ x_{21}(t) \\ \vdots \\ x_{n1}(t) \end{bmatrix}, \begin{bmatrix} x_{12}(t) \\ x_{22}(t) \\ \vdots \\ x_{n2}(t) \end{bmatrix}, \ldots, \begin{bmatrix} x_{1n}(t) \\ x_{2n}(t) \\ \vdots \\ x_{nn}(t) \end{bmatrix}$$

of the system (1.2) *are linearly independent and if* $[x_{1p}(t), x_{2p}(t), \ldots, x_{np}(t)]^T$ *is a* **particular** *solution of the* **inhomogeneous system** (1.1)*, then a* **general solution** *of* (1.1) *is given by*

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = c_1 \begin{bmatrix} x_{11}(t) \\ x_{21}(t) \\ \vdots \\ x_{n1}(t) \end{bmatrix} + c_2 \begin{bmatrix} x_{12}(t) \\ x_{22}(t) \\ \vdots \\ x_{n2}(t) \end{bmatrix} + \ldots + c_n \begin{bmatrix} x_{1n}(t) \\ x_{2n}(t) \\ \vdots \\ x_{nn}(t) \end{bmatrix} + \begin{bmatrix} x_{1p}(t) \\ x_{2p}(t) \\ \vdots \\ x_{np}(t) \end{bmatrix}$$

i.e.

$$\begin{aligned} x_1(t) &= c_1x_{11}(t) + c_2x_{12}(t) + \ldots + c_nx_{1n}(t) + x_{1p}(t) \\ x_2(t) &= c_1x_{21}(t) + c_2x_{22}(t) + \ldots + c_nx_{2n}(t) + x_{2p}(t) \\ &\vdots \\ x_n(t) &= c_1x_{n1}(t) + c_2x_{n2}(t) + \ldots + c_nx_{nn}(t) + x_{np}(t) \end{aligned}$$

where $c_1, c_2, \ldots, c_n$ *are arbitrary constants.*

Remark

In Theorem 1.7 we learn that a general solution of a system of n first–order linear differential equations in n unknown functions contains n arbitrary constants. As we shall see in the next section the method of elimination sometimes introduces *redundant constants.* The redundant constants can be eliminated by substituting the solutions into the system and equating coefficients of similar terms.

In general the following theorem determines the correct number of arbitrary constants in a general solution of a system of differential equations (which may be a higher–order system):

Theorem 1.9 *The correct number of arbitrary constants in a general solution of a system of differential equations is equal to the order of the determinant of the system, provided this determinant has non–zero value.*

In particular: The number of arbitrary constants in a general solution of the higher–order system

$$\begin{aligned} P_{11}(D)\,x_1 + P_{12}(D)\,x_2 + \ldots + P_{1n}(D)\,x_n &= h_1(t) \\ P_{21}(D)\,x_1 + P_{22}(D)\,x_2 + \ldots + P_{2n}(D)\,x_n &= h_2(t) \\ &\vdots \\ P_{n1}(D)\,x_1 + P_{n2}(D)\,x_2 + \ldots + P_{nn}(D)\,x_n &= h_n(t) \end{aligned} \tag{1.5}$$

is equal to the degree of

$$\Delta = \det \begin{bmatrix} P_{11}(D) & \ldots & P_{1n}(D) \\ P_{21}(D) & \ldots & P_{2n}(D) \\ & \vdots & \\ P_{n1}(D) & \ldots & P_{nn}(D) \end{bmatrix}. \tag{1.6}$$

If Δ is identically zero, then the system either has either infinitely many solutions, or no solutions.

We will return to the proof of the latter theorem in Section 1.4, where we will also discuss the degenerate case when $\Delta = 0$ in order to be able to decide when a degenerate system has an infinite number of solutions or no solution. We end this section with the Existence and Uniqueness Theorem for a linear system with initial conditions.

Theorem 1.10 (**Existence and Uniqueness Theorem for Linear Initial Value Problems**) *Assume that the coefficients a_{ij}, $i, j = 1, 2, \ldots, n$ and the functions $Q_k(t)$, $k = 1, 2, \ldots, n$ of the system* (1.1) *are all continuous on the interval J. Let t_0 be a point in J and let x_{10}, $x_{20}, \ldots, x_{n0}$, be n given constants. Then the initial value problem* (*IVP*) *consisting of the system* (1.1) *and the initial conditions*

$$x_1(t_0) = x_{10},\ x_2(t_0) = x_{20}, \ldots, x_n(t_0) = x_{n0}$$

has a unique solution $[x_1(t), x_2(t), \ldots, x_n(t)]^T$. Furthermore this unique solution is valid throughout the interval J.

We are now ready to treat the method of elimination (operator method).

1.3 THE METHOD OF ELIMINATION

In this section we show a method for solving *non–degenerate* systems of differential equations which resembles the method of elimination of a variable used for solving systems of algebraic equations. Since the method entails the application of differential operators to differential equations, the method is also known as the operator method.

Before carrying on with this section, you should revise Sections 4.1 - 4.5 of Z&W which were done in APM2611 as well as Section 4.9 of Z&W where the notion of differential operators are introduced and discussed. Also read through Appendix B at the end of this study guide.

1.3.1 POLYNOMIAL OPERATORS

First we define:

Definition 1.11 *An nth–order* **polynomial operator** $P(D)$ *is a linear combination of differential operators of the form*

$$a_n D^n + a_{n-1} D^{n-1} + \ldots + a_1 D + a_0$$

where $a_0, a_1, \ldots, a_n$ *are constants.*

The symbol $P(D)$ indicates that $P(D)$ is applied to an n–times differentiable function y. To emphasize the fact that $P(D)$ is an *operator*, it is actually preferable (but not compulsory) to write $P(D)[y]$.

Definition 1.12 (**Linearity property**) *The polynomial differential operator* $P(D)$ *is a* **linear operator** *in view of the fact that the following linearity properties are satisfied:*

$$\begin{aligned} P(D)[y_1 + y_2] &= P(D)[y_1] + P(D)[y_2], \\ P(D)[cy] &= cP(D)[y]. \end{aligned}$$

Definition 1.13 (**Sum and Product of polynomial differential operators**) *The* **sum** $P_1(D)+P_2(D)$ *of two operators,* $P_1(D)$ *and* $P_2(D)$ *is obtained by writing* P_1 *and* P_2 *as linear combinations of the* D *operator and adding coefficients of equal powers of* D.

The **product** $P_1(D)P_2(D)$ *of two operators* $P_1(D)$ *and* $P_2(D)$ *is obtained by applying the operator* $P_2(D)$ *and then applying the operator* $P_1(D)$.

This is interpreted as follows:

$$(P_1(D)P_2(D))[y] = P_1(D)[P_2(D)[y]].$$

Remark

Polynomial operators can be shown to satisfy all the laws of elementary Algebra with regards to the basic operations like addition and multiplication. They may therefore be handled in the same way as algebraic polynomials. We can, for instance, factorize polynomial differential operators by using methods similar to those used in factorizing algebraic polynomials.

Example 1.14

(a) If $P_1(D) = D^3 + D - 8$, $P_2(D) = 3D^2 - 5D + 1$, then $P_1(D) + P_2(D) = D^3 + 3D^2 - 4D - 7$.

(b) The product operator $(4D-1)\left(D^2+2\right)$ may be expanded to yield $4D^3 - D^2 + 8D - 2$.

(c) The operator $D^3 - 2D^2 - 15D$ may be factorized into the factors $D(D-5)(D+3)$.

(d) Any linear system of differential equations may be written in the form (1.5). For example, the system

$$\begin{aligned} \frac{d^2x}{dt^2} - 4x + \frac{dy}{dt} &= 0 \\ -4\frac{dx}{dt} + \frac{d^2y}{dt^2} + 2y &= 0 \end{aligned} \tag{1.7}$$

can be written in operator notation as

$$\begin{aligned} \left(D^2 - 4\right)[x] + D[y] &= 0 & (1) \\ -4D[x] + \left(D^2 + 2\right)[y] &= 0. & (2) \end{aligned}$$

We are now ready to do:

Example 1.15 Solve the system (1.7) by using the elimination method (operator method). Eliminate x first.

Solution
To eliminate x first, we apply the operator $4D$ to the first equation and the operator D^2-4 to the second equation of the system. For brevity's sake, we denote these operations by $4D\,[(1)]$ and $\left(D^2-4\right)[(2)]$. We thus have

$$4\left(D^3-4D\right)[x]+4D^2\,[y]=0 \qquad (3)$$
$$-4\left(D^3-4D\right)[x]+\left(D^4-2D^2-8\right)[y]=0. \qquad (4)$$

We now add the last two equations to obtain

$$\left(D^4+2D^2-8\right)[y]=0. \qquad (5)$$

From the auxiliary equation

$$m^4+2m^2-8=\left(m^2+4\right)\left(m^2-2\right)=0$$

we find

$$m^2=-4 \;\text{ or }\; m^2=2$$

so that

$$m=\pm 2i \;\text{ or }\; m=\pm\sqrt{2}.$$

The solution of equation (5) is therefore

$$y\,(t)=c_1e^{\sqrt{2}t}+c_2e^{-\sqrt{2}t}+c_3\cos 2t+c_4\sin 2t,$$

where $c_1,\ldots,c_4$ are arbitrary constants.
Substitution into (2) yields

$$-4D\,[x]+2c_1e^{\sqrt{2}t}+2c_2e^{-\sqrt{2}t}-4c_3\cos 2t-4c_4\sin 2t+2c_1e^{\sqrt{2}t}+2c_2e^{-\sqrt{2}t}+2c_3\cos 2t+2c_4\sin 2t=0$$

whence

$$D\,[x]=c_1e^{\sqrt{2}t}+c_2e^{-\sqrt{2}t}-\frac{1}{2}c_3\cos 2t-\frac{1}{2}c_4\sin 2t.$$

Integration yields

$$x\,(t)=\frac{\sqrt{2}}{2}c_1e^{\sqrt{2}t}-\frac{\sqrt{2}}{2}c_2e^{-\sqrt{2}t}-\frac{1}{4}c_3\sin 2t+\frac{1}{4}c_4\cos 2t+c_5,$$

with $c_i,\ i=1,2,\ldots,5$ arbitrary constants.

The solution, therefore, contains five arbitrary constants whereas the determinant

$$\det\begin{bmatrix} D^2-4 & D \\ -4D & D^2+2 \end{bmatrix}=\left(D^2-4\right)\left(D^2+2\right)+4D^2$$

is of the fourth order. Substitution of the solution $(x\,(t)\,,y\,(t))$ into (1) yields $c_5=0$. Consequently, a general solution of the system is given by

$$\begin{aligned} x\,(t) &= \frac{\sqrt{2}}{2}c_1e^{\sqrt{2}t}-\frac{\sqrt{2}}{2}c_2e^{-\sqrt{2}t}-\frac{1}{4}c_3\sin 2t+c_4\cos 2t \\ y\,(t) &= c_2e^{\sqrt{2}t}+c_2e^{-\sqrt{2}t}+c_3\cos 2t+c_4\sin 2t. \end{aligned}$$

Example 1.16 Solve the system

$$\begin{aligned} (D+1)[x]-(D+1)[y] &= e^t && (1)\\ (D-1)[x]+(2D+1)[y] &= 5 && (2)\end{aligned}$$

by using the elimination method (operator method). Eliminate y first.

Solution
Eliminating y first we find:

$$\begin{aligned} (2D+1)[(1)] &: \left(2D^2+3D+1\right)[x]-\left(2D^2+3D+1\right)[y]=(2D+1)\left[e^t\right]=3e^t && (3)\\ (D+1)[(2)] &: \left(D^2-1\right)[x]+\left(2D^2+3D+1\right)[y]=(D+1)[5]=5 && (4)\end{aligned}$$

$$(3)+(4) \quad : \quad 3D(D+1)[x]=3e^t+5. \qquad (5)$$

From the auxiliary equation

$$m(m+1)=0$$

we have $m=0$ or $m=-1$ so that the complementary function is given by

$$x_{\text{C.F.}}(t)=c_1+c_2e^{-t}$$

with c_1 and c_2 arbitrary constants. We find the particular integral $x_{\text{P.I.}}(t)$ by using the method of undetermined coefficients:

Let

$$x_{\text{P.I.}}(t)=Ae^t+Bt$$

so that

$$\begin{aligned} \dot{x}_{\text{P.I.}}(t) &= Ae^t+B\\ \ddot{x}_{\text{P.I.}}(t) &= Ae^t.\end{aligned}$$

Substituting these equations in (5) we find

$$6Ae^t+3B=3e^t+5.$$

Comparing coefficients we find

$$A=\frac{1}{2} \text{ and } B=\frac{5}{3}$$

so that

$$x_{\text{P.I.}}(t)=\frac{1}{2}e^t+\frac{5}{3}t.$$

Therefore

$$\begin{aligned} x(t) &= c_{\text{C.F.}}(t)+x_{\text{P.I.}}(t)\\ &= c_1+c_2e^{-t}+\frac{1}{2}e^t+\frac{5}{3}t.\end{aligned}$$

Substituting $x(t)$ in (1) yields

$$\begin{aligned}(D+1)[y] &= (D+1)[x] - e^t \\ &= c_1 + \frac{5}{3}t + \frac{5}{3}.\end{aligned} \tag{6}$$

From the auxiliary equation $m+1=0$ it follows that the complementary function $y_{\text{C.F.}}(t)$ is given by

$$y_{\text{C.F.}}(t) = c_3 e^{-t}.$$

Using the method of undetermined coefficients we solve the particular integral $y_{\text{P.I.}}(t)$ as follows:
Let

$$y_{\text{P.I.}}(t) = At + B.$$

Then

$$\dot{y}_{\text{P.I.}}(t) = A.$$

Substitution of these equations in (6) yields

$$At + (A+B) = c_1 + \frac{5}{3}t + \frac{5}{3}.$$

Comparing coefficients we find that

$$A = \frac{5}{3} \quad \text{and} \quad B = c_1$$

so that

$$y_{\text{P.I.}}(t) = \frac{5}{3}t + c_1.$$

Therefore

$$\begin{aligned}y(t) &= y_{\text{C.F.}}(t) + y_{\text{P.I.}}(t) \\ &= c_3 e^{-t} + c_1 + \frac{5}{3}t.\end{aligned}$$

The solution therefore contains three arbitrary constants c_1, c_2 and c_3 whereas the determinant

$$\det \begin{bmatrix} D+1 & -D-1 \\ D-1 & 2D+1 \end{bmatrix}$$

is of the *second* order. Substitution of the solution $(x(t), y(t))$ into (2) yields $c_3 = -2c_2$. Consequently a general solution of the system is given by

$$\begin{aligned}x(t) &= c_1 + c_2 e^{-t} + \frac{1}{2}e^t + \frac{5}{3}t \\ y(t) &= c_1 - 2c_2 e^{-t} + \frac{5}{3}t.\end{aligned}$$

Exercise 1.17

(1) Find a general solution of each of the following systems:

(a) $\dot{x} = 5x - 6y + 1$
$\dot{y} = 6x - 7y + 1$

(b) $\dot{x} = 3x - 2y + 2t^2$
$\dot{y} = 5x + y - 1$

(c) $\ddot{x} - \dot{y} = t$
$\dot{x} + 3x + \dot{y} + 3y = 2.$

(2) Solve the following initial value problem:

$$\begin{aligned} \ddot{y} &= -y + \dot{y} + z \\ \dot{z} &= y - z - 1 \\ y(0) &= 1, \;\; \dot{y}(0) = 0, \;\; z(0) = 1. \end{aligned}$$

(3) Use the operator method (method of elimination) to find a general solution of each of the following systems of ordinary differential equations:

(a) $(3D + 2)[x] + (D - 6)[y] = 5e^t$
$(4D + 2)[x] + (D - 8)[y] = 5e^t + 2t - 3$

NB: Eliminate y first.

(b) $\dot{x} = 4x - \ddot{y} + t^2$
$\dot{y} = -x - \dot{x}$

NB: Eliminate x first.

1.3.2 EQUIVALENT TRIANGULAR SYSTEMS

As noted previously, the method of elimination sometimes introduces redundant arbitrary constants into a general solution of a system of differential equations, the reason being, as we have in fact seen in the above example, that a polynomial operator in D operates on each equation. Since the elimination of such constants is generally a tedious and time consuming process, we shall develop a method which immediately yields the correct number of constants in a general solution of a (non–degenerate) system of differential equations. The method amounts to a reduction of a system of differential equations to an *equivalent triangular system.*

Definition 1.18 *If the system*

$$\begin{aligned} P_{11}(D)\,x_1 + P_{12}(D)\,x_2 + \ldots + P_{1n}(D)\,x_n &= h_1(t) \\ P_{21}(D)\,x_1 + P_{22}(D)\,x_2 + \ldots + P_{2n}(D)\,x_n &= h_2(t) \\ &\vdots \\ P_{n1}(D)\,x_1 + P_{n2}(D)\,x_2 + \ldots + P_{nn}(D)\,x_n &= h_n(t) \end{aligned} \tag{1.8}$$

is reducible by a method analogous to that of row operations in Matrix Algebra to the system

$$\begin{aligned}
&P_{11}(D)\,x_1 = H_1(t)\\
&P_{21}(D)\,x_1 + P_{22}(D)\,x_2 = H_2(t)\\
&\qquad\qquad\vdots\\
&P_{n1}(D)\,x_1 + P_{n2}(D)\,x_2 + \ldots + P_{nn}(D)\,x_n = H_n(t),
\end{aligned} \tag{1.9}$$

then the system (1.9) *is known as a* **triangular system equivalent to** (1.8).

A linear system of differential equations is thus in triangular form if each succeeding equation in the system has at least one unknown function more (or less) than the previous equation.

The concept of triangular system thus corresponds to that of triangular matrix.

From the definition and from the properties of determinants, the following conclusion can immediately be drawn:

If a system of differential equations is reduced to an equivalent triangular system, the determinant and the order of the determinant of the system remain unchanged.

In Theorem 1.22 we will show that a solution of an equivalent triangular system is also a solution of the original system of differential equations. The proof contains the proof of Theorem 1.9 for $n = 2$.

In the next example we show how a linear system of differential equations can be reduced to an equivalent triangular system.

Example 1.19 Reduce the following system of equations to an equivalent triangular system

$$\begin{aligned}
\left(D^3 + 2D^2 - D + 1\right)[x] + \left(D^4 - 2D^2 + 2\right)[y] &= 0 \qquad (1)\\
(D - 3)[x] + \left(D^3 + 1\right)[y] &= 0 \qquad (2)
\end{aligned}$$

of the form

$$\begin{aligned}
P_1(D)[y] &= 0 \qquad (A)\\
P_2(D)[x] + P_3(D)[y] &= 0. \qquad (B)
\end{aligned}$$

Solution

Consider the polynomials in D, i.e. the polynomials $P(D)$; retain the equation in which the order of $P(D)$ is the *lowest* — in this case (2). Our aim is to reduce the order of the coefficient polynomial $P(D)$ of x in (1) step by step: first we get rid of the coefficient D^3 of x. The notation $(1) \to (1) + k(D)[(2)]$ denotes that (1) is replaced by $(1) + k(D)[(2)]$ with $k(D)$ a polynomial in D. By executing the operations as indicated, we obtain the following systems, of which the last is the required triangular system.

$$\begin{aligned}
(1) \to (1) - D^2[(2)] : \left(5D^2 - D + 1\right)[x] + \left(-D^5 + D^4 - 3D^2 + 2\right)[y] &= 0 \qquad (3)\\
(D - 3)[x] + \left(D^3 + 1\right)[y] &= 0. \qquad (2)
\end{aligned}$$

Note that the order of the coefficient polynomial $P(D)$ of x in (3) has been reduced by one; it is no longer a cubic polynomial, but a quadratic polynomial. *This is the basic difference between the operator method* *and the method of triangularization.* (If we were applying the operator method, we would have executed the operation

$$\left(D^3+1\right)[1]-\left(D^4-2D^2+2\right)[2]$$

to eliminate y, or the operation

$$(D-3)[1]-\left(D^3+2D^2-D+1\right)[2]$$

to eliminate x.)

Instead, we did not eliminate x, but only got rid of the coefficient D^3 of x. The next step would be to reduce the order of coefficient polynomial $P(D)$ of x in (3) (i.e. $5D^2-D+1$) by one which means that we have to get rid of the coefficient $5D^2$:

$$(3)\to(3)-5D\,[(2)]:(14D+1)[x]+\left(-D^5-4D^4-3D^2-5D+2\right)[y]=0 \qquad (4)$$
$$(D-3)[x]+\left(D^3+1\right)[y]=0. \qquad (2)$$

The order of the coefficient polynomial $P(D)$ of x in the two equations is now the same. Retain either of the two equations.

$$(4)-14\,(2):43x-\left(D^5+4D^4+14D^3+3D^2+5D+12\right)[y]=0 \qquad (5)$$
$$(D-3)[x]+\left(D^3+1\right)[y]=0. \qquad (2)$$

From (5) we get

$$x=\frac{1}{43}\left(D^5+4D^4+14D^3+3D^2+5D+12\right)[y].$$

Substitute into (2). This yields the system

$$43x-\left(D^5+4D^4+14D^3+3D^3+5D+12\right)[y]=0 \qquad (5)$$
$$\tfrac{D-3}{43}\left(D^5+4D^4+14D^3+3D^2+5D+12\right)[y]+\left(D^3+1\right)[y]=0. \qquad (6)$$

Simplification of (6) yields the triangular system

$$\left(D^6+D^5+2D^4+4D^3-4D^2-3D+7\right)[y]=0 \qquad (7)$$
$$43x-\left(D^5+4D^4+14D^3+3D^2+5D+12\right)[y]=0. \qquad (5)$$

Note that in the triangular system above, the polynomial coefficient of x in (5) is a constant.

Remark

An equivalent triangular system is not unique. The system

$$(D-2)[y_1]+2D[y_2]=2-4e^{2x}$$
$$(2D-3)[y_1]+(3D-1)[y_2]=0$$

can be reduced to two equivalent systems, namely

$$\left(D^2+D-2\right)[y_1]=2+20e^{2x}$$
$$D[y_1]-2y_2=-6+12e^{2x}$$

and

$$\left(D^2+D-2\right)[y_2]=-6-4e^{2x}$$
$$y_1-(D+1)[y_2]=-4+8e^{2x}.$$

Exercise 1.20 Derive the two triangular systems mentioned in the above remark. Use the method of elimination to solve these two systems as well as the original system of differential equations. What conclusion can you draw with regard to the solutions?

Example 1.21 Reduce the following system of equations to an equivalent triangular system

$$\begin{aligned} \ddot{x} + x + 4\dot{y} - 4y &= 4e^t \\ \dot{x} - x + \dot{y} + 9y &= 0 \end{aligned}$$

of the form

$$\begin{aligned} P_1(D)[y] &= f_1(t) \\ P_2(D)[x] + P_3(D)[y] &= f_2(t) \end{aligned}$$

and SOLVE.

Solution

Note that there are two triangular forms to which you can reduce the system: either

$$\left.\begin{aligned} P_1(D)[y] &= f_1(t) \\ P_2(D)[x] + P_3(D)[y] &= f_2(t) \end{aligned}\right\} A$$

or

$$\left.\begin{aligned} &P_1(D)[x] = f_1(t) \\ &P_2(D)[x] + P_3(D)[y] = f_2(t) \end{aligned}\right\} B$$

In this case you are asked to reduce the system to the first triangular form.

Also note that the polynomial $P_2(D)$ should be a constant in the triangular form A. This implies that once you have solved for $y(t)$ from the first equation in A, you will be able to solve for $x(t)$ immediately from the second equation in A. Therefore there won't be any redundant constants as in the case of the operator method. (In the case of the triangular form B, the polynomial $P_3(D)$ should be a constant so that $y(t)$ can be solved immediately once you have obtained $x(t)$ from the first equation.)

In operator notation the system becomes

$$\begin{aligned} (D^2+1)[x] + 4(D-1)[y] &= 4e^t & (1) \\ (D-1)[x] + (D+9)[y] &= 0. & (2) \end{aligned}$$

We retain the equation in which the order of $P(D)$ is the lowest, in this case (2). Furthermore, we execute the following operations in order to obtain the required triangular system:

$(1) \to (1) - D[(2)]$:

$$\begin{aligned} (D^2+1-D(D-1))[x] + (4(D-1) - D(D+9))[y] = (D+1)[x] - (D^2+5D+4)[y] &= 4e^t & (3) \\ (D-1)[x] + (D+9)[y] &= 0 & (2) \end{aligned}$$

$(3) \to (3) - (2)$:

$$\begin{aligned} (D+1-(D-1))[x] + (-D^2-5D-4-(D+9))[y] = 2x - (D^2-6D+13)[y] &= 4e^t & (4) \\ (D-1)[x] + (D+9)[y] &= 0. & (2) \end{aligned}$$

Substituting

$$x = 2e^t + \frac{1}{2}(D^2 - 6D + 13)[y] \qquad \text{(from(4))}$$

into (2) we find

$$(D-1)\left[2e^t + \frac{1}{2}(D^2-6D+13)[y]\right] + (D+9)[y] = 0.$$

The triangular system is therefore given by

$$\begin{aligned}(D^3+5D^2+9D+5)[y] &= 0 & (5)\\ 2x-(D^2-6D+13)[y] &= 4e^t & (4)\end{aligned}$$

We now solve for y using equation (5):
From the auxiliary equation

$$m^3+5m^2+9m+5 = (m+1)(m^2+4m+5) = 0,$$

we get

$$m=-1 \quad \text{and } m=-2\pm i,$$

so that

$$y(t) = y_{C.F.}(t) + y_{P.I.}(t) = e^{-2t}(A\cos t + B\sin t) + Ce^{-t}.$$

Substitution of $y(t)$ in (4) then gives

$$x(t) = 2e^t + e^{-2t}\left((2A+B)\cos t + (2B-A)\sin t\right) + 4Ce^{-t}.$$

We now proceed to the proof of Theorem 1.22. We formulate and prove the theorem only for a linear system consisting of two equations, since the proof is completely analogous for systems consisting of more equations.

Theorem 1.22 *Suppose that the system*

$$\begin{aligned}f_1(D)[x] + g_1(D)[y] &= h_1(t) & (1)\\ f_2(D)[x] + g_2(D)[y] &= h_2(t) & (2)\end{aligned} \qquad (1.10)$$

has been reduced to the equivalent system

$$\begin{aligned}F_1(D)[x] &= H_1(t) & (3)\\ F_2(D)[x] + G_2(D)[y] &= H_2(t)\,. & (4)\end{aligned} \qquad (1.11)$$

A general solution of (1.11) *will then be a general solution of* (1.10).

Proof. The proof is in the form of a discussion:
Suppose we solve equation (3) for x. The correct number of arbitrary constants in the solution is equal to the order of $F_1(D)$. Substitute this solution into equation (4), which is then reduced to an equation in y of which the order equals that of $G_2(D)$. The number of arbitrary constants in the solution of (4) equals the order of $G_2(D)$. Consequently the total number of constants in the solution (x, y) of (1.11) is equal to the sum of the orders of $F_1(D)$ and $G_2(D)$. This is, however, precisely the order of the determinant of (1.11), i.e. the order of $F_1(D)\,G_2(D)$. We have thus shown that, if we solve the system as described, the solution contains the correct number of constants, (since the pair satisfies each equation identically) and that this number equals the order of the determinant of (1.11). But the order of the determinant of the reduced system (1.11) is the same as the order of the determinant of the original system (1.10). Thus we have not only proved that a solution of (1.11) will also be a solution of (1.10) and contain the correct number of arbitrary constants, but have at the same time proved Theorem 1.9. ■

Exercise 1.23 Triangularize and solve the following system

$$(D^2+1)[x]-2D[y]=2t$$
$$(2D-1)[x]+(D-2)[y]=7.$$

NB: Reduce to the form

$$y=f_1(t)$$
$$P_2(D)[x]+P_3(D)[y]=f_2(t)$$

and SOLVE.

In all the examples that we have studied, the systems of differential equations were non–degenerate. We now study:

1.4 DEGENERATE SYSTEMS OF DIFFERENTIAL EQUATIONS

We recall that a system of differential equations is *degenerate* when its determinant is zero. As in the case of systems of algebraic equations, a degenerate system of differential equations has either no solution or otherwise an infinite number of solutions.

By systematizing the elimination procedure and letting it assume Cramer's rule, we will be able to decide when a degenerate system of differential equations has no solution or infinitely many. For simplicity's sake, we consider the system

$$\begin{aligned} P_{11}(D)[y_1]+P_{12}(D)[y_2]&=f_1(t)\\ P_{21}(D)[y_1]+P_{22}(D)[y_2]&=f_2(t)\,. \end{aligned} \tag{1.12}$$

By applying $P_{22}(D)$ to the first equation of (1.12) and $P_{12}(D)$ to the second equation of (1.12), and by subtracting the resulting equations, we obtain

$$(P_{11}(D)\,P_{22}(D)-P_{12}(D)\,P_{21}(D))\,[y_1]=P_{22}(D)\,[f_1(x)]-P_{12}(D)\,[f_2(x)]\,.$$

This can be written in the form

$$\det\begin{bmatrix} P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}[y_1]=\det\begin{bmatrix} f_1 & P_{12}\\ f_2 & P_{22}\end{bmatrix}, \tag{1.13}$$

where the determinant on the right hand side is interpreted to mean $P_{22}[f_1]-P_{12}[f_2]$.
Similarly we have

$$\det\begin{bmatrix} P_{11} & P_{12}\\ P_{21} & P_{22}\end{bmatrix}[y_2]=\det\begin{bmatrix} P_{11} & f_1\\ P_{21} & f_2\end{bmatrix}. \tag{1.14}$$

By using equations (1.13) and (1.14), we can make the last part of Theorem 1.9 precise for the case $n=2$, by stating:

If $\Delta=P_{11}P_{22}-P_{12}P_{21}=0$, *then the system* (1.12) *has infinitely many solutions if the determinants on the right hand sides of equations* (1.13) *and*(1.14) *vanish; otherwise there is no solution.*

The theorem can easily be generalized to hold for the case $n>3$.

We illustrate the validity of the theorem with a few examples:

Example 1.24 Find a general solution, if it exists, of the system

$$\begin{aligned} D[x] - D[y] &= e^t \\ 3D[x] - 3D[y] &= 3e^t. \end{aligned}$$

Solution
The system is clearly degenerate since

$$\Delta = \begin{vmatrix} D & -D \\ 3D & -3D \end{vmatrix} = D(-3D) - (-D)(3D) = 0.$$

Furthermore,

$$\begin{vmatrix} D & e^t \\ 3D & 3e^t \end{vmatrix} = D\left[3e^t\right] - 3D\left[e^t\right] = 3e^t - 3e^t = 0$$

and similarly

$$\begin{vmatrix} e^t & D \\ 3e^t & 3D \end{vmatrix} = 3D\left[e^t\right] - D\left[3e^t\right] = 0$$

so that an infinite number of solutions exist (note that the system is equivalent to a single equation in two unknowns). We could for instance choose

$$x(t) = e^t + c, \quad y(t) = c, \quad \text{or} \quad x(t) = t + e^t + c, \quad y(t) = t + c.$$

More generally, let $y(t)$ be any arbitrary function of t. Then

$$x(t) = y(t) + e^t + c$$

where c is an arbitrary constant.

Example 1.25 Find a general solution, if it exists, of the system

$$\begin{aligned} D[x] + 2D[y] &= e^t \\ D[x] + 2D[y] &= t. \end{aligned}$$

Solution
The system is degenerate since

$$\Delta = \begin{vmatrix} D & 2D \\ D & 2D \end{vmatrix} = D(2D) - 2D(D) = 0.$$

However, since

$$\begin{vmatrix} D & e^t \\ D & t \end{vmatrix} = D(t) - D\left(e^t\right) = 1 - e^t \neq 0,$$

the right hand side of the system does not become zero when x is eliminated. (In this case, the same is true when y is eliminated.) Hence the system has no solution. (This could also quite easily be seen by noting that, since the left hand sides of the two equations are identical, we must have that the right hand sides should also be equal, i.e. $e^t = t$ for all values of t!)

Exercise 1.26 In the following exercises determine if the given systems have no solutions or infinitely many solutions. If a system has infinitely many solutions, indicate how the set of solutions may be obtained.

(1)
$$\begin{aligned} \dot{y}_1 + \dot{y}_2 &= 3 \\ \ddot{y}_1 - \dot{y}_1 + \ddot{y}_2 - \dot{y}_2 &= x \end{aligned}$$

(2)
$$\begin{aligned} \dot{y}_1 - y_1 + \dot{y}_2 &= x \\ \ddot{y}_1 - \dot{y}_1 + \ddot{y}_2 &= 1 \end{aligned}$$

We conclude this chapter with a few applications.

1.5 APPLICATIONS[2]

As stated in Section 1.1, linear systems of differential equations are encountered in a large variety of physical problems. We will treat a problem which occurs in the study of electrical circuits, as well as a mixture problem, and finally an example in which the love and hate between two lovers are modelled (believe it if you can!) Many other applications of linear differential equations may be found in books on differential equations, a few of which will be mentioned at the end of this section.

Section 1.5.1 as well as Section 1.5.2 are adopted from Section 9–8 of Bernard J. Rice and Jerry D. Strange's book *Ordinary Differential Equations with Applications*, Second Edition, Brooks/Cole Publishing Company, Pacific Grove, California, 1989, since these sections effectively illustrate the use of systems of differential equations in modelling physical processes. Only numbers of equations and units (we use metric units) have been changed.

1.5.1 ELECTRICAL CIRCUITS

The current in the electrical circuit shown in Figure 1.1 can be found by applying the elements of circuit analysis (which you are, for the purpose of this module, not expected to be familiar with). Essential to the analysis is *Kirchoff's law*, which states that *the sum of the voltage drops around any closed loop is zero.* In applying Kirchoff's law, we use the fact that the voltage across an inductance is

$$v_L = L\frac{dI}{dt}$$

and the voltage across a resistance is $v_R = IR$.

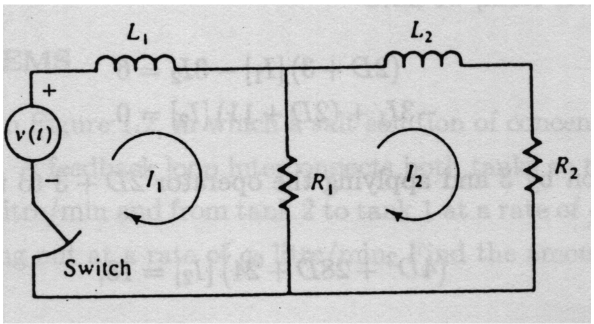

Figure 1.1: An electrical circuit

[2]Students are not expected to be able to derive the mathematical models encountered in this section.

Let the current in the left loop of the circuit be I_1, and the current in the right loop be I_2. From the figure we conclude that the current in the resistor R_1 is $I_1 - I_2$ relative to the left loop, and $I_2 - I_1$ relative to the right loop. Applying Kirchoff's law to the left loop, we get

$$L_1\dot{I}_1 + R_1\left(I_1 - I_2\right) = v\left(t\right) \tag{1.15}$$

where $\dot{I}_1 = \dfrac{dI_1}{dt}$. Similarly, the sum of the voltage drops around the right loop yields

$$L_2\dot{I}_2 + R_2 I_2 + R_1\left(I_2 - I_1\right) = 0. \tag{1.16}$$

If the components of the circuit are given, the values of I_1 and I_2 can be found by solving this system of differential equations.

Example 1.27 Consider the circuit shown in Figure 1.1. Determine I_1 and I_2 when the switch is closed if

$$\begin{aligned} L_1 &= L_2 = 2 \text{ henrys,} \\ R_1 &= 3 \text{ ohms,} \\ R_2 &= 8 \text{ ohms, and} \\ v\left(t\right) &= 6 \text{ volts.} \end{aligned}$$

Assume the initial current in the circuit is zero.

Solution
As previously noted, the circuit is described by the system

$$\begin{aligned} L_1\dot{I}_1 + R_1\left(I_1 - I_2\right) &= v\left(t\right) \\ L_2\dot{I}_2 + R_2 I_2 + R_1\left(I_2 - I_1\right) &= 0. \end{aligned}$$

Substituting the given values yields the initial value problem

$$\begin{aligned} 2\dot{I}_1 + 3I_1 - 3I_2 &= 6 \\ 2\dot{I}_1 + 11I_2 - 3I_1 &= 0 \\ I_1\left(0\right) = I_2\left(0\right) &= 0. \end{aligned}$$

Writing the system in operator form, we have

$$\begin{aligned} \left(2D+3\right)\left[I_1\right] - 3I_2 &= 6 \\ -3I_1 + \left(2D+11\right)\left[I_2\right] &= 0. \end{aligned}$$

Multiplying the first equation by 3 and applying the operator $2D + 3$ to the second and then adding the two equations, we obtain

$$\left(4D^2 + 28D + 24\right)\left[I_2\right] = 18,$$

or dividing by 4,

$$\left(D^2 + 7D + 6\right)\left[I_2\right] = \frac{9}{2}.$$

The solution of this second–order non–homogeneous linear differential equation consists of the sum of a general solution of the corresponding homogeneous equation and a particular solution of the given non–homogeneous equation. The solution of $\left(D^2 + 7D + 6\right)\left[I_2\right] = 0,$ is $c_1e^{-t} + c_2e^{-6t}$.

The method of undetermined coefficients can now be used to show that

$$I_{p2} = \frac{3}{4}$$

is a particular solution of the non–homogeneous equation. Thus the expression for I_2 is

$$I_2 = c_1 e^{-t} + c_2 e^{-6t} + \frac{3}{4}.$$

To find an expression for I_1, we substitute I_2 into the second equation of the system:

$$2\left(-c_1 e^{-t} - 6c_2 e^{-6t}\right) + 11\left(c_1 e^{-t} + c_2 e^{-6t} + \frac{3}{4}\right) - 3I_1 = 0.$$

Solving this equation for I_1 yields

$$I_1 = 3c_1 e^{-t} - \frac{1}{3}c_2 e^{-6t} + \frac{11}{4}.$$

Finally, using $I_1(0) = 0$ in the expressions for I_1 and I_2, we get

$$0 = 3c_1 - \frac{1}{3}c_2 + \frac{11}{4}, \qquad (\mathrm{I}_1)$$

$$0 = c_1 + c_2 + \frac{3}{4}. \qquad (\mathrm{I}_2)$$

Solving this system, we get $c_1 = -\dfrac{9}{10}$ and $c_2 = \dfrac{3}{20}$, so the desired currents are

$$I_1 = -\frac{27}{10}e^{-t} - \frac{1}{20}e^{-6t} + \frac{11}{4},$$

$$I_2 = -\frac{9}{10}e^{-t} + \frac{3}{20}e^{-6t} + \frac{3}{4}.$$

Exercise 1.28 This exercise is Exercise 3 from Section 9–8 of the quoted book by Rice and Strange. Referring to Figure 1.1, we assume that

$$L_1 = L_2 = 1, \quad R_1 = 6, \quad R = 1, \quad \text{and} \quad R_2 = 3.$$

Prior to $t = 0$, the generator has an output of $6v$ which at $t = 0$ is increased to $12v$. Determine the initial current and the current for $t > 0$.

1.5.2 MIXTURE PROBLEMS

Consider the two tanks shown in Figure 1.2, in which a salt solution of concentration c_i kg/litre is pumped into tank 1 at a rate g_i kg/min. A feedback loop interconnects both tanks so that solution is pumped from tank 1 to tank 2 at a rate of g_1 litre/min and from tank 2 to tank 1 at a rate of g_2 litre/min. Simultaneously, the solution in tank 2 is draining out at a rate of g_0 litre/min. Find the amount of salt in each tank.
Let y_1 and y_2 represent the amount of salt (in kilograms) in tanks 1 and 2 at any time and G_1 and G_2 represent the amount of liquid in the tank at any time t. Then the concentration in each tank is given by

$$c_1 = \frac{y_1}{G_1} = \text{concentration of salt in tank 1}$$

$$c_2 = \frac{y_2}{G_2} = \text{concentration of salt in tank 2.}$$

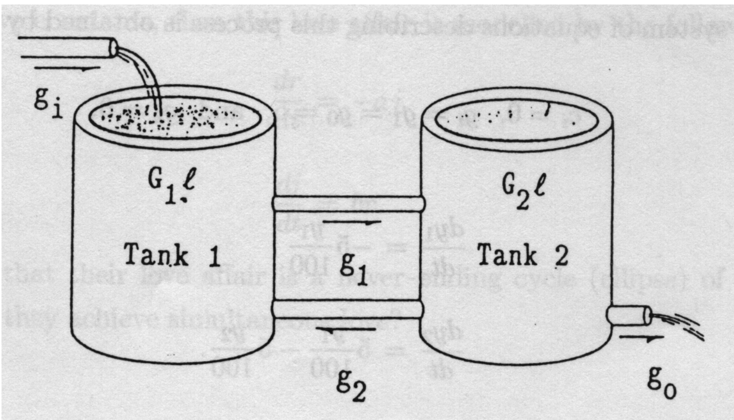

Figure 1.2: Flow between two tanks

The rate at which the salt is changing in tank 1 is

$$\begin{aligned}\frac{dy_1}{dt} &= c_i g_i - c_1 g_1 + c_2 g_2 \\ &= c_i g_i - \frac{y_1}{G_1} g_1 + \frac{y_2}{G_2} g_2 \qquad (1.17)\end{aligned}$$

Also, in tank 2, we have

$$\frac{dy_2}{dt} = \frac{y_1}{G_1} g_1 - \frac{y_2}{G_2} g_2 - \frac{y_2}{G_2} g_0. \qquad (1.18)$$

Equations (1.17) and (1.18) form a system in y_1 and y_2. If G_{10} and G_{20} are the initial volumes of the two tanks, then

$$\begin{aligned}G_1 &= G_{10} + g_i t - g_1 t + g_2 t \\ G_2 &= G_{20} + g_0 t - g_2 t + g_1 t.\end{aligned}$$

Comment

In many mixture problems the rates are assumed to be balanced; that is

$$g_i = g_0, \quad g_i + g_2 = g_1; \text{ and } g_0 + g_2 = g_1.$$

In this problem we assume the volumes are constant, that is $G_1 = G_{10}$ and $G_2 = G_{20}$.

Example 1.29 Assume both tanks in Figure 1.2 are filled with 100 litres of salt solutions of concentrations c_1 and c_2, respectively. Pure water is pumped into tank 1 at the rate of 5 litres/min. The solution is thoroughly mixed and pumped into and out of tank 2 at the rate of 5 litres/min. Assume no solution is pumped from tank 2 to tank 1. The system of equations describing this process is obtained by using equations (1.17) and (1.18) with

$$c_i = 0, \;\; g_i = g_1 = g_0 = 5, \text{ and } g_2 = 0.$$

Thus

$$\begin{aligned}\frac{dy_1}{dt} &= -5\frac{y_1}{100} \\ \frac{dy_2}{dt} &= 5\frac{y_1}{100} - 5\frac{y_2}{100}.\end{aligned}$$

Exercise 1.30 The mixture problems below are Exercises 11–13 from Section 9–8, again from the cited book by Rice and Strange.

(1) Solve for the amount of salt at any time in two 200–litre tanks if the input is pure water and

$$g_i = g_0 = 5, \quad g_1 = 7, \quad and \quad g_2 = 2.$$

Assume $y_1(0) = y_2(0) = 0$.

(2) Solve for the amount of salt at any time in two 200–litre tanks if the input is a salt solution with a concentration of 0.5 kilogram/litre and $g_i = g_1 = g_0 = 12$ litres/minute. Assume $g_2 = 0$ and $y_1(0) = y_2(0) = 0$.

(3) Solve for the amount of salt at any time in two 400–litre tanks if the input is a salt solution with a concentration of 0.5 kilogram/litre,

$$g_i = g_0 = 5, \quad g_1 = 7, \quad \text{and} \quad g_2 = 2.$$

Assume $y_1(0) = y_2(0) = 0$.

Our final application on linear systems of differential equations is a rather humoristic one, just in case none of the above applications appeal to you.

1.5.3 LOVE AFFAIRS

This application is due to Steven H. Strogaz, "Love Affairs and Differential Equations", *Math Magazine* **61** (1988): 35.

Two lovers, Romeo and Juliet, are such that the more Juliet loves Romeo, the more he begins to dislike her. On the other hand, when Juliet's love for Romeo begins to taper off, the more his affection for her begins to grow. For Juliet's part, her love for Romeo grows when he loves her and dissipates when he dislikes her. Introduce $r(t)$ to represent Romeo's love/hate for Juliet at time t and $j(t)$ to represent Juliet's love/hate for Romeo at time t. For either function a positive value represents love and a negative value represents hate. If a and b are positive constants, then this love affair is modelled by the following system:

$$\frac{dr}{dt} = -aj,$$

$$\frac{dj}{dt} = br.$$

Solve this system. Show that their love affair is a never–ending cycle (ellipse) of love and hate. What percentage of the time do they achieve simultaneous love?

For more applications please read Section 3.3 of Z&W.

CHAPTER 2

Eigenvalues and Eigenvectors and Systems of Linear Equations with Constant Coefficients

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding the eigenvalue–eigenvector method of solving linear systems of DE's:

- the characteristic equation $C(\lambda) = 0$ of the matrix $\mathbf{A}$;
- finding eigenvalues and their corresponding eigenvectors;
- the linear independence of eigenvectors;
- complex eigenvalues and their corresponding eigenvectors;
- multiple roots of $C(\lambda) = 0$ – finding a second linearly independent eigenvector corresponding to a double root of $C(\lambda) = 0$;
- solving initial value problems using the eigenvalue–eigenvector method.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- find the eigenvalues of a matrix $\mathbf{A}$;
- determine the corresponding eigenvectors;
- use the eigenvalue–eigenvector method to find solutions for a system of DE's;
- determine the linear dependance/independance of the solutions;
- find a general solution for the system of DE's;
- find real solutions for a system of DE's which has complex eigenvalues and eigenvectors;
- find two linear independent eigenvectors which correspond to a double (repeated) root of $C(\lambda) = 0$;
- solve initial value problems using the eigenvalue–eigenvector method.

2.1 INTRODUCTION

We have already pointed out that a system of n linear equations in functions $x_1(t), x_2(t), \ldots, x_n(t)$ and with constant coefficients a_{ij}, may be expressed in the form $\dot{\mathbf{X}} = \mathbf{AX}$, with $\mathbf{A}$ the $n \times n$ coefficient matrix and $\mathbf{X}$ the vector $[x_1, \ldots, x_n]^T$. Thus the system

$$\begin{aligned} \dot{x} - y &= x \\ \dot{y} + y &= 0 \\ \dot{z} - 2z &= 0 \end{aligned}$$

may be expressed in the form

$$\dot{\mathbf{X}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \mathbf{X} \text{ with } \mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

In this chapter it will be shown that the solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ may be found by studying the eigenvalues and eigenvectors of $\mathbf{A}$. This approach is known as the *eigenvalue–eigenvector method*. If this method yields n linearly independent solutions of a system of differential equations, then a general solution is a linear combination of these solutions — the same as that obtained by applying the method of elimination which was described in the previous chapter.

In what follows, $\mathbf{A}$ is an $n \times n$ matrix with *real* entries. For the sake of completeness we give the following definitions and theorem with which you should be familiar:

Definition 2.1 *If a vector* $\mathbf{U} \neq \mathbf{0}$ *and a real number* λ *exists such that*

$$\mathbf{AU} = \lambda \mathbf{U},$$

i.e.

$$(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{U} = \mathbf{0},$$

where $\mathbf{I}$ *is the identity matrix, then* λ *is said to be an* **eigenvalue** *of* $\mathbf{A}$*. The vector* $\mathbf{U}$ *is called the* **associated** *or* **corresponding eigenvector***.*

Theorem 2.2 *The number* λ *is an eigenvalue of* $\mathbf{A}$ *if and only if*

$$\det(\mathbf{A} - \lambda \mathbf{I}) \equiv |\mathbf{A} - \lambda \mathbf{I}| = 0. \tag{2.1}$$

Equation (2.1) is called the *characteristic equation* of $\mathbf{A}$.

Definition 2.3

$$C(\lambda) = |\mathbf{A} - \lambda \mathbf{I}| = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{bmatrix}$$

is known as the characteristic polynomial *of* $\mathbf{A}$*.*

Exercise 2.4 Find the eigenvalues and eigenvectors of $\mathbf{A}$ where

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

(b) $\mathbf{A} = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}$

2.2 THE EIGENVALUE–EIGENVECTOR METHOD

As far as solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ are concerned, the analogy between $\dot{x} = ax$ (a a constant) and $\dot{\mathbf{X}} = \mathbf{AX}$, gives rise to the conjecture that exponentials should satisfy the equation $\dot{\mathbf{X}} = \mathbf{AX}$. Let us assume that $\mathbf{X}(t) = e^{\lambda t}\mathbf{U}$ is a solution. Since $\dot{\mathbf{X}} = \lambda e^{\lambda t}\mathbf{U}$, $\mathbf{U}$ and λ must satisfy (from $\dot{\mathbf{X}} = \mathbf{AX}$)

$$\lambda e^{\lambda t}\mathbf{U} = \mathbf{A}e^{\lambda t}\mathbf{U} = e^{\lambda t}\mathbf{AU},$$

i.e.

$$e^{\lambda t}(\mathbf{A} - \lambda\mathbf{I})\mathbf{U} = \mathbf{0},$$

i.e.

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{U} = \mathbf{0},$$

since $e^{\lambda t} \neq 0$. We note that this equation expresses precisely the connection between an eigenvalue λ of $\mathbf{A}$ and the associated eigenvector $\mathbf{U}$. If λ is real, then $\mathbf{U} \neq \mathbf{0}$ is the solution of a homogeneous system of algebraic equations with real coefficients and may, therefore, be chosen as a real vector.
The following theorem can now be formulated:

Theorem 2.5 *If* $\mathbf{U}_1$ *is a real eigenvector of* $\mathbf{A}$ *corresponding to the real eigenvalue* λ_1*, then* $\mathbf{X}_1(t) = e^{\lambda_1 t}\mathbf{U}_1$ *is a solution of* $\dot{\mathbf{X}} = \mathbf{AX}$ *for all* t.

Exercise 2.6

(1) Prove Theorems 2.2 and 2.5.

(2) Find solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ with

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}.$$

Answer:

$$\mathbf{U}_1 = k_1[0,\ 1, -1]^T, \quad \mathbf{U}_2 = e^{-2t}k_2[-2,\ 1,\ 0]^T, \quad \mathbf{U}_3 = e^{-3t}k_3[1,\ 0, -1]^T.$$

(In general we will not provide answers, as solutions can be checked by substituting into the original equation. Do this as an exercise.)

As in Chapter 1 we have

Theorem 2.7 (The Superposition Principle) *Assume that* $\mathbf{X}_1(t), \ldots, \mathbf{X}_r(t)$ *are solutions of* $\dot{\mathbf{X}} = \mathbf{AX}$. *Then*

$$\mathbf{X}(t) = k_1\mathbf{X}_1(t) + \ldots + k_r\mathbf{X}_r(t)$$

with k_i, $i = 1, 2, \ldots, r$ *arbitrary constants, is also a solution.*

Exercise 2.8 Prove Theorem 2.7.

Remark 2.9

(1) In the proof use is made of the fact that for any $n \times n$ matrix $\mathbf{A}$

$$\mathbf{A}\left\{k_1\mathbf{X}_1(t) + k_2\mathbf{X}_2(t) + \ldots + k_r\mathbf{X}_r(t)\right\} = k_1\mathbf{A}\mathbf{X}_1(t) + \ldots + k_r\mathbf{A}\mathbf{X}_r(t),$$

i.e.

$$\mathbf{A}\left\{\sum_{i=1}^{r} k_i\mathbf{X}_i(t)\right\} = \sum_{i=1}^{r} k_i\mathbf{A}\mathbf{X}_i(t).$$

(2) It follows from Theorem 2.7 that the set of solutions of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ forms a vector space (choose the trivial solution $\mathbf{X} \equiv \mathbf{0}$ as the zero element). We may, therefore, now use the expression "*the space of solutions*" or "*the solution space*".

It will be shown later on that the dimension of the solution space corresponds with that of $\mathbf{A}$, i.e. if $\mathbf{A}$ is an $n \times n$ matrix, the space of solutions $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ will be n–dimensional.

2.3 LINEAR INDEPENDENCE OF EIGENVECTORS

We recall

Definition 2.10 *The vectors* $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *are* **linearly independent** *if it follows from*

$$c_1\mathbf{X}_1 + \ldots + c_n\mathbf{X}_n = \mathbf{0}$$

that $c_i = 0$ *with* $c_i,\ \ i = 1, \ldots, n$ *arbitrary constants.*

The following theorem supplies a condition for eigenvectors $\mathbf{U}_1$, $\mathbf{U}_2, \ldots, \mathbf{U}_n$, corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\mathbf{A}$, to be linearly independent, in other words, a condition for solutions

$$\mathbf{X}_1(t), \ldots, \mathbf{X}_n(t) \quad \text{with} \quad \mathbf{X}_i(t) = e^{\lambda_i t}\mathbf{U}_i, \ \ i = 1, \ldots, n$$

to be linearly independent.

Theorem 2.11 *Suppose the* $n \times n$ *matrix* $\mathbf{A}$ *has* **different** (*or* **distinct**) *eigenvalues* $\lambda_1, \ldots, \lambda_n$ *corresponding to eigenvectors* $\mathbf{U}_1, \ldots \mathbf{U}_n$*. Then* $\mathbf{U}_1, \ldots, \mathbf{U}_n$ *are linearly independent.*

Proof. We prove the statement by means of the induction principle. We start with $n = 2$. Suppose

$$c_1\mathbf{U}_1 + c_2\mathbf{U}_2 = \mathbf{0}. \qquad (1)$$

If $\mathbf{A}$ is applied to both sides of (1), we find

$$\mathbf{0} = \mathbf{A}(c_1\mathbf{U}_1 + c_2\mathbf{U}_2) = c_1\mathbf{A}\mathbf{U}_1 + c_2\mathbf{A}\mathbf{U}_2$$

or

$$c_1\lambda_1\mathbf{U}_1 + c_2\lambda_2\mathbf{U}_2 = \mathbf{0}, \qquad \text{since } \mathbf{A}\mathbf{U}_i = \lambda_i\mathbf{U}_i. \qquad (2)$$

We now multiply (1) by λ_1 and subtract from (2). This gives

$$(c_1\lambda_1\mathbf{U}_1 + c_2\lambda_2\mathbf{U}_2) - (c_1\lambda_1\mathbf{U}_1 + c_2\lambda_1\mathbf{U}_2) = \mathbf{0}$$

or

$$c_2(\lambda_2 - \lambda_1)\mathbf{U}_2 = \mathbf{0}.$$

Since $\mathbf{U}_2 \neq \mathbf{0}$ (from the definition of an eigenvector) and since $\lambda_2 \neq \lambda_1$, we derive that $c_2 = 0$. By substituting $c_2 = 0$ in (1), we find that $c_1 = 0$, which proves the validity of the theorem for $n = 2$. Now suppose the theorem holds for $n = k$, i.e. assume that any k eigenvectors which correspond to k distinct eigenvalues are linearly independent. We now prove that the theorem holds for $n = k + 1$. Hence, we assume that

$$c_1\mathbf{U}_1 + c_2\mathbf{U}_2 + \ldots + c_k\mathbf{U}_k + c_{k+1}\mathbf{U}_{k+1} = \mathbf{0}. \qquad (3)$$

By applying $\mathbf{A}$ to both sides of (3) and making use of the fact that $\mathbf{A}\mathbf{U}_i = \lambda_i\mathbf{U}_i$, we find that

$$c_1\lambda_1\mathbf{U}_1 + c_2\lambda_2\mathbf{U}_2 + \ldots + c_k\lambda_k\mathbf{U}_k + c_{k+1}\lambda_{k+1}\mathbf{U}_{k+1} = \mathbf{0}. \qquad (4)$$

Next, multiply both sides of (3) by λ_{k+1} and subtract from (4):

$$c_1(\lambda_1 - \lambda_{k+1})\mathbf{U}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{U}_2 + \ldots + c_k(\lambda_k - \lambda_{k+1})\mathbf{U}_k = \mathbf{0}.$$

From the induction assumption we have that $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_k$ are linearly independent. Thus

$$c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \ldots = c_k(\lambda_k - \lambda_{k+1}) = 0,$$

and since $\lambda_i \neq \lambda_{k+1}$ for $i = 1, 2, \ldots, k$, we derive that $c_1 = c_2 = \ldots = c_k = 0$. Thus from (3) it follows that $c_{k+1} = 0$. The statement therefore holds for $m = k + 1$, thereby completing the proof.

Exercise 2.12

Find an $n \times n$ matrix $\mathbf{A}$ such that although $\mathbf{A}$ does not have n different eigenvalues, n linearly independent corresponding eigenvectors still exist.

A natural question to ask is: which $n \times n$ matrices have n linearly independent eigenvectors? An important class of matrices with this property, is the class of *symmetric matrices*. These matrices appear frequently in applications. We will return to this class of matrices in Chapter 3.

Exercise 2.13

(1) Study Appendix A at the end of the study guide which deals with symmetric matrices.

(2) For each of the following matrices, find three linearly independent eigenvectors.

(a) $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & \sqrt{2} \\ 0 & \sqrt{2} & 2 \end{bmatrix}$ (c) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

(3) Write out the details which prove $\overline{\mathbf{A}\mathbf{u}} = \mathbf{A}\overline{\mathbf{u}}$ if $\mathbf{A}$ is real.

(4) Solve $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, $\mathbf{x}(0) = \mathbf{x_0}$ for the matrix $\mathbf{A}$ in 2(c) above, when

(a) $\mathbf{x_0} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ (b) $\mathbf{x_0} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

(5) If $\mathbf{A}$ and $\mathbf{B}$ are symmetric and $\mathbf{AB} = \mathbf{BA}$, prove that $\mathbf{AB}$ is symmetric.

(6) $\mathbf{A}$ is antisymmetric if $\mathbf{A}^T = -\mathbf{A}$. Show that if $\mathbf{A}$ is antisymmetric, then $\mathbf{u}^T\mathbf{A}\overline{\mathbf{v}} = -\overline{\mathbf{v}}^T\mathbf{A}\mathbf{u}$.

(7) Show that the eigenvalues of a real, antisymmetric matrix are pure imaginary (0 may be considered a pure imaginary number).

(8) Prove that if $\mathbf{A}$ is a real, $(2n+1)\times(2n+1)$ antisymmetric matrix, then $\lambda = 0$ is an eigenvalue of $\mathbf{A}$.

We recall that a convenient method of determining whether solutions

$$\mathbf{X}_1(t) = \begin{bmatrix} x_{11}(t) \\ x_{21}(t) \\ \vdots \\ x_{n1}(t) \end{bmatrix}, \quad \mathbf{X}_2(t) = \begin{bmatrix} x_{12}(t) \\ x_{22}(t) \\ \vdots \\ x_{n2}(t) \end{bmatrix}, \ldots, \ \mathbf{X}_n(t) = \begin{bmatrix} x_{1n}(t) \\ x_{2n}(t) \\ \vdots \\ x_{nn}(t) \end{bmatrix},$$

of a system of linear equations are linearly independent, is to find the *Wronskian* of the solutions. We recall that the Wronskian $W(t)$ of the solutions $\mathbf{X}_1(t), \ldots, \mathbf{X}_n(t)$ is the determinant

$$W(t) = |\mathbf{X}_1(t), \ \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)| = \det \begin{bmatrix} x_{11}(t) & \ldots & x_{1n}(t) \\ \vdots & \vdots & \vdots \\ x_{n1}(t) & \ldots & x_{nn(t)} \end{bmatrix}.$$

By recalling that the value of a determinant in which one of the rows or columns is a linear combination of another row or column, is zero, it follows immediately that $W = 0$ if and only if $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are linearly dependent. Therefore, if $W \neq 0$, the solutions $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are linearly independent.

Since every $\mathbf{X}_i, \ i = 1, \ldots, n$ is a function of t, we must investigate the questions whether it is possible that on an interval $I \subset R$, W could be zero in some points and in others not. The following theorem, to which we shall return in Chapter 4, settles the question.

Theorem 2.14 *If W, the Wronskian of the solutions of a system of differential equations, is zero at a point t_0 in an interval I, it is zero for all $t \in I$.*

The statement of the theorem is equivalent to the following statement: Either

$$W(t) = 0 \ \ \forall t \in I$$

or

$$W(t) \neq 0 \ \ \forall t \in I.$$

Example 2.15 Solve the initial value problem

$$\dot{\mathbf{X}} = \begin{bmatrix} 6 & 8 \\ -3 & -4 \end{bmatrix} \mathbf{X} \equiv \mathbf{AX}, \text{ with } \mathbf{X}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Solution

We solve the characteristic equation

$$C(\lambda) = \begin{vmatrix} 6-\lambda & 8 \\ -3 & -4-\lambda \end{vmatrix} = 0$$

to find the eigenvalues of the matrix. This gives

$$C(\lambda) = (6-\lambda)(-4-\lambda) + 24 = \lambda(\lambda - 2) = 0$$

so that $\lambda_1 = 0$ and $\lambda_2 = 2$.
To find the eigenvector $\mathbf{U}_1$ corresponding to $\lambda_1 = 0$ we solve (using the equation $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{U} = \mathbf{0}$ with $\lambda = \lambda_1 = 0$)

$$\begin{bmatrix} 6 & 8 \\ -3 & -4 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which is equivalent to the equation

$$3u_1 + 4u_2 = 0.$$

Choosing, for example $u_1 = 4c_1$ where c_1 is an arbitrary constant, it follows that

$$u_2 = -\frac{3}{4}u_1 = -\frac{3}{4}(4c_1) = -3c_1$$

so that

$$\mathbf{U}_1 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = c_1\begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

and consequently

$$\mathbf{X}_1(t) = c_1\begin{bmatrix} 4 \\ -3 \end{bmatrix}e^{ot} = c_1\begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ corresponding to $\lambda = 0$.
Similarly, for $\lambda_2 = 2$ we solve (again using the equation $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{U} = \mathbf{0}$with $\lambda = \lambda_2 = 2$)

$$\begin{bmatrix} 4 & 8 \\ -3 & -6 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

to find

$$\mathbf{U}_2 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = c_2\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

so that

$$\mathbf{X}_2(t) = c_2\begin{bmatrix} -2 \\ 1 \end{bmatrix}e^{2t}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ corresponding to $\lambda = 2$.
The general solution is therefore given by

$$\mathbf{X}(t) = c_1\begin{bmatrix} 4 \\ -3 \end{bmatrix} + c_2\begin{bmatrix} -2 \\ 1 \end{bmatrix}e^{2t}.$$

From the initial value $\mathbf{X}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, we now solve for c_1 and c_2:

$$\begin{aligned} 1 &= 4c_1 - 2c_2 \\ 2 &= -3c_1 + c_2 \end{aligned}$$

which yields $c_1 = -\dfrac{5}{2}$ and $c_2 = -\dfrac{11}{2}$.

The solution of the initial value problem is therefore given by

$$\mathbf{X} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = -\frac{5}{2}\begin{bmatrix} 4 \\ -3 \end{bmatrix} - \frac{11}{2}\begin{bmatrix} -2 \\ 1 \end{bmatrix} e^{2t}$$

$$= \begin{bmatrix} -10 + 11e^{2t} \\ \\ \dfrac{15}{2} - \dfrac{11}{2}e^{2t} \end{bmatrix}.$$

Exercise 2.16

(1) Study Sections 8.1 and 8.2 up to p. 337 of Z&W.

(2) Do problems 1 – 14 of Exerciss 8.2.1 on p. 346 of Z&W.

By using techniques to be developed in the next section, n linearly independent solutions $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots,$ $\mathbf{X}_n(t)$ of the equation $\dot{\mathbf{X}} = \mathbf{AX}$, can always be found, even if some of, or all, the roots of the characteristic equation are coincident. This implies that the space of solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ is n–dimensional: the solution space is namely a subspace of an n–dimensional space, so that its dimension cannot exceed n. That its dimension is exactly n, now follows from the existence of n linearly independent solutions.

From the above it follows that any solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is of the form $\mathbf{X}(t) = k_1\mathbf{X}_1(t) + \ldots + k_n\mathbf{X}_n(t)$ with $\mathbf{X}_i(t)$, $i = 1, \ldots, n$, the above solutions.

Since any other set of n linearly independent solutions (known as a fundamental set of solutions) also forms a base for the solution space, it is sufficient, in order to determine a general solution of $\dot{\mathbf{X}} = \mathbf{AX}$, to find in any manner any set of n linearly independent solutions.

If one verifies, by differentiation and by evaluating the Wronskian, that they are actually linearly independent solutions, no theoretical justification for one's method is needed.

From the above, it is clear that if the equation $C(\lambda) = 0$ has m–fold roots $(m < n)$, a method will have to be found for determining the $(n - m)$ solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ that may still be lacking and which, together with the solutions already found, constitute a set of n linearly independent solutions. The technique for finding these solutions will be described in Sections 2.5 as well as in Chapter 3.

2.4 COMPLEX EIGENVALUES AND EIGENVECTORS

Even if all the entries of $\mathbf{A}$ are real, it often happens that the eigenvalues of $\mathbf{A}$ are complex. In order to extend Theorem 2.5 to the case where λ is complex, we must define $e^{\lambda t}$ for complex λ. Let $\lambda = a + ib$ with a and b real numbers. In view of Euler's equation we have

$$e^{(a+ib)t} = e^{at}(\cos bt + i \sin bt)$$

– a complex number for every real value of t.

Let λ be a complex eigenvalue with associated eigenvector $\mathbf{U}$. We now have

$$\frac{d}{dt}\left(e^{\lambda t}\mathbf{U}\right) = \lambda e^{\lambda t}\mathbf{U} = e^{\lambda t}\lambda\mathbf{U} = e^{\lambda t}\mathbf{AU} = \mathbf{A}e^{\lambda t}\mathbf{U}.$$

Thus $e^{\lambda t}\mathbf{U}$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ as in the real case.

Note that $e^{\lambda t} \neq 0$ in view of

$$\begin{aligned} \left|e^{\lambda t}\right|^2 = \left|e^{(a+ib)t}\right|^2 &= \left|e^{at}\right|^2 \left|e^{ibt}\right|^2 \\ &= \left|e^{at}\right|^2 \left(\cos^2 bt + \sin^2 bt\right) \\ &= \left|e^{at}\right|^2 \\ &> 0 \end{aligned}$$

for all values of t.

Exercise 2.17 Prove that

$$\frac{d}{dt}\left(e^{(a+ib)t}\right) = (a+ib)\, e^{(a+ib)t},$$

i.e.

$$\frac{d}{dt}\left(e^{\lambda t}\right) = \lambda e^{\lambda t}$$

for λ complex.

Consider again the complex vector $e^{\lambda t}\mathbf{U}$. If $\mathbf{U} = \mathbf{V} + i\mathbf{W}$ with $\mathbf{V}, \mathbf{W} \in R^n$, $\lambda = a + ib$, then

$$\begin{aligned} e^{\lambda t}\mathbf{U} &= e^{at}\left(\cos bt + i \sin bt\right)\left(\mathbf{V} + i\mathbf{W}\right) \\ &= e^{at}\left(\mathbf{V}\cos bt - \mathbf{W}\sin bt\right) + ie^{at}\left(\mathbf{V}\sin bt + \mathbf{W}\cos bt\right). \end{aligned}$$

Consequently

$$\operatorname{Re}\left(e^{\lambda t}\mathbf{U}\right) = e^{at}\left(\mathbf{V}\cos bt - \mathbf{W}\sin bt\right)$$

and

$$\operatorname{Im}\left(e^{\lambda t}\mathbf{U}\right) = e^{at}\left(\mathbf{V}\sin bt + \mathbf{W}\cos bt\right).$$

The following theorem shows that complex eigenvalues of $\mathbf{A}$ may lead to real solutions of $\dot{\mathbf{X}} = \mathbf{AX}$.

Theorem 2.18 *If $\mathbf{X}(t)$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{A}$ real, then $\operatorname{Re}\mathbf{X}(t)$ and $\operatorname{Im}\mathbf{X}(t)$ are solutions.*

Proof. We must show that

$$\frac{d}{dt}\left\{\operatorname{Re}\mathbf{X}(t)\right\} = \mathbf{A}\left\{\operatorname{Re}\mathbf{X}(t)\right\}$$

and similarly for $\operatorname{Im}\mathbf{X}(t)$.
From our assumption we have

$$\frac{d}{dt}\left\{\mathbf{X}(t)\right\} = \mathbf{AX}(t).$$

Therefore

$$\frac{d}{dt}\left\{\operatorname{Re}\mathbf{X}(t)\right\} = \operatorname{Re}\frac{d}{dt}\left\{\mathbf{X}(t)\right\} = \operatorname{Re}\left\{\mathbf{AX}(t)\right\}$$

since the elements of $\mathbf{A}$ are real.

The proof for $\operatorname{Im}\mathbf{X}(t)$ is similar. ■

If $\lambda = a + ib$ and $\mathbf{U} = \mathbf{V} + i\mathbf{W}$, the theorem may be reformulated:

Theorem 2.19 *If $\lambda = a + ib$ is an eigenvalue of* $\mathbf{A}$ *and* $\mathbf{U} = \mathbf{V} + i\mathbf{W}$ *the corresponding eigenvector, then two solutions of* $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ *are*

$$\begin{aligned} \mathbf{X}_1(t) &= e^{at}(\mathbf{V}\cos bt - \mathbf{W}\sin bt) \\ \mathbf{X}_2(t) &= e^{at}(\mathbf{V}\sin bt + \mathbf{W}\cos bt). \end{aligned} \tag{2.2}$$

Example 2.20 Find real–valued fundamental solutions of

$$\dot{\mathbf{X}} = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix} \mathbf{X}.$$

Solution

The characteristic equation is $\lambda^2 - 4\lambda + 13 = 0$. The roots are $\lambda_1 = 2+3i$ and $\lambda_2 = 2-3i$. Choose $\lambda = 2+3i$. The eigenvector equation is now

$$\begin{bmatrix} -3i & 3 \\ -3 & -3i \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -3iv_1 + 3v_2 \\ -3v_1 - 3iv_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

A solution of this linear system is $v_2 = iv_1$, where v_1 is any nonzero constant (note that the two equations are equivalent). We take $v_1 = 1$ so that $\mathbf{V} = (1, i)^T$ is an eigenvector. A *complex–valued* solution is

$$\begin{aligned} \mathbf{X}_1(t) &= e^{(2+3i)t} \begin{bmatrix} 1 \\ i \end{bmatrix} \\ &= e^{2t}(\cos 3t + i\sin 3t) \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \\ &= e^{2t} \left\{ \begin{bmatrix} \cos 3t \\ -\sin 3t \end{bmatrix} + i \begin{bmatrix} \sin 3t \\ \cos 3t \end{bmatrix} \right\}. \end{aligned}$$

Two linearly independent *real–valued* solutions are

$$e^{2t} \begin{bmatrix} \cos 3t \\ -\sin 3t \end{bmatrix} \quad \text{and} \quad e^{2t} \begin{bmatrix} \sin 3t \\ \cos 3t \end{bmatrix}.$$

As for solutions corresponding to $\lambda = 2 - 3i$: This eigenvalue yields identical solutions up to a constant, i.e. solutions which can be obtained by multiplying the solutions already found by a constant. **This happens whenever eigenvalues occur in complex conjugate pairs**. The implication of this is that you need not construct solutions corresponding to the second complex eigenvalue as this yields nothing new.

A general solution of the given problem is

$$c_1 e^{2t} \begin{bmatrix} \cos 3t \\ -\sin 3t \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} \sin 3t \\ \cos 3t \end{bmatrix}.$$

Exercise 2.21 Study Section 8.2.3 on pp. 342 –346 of Z&W.

Exercise 2.22 Do Problems 33 – 46 of Exercise 8.2.3 on pp. 347 – 348 of Z&W.

2.5 NEW ROOTS OF $\dot{\mathbf{X}} = \mathbf{AX}$

Assume that the equation $C(\lambda) = 0$ has multiple roots. In order to find the general solution of $\dot{\mathbf{X}} = \mathbf{AX}$, we seek a method for determining the eigenvectors which may be lacking.

Consider the system

$$\dot{\mathbf{X}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{X}.$$

Then

$$C(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2$$

so that $\lambda = 1$ is a double root of $C(\lambda) = 0$. An associated eigenvector is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Thus $\mathbf{X}_1(t) = e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a solution.

Suppose that $\widehat{\mathbf{X}}(t) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a second solution. Then

$$\frac{d}{dt}\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 \end{bmatrix}.$$

This yields the system

$$\begin{aligned} \dot{x}_1 &= x_1 + x_2 && (1) \\ \dot{x}_2 &= x_2 && (2) \end{aligned}$$

for which we find (from (2)) that

$$x_2(t) = e^t. \qquad (3)$$

Substitution of (3) in (2) gives

$$\dot{x}_1 = x_1 + e^t$$

which has the complementary function

$$x_{1_{\text{C.F.}}}(t) = e^t$$

and the particular integral

$$x_{1_{\text{P.I.}}}(t) = te^t \;\; \text{(from the shift property – see Appendix B)},$$

so that $x_1(t) = e^t + te^t$ together with $x_2(t) = e^t$ is a solution of the system. Consequently

$$\widehat{\mathbf{X}}(t) = \begin{bmatrix} e^t(t+1) \\ e^t \end{bmatrix} = te^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} + e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

By differentiation we see that $\widehat{\mathbf{X}}(t)$ is indeed a solution. We check for linear independence:

$$\det \begin{bmatrix} \mathbf{X}_1(t) & \widehat{\mathbf{X}}(t) \end{bmatrix} = \det \begin{bmatrix} e^t & e^t(t+1) \\ 0 & e^t \end{bmatrix} = \left(e^t\right)^2 \neq 0.$$

The general solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is therefore

$$\mathbf{X}(t) = k_1\mathbf{X}_1(t) + k_2\widehat{\mathbf{X}}(t),$$

with k_1, k_2 arbitrary constants. If we compare $\mathbf{X}_1(t)$ and $\widehat{\mathbf{X}}(t)$, we see that $\widehat{\mathbf{X}}(t)$ is of the form

$$e^t(\mathbf{B} + t\mathbf{C}) \quad \text{with} \quad \mathbf{C} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

an eigenvector of $\mathbf{A}$, whereas $\mathbf{B}$ has to be determined. If λ is a triple root of $C(\lambda) = 0$, it can be deduced in a similar way that if $\mathbf{X}(t) = e^{\lambda t}\mathbf{C}$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$, then

$$(\mathbf{C}t + \mathbf{B})e^{\lambda t} \quad \text{and} \quad \left(\mathbf{C}\frac{t^2}{2} + \mathbf{B}t + \mathbf{D}\right)e^{\lambda t}$$

are also solutions. We return to this method for finding new solutions in Chapter 3.

Example 2.23 Find a general solution of

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix} \mathbf{X} \equiv \mathbf{AX}.$$

Solution

Put $C(\lambda) = 0$, i.e.

$$\det \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -2 & -5 & -4-\lambda \end{bmatrix} = 0.$$

Then $-4\lambda^2 - \lambda^3 - 5\lambda - 2 = 0$ from which we obtain $\lambda = -1, -1$, or -2. We first determine an eigenvector

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

for $\lambda = -2$. Put

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ -2 & -5 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = 0.$$

Hence

$$\begin{aligned} 2u_1 + u_2 \quad &= 0 \\ 2u_2 + u_3 &= 0 \\ -2u_1 - 5u_2 - 2u_3 &= 0. \end{aligned}$$

This system is satisfied by $u_1 = k, \quad u_2 = -2k, \quad u_3 = 4k$. Consequently

$$\mathbf{X}_1(t) = k_1 e^{-2t} \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$. For $\lambda = -1$,

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

is an eigenvector (check this!) so that

$$\mathbf{X}_2(t) = k_2 e^{-t} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

is a second solution of $\dot{\mathbf{X}} = \mathbf{AX}$. From the preceding, a third solution is

$$\mathbf{X}_3(t) = e^{-t}\mathbf{B} + te^{-t} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \equiv e^{-t}\mathbf{B} + te^{-t}\mathbf{C}.$$

In order to determine $\mathbf{B}$, substitute $\mathbf{X}_3(t)$ in $\dot{\mathbf{X}} = \mathbf{AX}$. This yields

$$\begin{aligned} -e^{-t}\mathbf{B} + e^{-t}\mathbf{C} - te^{-t}\mathbf{C} &= \mathbf{A}\left(e^{-t}\mathbf{B} + te^{-t}\mathbf{C}\right) \\ &= e^{-t}\mathbf{AB} + te^{-t}\mathbf{AC}. \end{aligned}$$

By equating coefficients of like powers of t, we obtain

$$-\mathbf{C} = \mathbf{AC} \tag{2.3}$$

$$-\mathbf{B} + \mathbf{C} = \mathbf{AB}. \tag{2.4}$$

This yields

$$-\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{B},$$

i.e.

$$\begin{aligned} -b_1 + 1 &= b_2 \\ -b_2 - 1 &= b_3 \\ -b_3 + 1 &= -2b_1 - 5b_2 - 4b_3. \end{aligned}$$

A solution of this system is $b_1 = 1, \ \ b_2 = 0, \ \ b_3 = -1$ (by eliminating either b_1 or b_3, it appears that the third equation yields nothing new). Therefore

$$\mathbf{X}_3(t) = e^{-t} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + te^{-t} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

and the general solution is

$$\mathbf{X}(t) = c_1\mathbf{X}_1(t) + c_2\mathbf{X}_2(t) + c_3\mathbf{X}_3(t)$$

with c_1, c_2, c_3 arbitrary constants. That $\mathbf{X}_1(t)$, $\mathbf{X}_2(t)$ and $\mathbf{X}_3(t)$ are linearly independent, can be checked by calculating their Wronskian.

Note that in this example $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \mathbf{B}$ is not an eigenvector associated with the eigenvalue $\lambda = -1$ of

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix},$$

since

$$(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -2 & -5 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \mathbf{C} \neq \mathbf{0}.$$

$\mathbf{B}$ is known as a *generalised eigenvector* or *root vector* of $\mathbf{A}$. This subject is dealt with in Chapter 3.

Exercise 2.24
Find a general solution of $\dot{\mathbf{X}} = \mathbf{AX}$ if

$$\text{(a)} \quad \mathbf{A} = \begin{bmatrix} -3 & 4 \\ -1 & 1 \end{bmatrix}, \qquad \text{(b)} \quad \mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}, \qquad \text{(c)} \quad \mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

2.6 INITIAL VALUE PROBLEMS; SOLUTION OF INITIAL VALUE PROBLEMS BY THE EIGENVALUE–EIGENVECTOR METHOD

Definition 2.25 *The system* $\dot{\mathbf{X}} = \mathbf{AX}$ *together with an (initial) condition* $\mathbf{X}(t_0) = \mathbf{X}_0$ *is known as an* **initial value problem**. *A solution to the problem is a differentiable function* $\mathbf{X}(t)$ *that is a solution of* $\dot{\mathbf{X}} = \mathbf{AX}$ *and such that* $\mathbf{X}(t_0) = \mathbf{X}_0$.

For the sake of convenience, t_0 is often (but not always) taken as 0.

Example 2.26 Solve

$$\dot{\mathbf{x}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Solution
From $C(\lambda) = (1-\lambda)^2 = 0$ it follows that $\lambda = 1$ (twice). Thus

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{u} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

from which we get $u_2 = 0$ and u_1 is arbitrary. Hence $\mathbf{u} = c \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector and therefore $\mathbf{x}(t) = ce^t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is the corresponding solution. But $\mathbf{x}(0) = \begin{bmatrix} c \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This implies that cannot yet find a solution that satisfies the given initial condition with what we know so far.

This example illustrates that application of the eigenvalue–eigenvector method does not always lead to a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ which satisfies the initial condition. This problem, as can be expected, is experienced when the characteristic equation has multiple roots. The following theorems contain sufficient conditions under which a solution for the initial value problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$ can be found by means of the eigenvalue–eigenvector method.

Theorem 2.27 *If n linearly independent eigenvectors exist with corresponding real (not necessarily different) eigenvalues of the $n \times n$ matrix $\mathbf{A}$, then the initial value problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$ has a solution for every $\mathbf{X}_0 \in R^n$.*

Exercise 2.28 Prove Theorem 2.27.

From Theorem 2.27 we can derive

Theorem 2.29 *If $\mathbf{A}$ has n real different eigenvalues, then $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$, has a solution for every $\mathbf{X}_0 \in R^n$.*

Exercise 2.30

(1) Study Example 3 on p. 338 of the Z&W.

(2) Solve

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix} \mathbf{X}, \quad \mathbf{X}(0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Why is there a unique solution to the above system?

Example 2.31 Find a solution of

$$\dot{\mathbf{X}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \mathbf{AX}, \quad \mathbf{X}_0 = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}.$$

Solution

The characteristic equation has roots $\lambda = 1, 1, 2$. If we put $\lambda = 1$ into $(\mathbf{A} - \lambda \mathbf{I})\mathbf{U} = \mathbf{0}$, we have

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Now

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = k_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

satisfies this equation and

$$\mathbf{X}_1(t) = e^t k_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ which, however, does not satisfy the initial condition, since

$$\mathbf{X}_1(t) = e^t k_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

can, for no values of t and k_1, yield the vector $\begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$. Next, put $\lambda = 2$ into $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{V} = \mathbf{0}$. We then have

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

from which follows

$$\begin{aligned} -v_1 + v_2 &= 0 \\ -v_2 &= 0. \end{aligned}$$

Therefore, $v_1 = 0$, $v_2 = 0$, while, since v_3 does not appear in the equations, v_3 may be chosen arbitrarily. Put $v_3 = k_2$. Then

$$\mathbf{V} = k_2 e^{2t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ which, however, does not satisfy the initial condition. Consequently we must check whether $\mathbf{X}_0$ belongs to the span of the eigenvectors $\mathbf{U}$ and $\mathbf{V}$.

Put $m\mathbf{U} + n\mathbf{V} = \mathbf{X}$ where m and n are scalars. By putting $t = 0$ we have

$$m \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + n \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix} \tag{2.5}$$

whence

$$\begin{aligned} m &= 2, \\ n &= -2. \end{aligned}$$

Consequently

$$\mathbf{X}(t) = 2e^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 2e^{2t} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ which also satisfies the initial condition.

Exercise 2.32 The $n \times n$ matrix $\mathbf{A}$ in each of the following initial value problems $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, $\mathbf{x}(t_0) = \mathbf{x_0}$ has n distinct, real eigenvalues. Find a general solution containing n arbitrary constants and then determine the values of the constants using the initial conditions $\mathbf{x}(t_0) = \mathbf{x_0}$.

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$, $\mathbf{x}(-1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (b) $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$, $\mathbf{x}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

(c) $\mathbf{A} = \begin{bmatrix} 4 & -3 & -2 \\ 2 & -1 & -2 \\ 3 & -3 & -1 \end{bmatrix}$, $\mathbf{x}(1) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ (d) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{x}(0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

CHAPTER 3

Generalised Eigenvectors (Root Vectors) and Systems of Linear Differential Equations

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding root vectors and systems of linear DE's:

- a root vector of order k;
- finding root vectors of order k corresponding to a multiple root of $C(\lambda) = 0$;
- linear independence of root vectors;
- existence of a solution of an initial value problem when $\mathbf{A}$ has constant entries.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- determine a root vector of order of k of $\mathbf{A}$;
- use root vectors to find linear independent solutions of $\dot{\mathbf{X}} = \mathbf{AX}$;
- solve initial value problems $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$ where the matrix $\mathbf{A}$ has multiple roots;
- know why real symmetric matrices has no root vectors of order ≥ 2.

3.1 GENERALISED EIGENVECTORS OR ROOT VECTORS

Suppose $\mathbf{A}$ is a square matrix and λ a multiple root of the characteristic equation with corresponding eigenvector $\mathbf{C} \neq \mathbf{0}$, i.e. $e^{\lambda t}\mathbf{C}$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$. In Section 5 of Chapter 2 we derived a second solution, $(\mathbf{B} + \mathbf{C}t)\, e^{\lambda t}$ where $\mathbf{B}$ is the solution of the equation

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{B} = \mathbf{C} \neq \mathbf{0}. \tag{3.1}$$

As has been previously observed, $\mathbf{B}$ is not an eigenvector corresponding to λ. This is clear from (3.1), the defining equation for $\mathbf{B}$. Equations of the form (3.1) in which $\mathbf{B}$ has to be determined and $\mathbf{C}$ is an eigenvector leads to the concept of *a generalised eigenvector* or *root vector*.

Before formally introducing the concept of a generalised eigenvector or root vector, we observe that equation (3.1) may be expressed in a form in which $\mathbf{C}$ does not appear at all and in which the right hand side equals zero, viz

$$(\mathbf{A}-\lambda\mathbf{I})^2\,\mathbf{B}=(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{C}=\mathbf{0} \tag{3.2}$$

by applying $\mathbf{A}-\lambda\mathbf{I}$ to (3.1). If λ is at least a triple root of the characteristic equation, then

$$\left(\frac{1}{2}\mathbf{C}^2t+\mathbf{B}t+\mathbf{D}\right)e^{\lambda t}$$

is, as in Section 2.5, a third solution, provided that the equations

$$\begin{aligned}&(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{C}=\mathbf{0},\quad \mathbf{C}\neq\mathbf{0}\\&(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{B}=\mathbf{C}\\&(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{D}=\mathbf{B}\end{aligned}$$

are satisfied. The last two equations may, by applying $(\mathbf{A}-\lambda\mathbf{I})$ and $(\mathbf{A}-\lambda\mathbf{I})^2$, be written as

$$(\mathbf{A}-\lambda\mathbf{I})^2\,\mathbf{B}=\mathbf{0} \tag{3.3}$$

$$(\mathbf{A}-\lambda\mathbf{I})^3\,\mathbf{D}=\mathbf{0}, \tag{3.4}$$

because

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{I})^2\,\mathbf{B}&=(\mathbf{A}-\lambda\mathbf{I})\,(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{B}=(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{C}=\mathbf{0}\\(\mathbf{A}-\lambda\mathbf{I})^3\,\mathbf{B}&=(\mathbf{A}-\lambda\mathbf{I})\,(\mathbf{A}-\lambda\mathbf{I})^2\,\mathbf{B}=(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{0}=\mathbf{0}.\end{aligned}$$

Note that on the other hand, we have $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{B}\neq\mathbf{0}$and $(\mathbf{A}-\lambda\mathbf{I})^2\,\mathbf{D}\neq\mathbf{0}$, because $\mathbf{B}$ and $\mathbf{D}$ are not eigenvectors corresponding to λ.

Generally, we now consider equations of the form $(\mathbf{A}-\lambda^*\mathbf{I})^k\,\mathbf{U}=\mathbf{0}$, $\mathbf{U}\neq\mathbf{0}$, with k the smallest number for which this relation holds, in other words, for $i=1,\ldots,k$, the relationship

$$(\mathbf{A}-\lambda^*\mathbf{I})^{k-i}\,\mathbf{U}\neq\mathbf{0}$$

holds.

Definition 3.1 *The non–zero vector* $\mathbf{V}$ *is said to be a* **root vector of order** k, $k\geq 1$, *of the matrix* $\mathbf{A}$, *if a number* λ^* *exists such that*

$$(\mathbf{A}-\lambda^*\mathbf{I})^k\,\mathbf{V}\;=\;\mathbf{0} \tag{3.5}$$

$$(\mathbf{A}-\lambda^*\mathbf{I})^{k-1}\,\mathbf{V}\;\neq\;\mathbf{0}. \tag{3.6}$$

If we bear in mind that $\mathbf{V}$ is said to be *annihilated* by $(\mathbf{A}-\lambda^*\mathbf{I})^k$ if $(\mathbf{A}-\lambda^*\mathbf{I})^k\,\mathbf{V}=\mathbf{0}$, Definition 3.1 can be reformulated as:

Definition 3.2 *The vector* $\mathbf{V}$ *is a* **root vector of order** k, $k\geq 1$, *of* $\mathbf{A}$, *if a number* λ^* *exists such that* $\mathbf{V}$ *is annihilated by* $(\mathbf{A}-\lambda^*\mathbf{I})^k$, *but not by* $(\mathbf{A}-\lambda^*\mathbf{I})^{k-1}$.

Note that unless $(\mathbf{A}-\lambda^*\mathbf{I})$ is singular (non–invertible), $(\mathbf{A}-\lambda^*\mathbf{I})^k$ would be a nonzero vector for all k. Hence the existence of k, $\mathbf{V}\neq\mathbf{0}$and λ^* satisfying (3.5) – (3.6) implies $\mathbf{A}-\lambda^*\mathbf{I}$is singular, which means that λ^* must be an eigenvalue of $\mathbf{A}$. It is now clear, by putting $k=1$, that an eigenvector is a root vector of order 1 while the vectors $\mathbf{B}$ and $\mathbf{D}$ above are root vectors of order 2 and 3 respectively.

Example 3.3 Verify that $[2,0,1]^T$ is a root vector of order 3 corresponding to the eigenvalue $\lambda=0$ of

$$\mathbf{A}=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}.$$

Solution Let $\mathbf{v}_3=[2,0,1]^T$. Since $\lambda=0$, $\mathbf{A}-\lambda\mathbf{I}=\mathbf{A}$ we get

$$\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\mathbf{v}_3=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}=\mathbf{v}_2$$

from which

$$\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\mathbf{v}_2=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}=\mathbf{v}_1$$

and finally

$$\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\mathbf{v}_1=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

This implies that $\mathbf{v}_3$, $\mathbf{v}_2$ and $\mathbf{v}_1$ are the root vectors of order 3, 2, and 1, respectively.

The following theorem provides a condition under which a vector $\mathbf{V}$ is a root vector of order k, while it also supplies a method for constructing a sequence of root vectors such that the order of each root vector is one less than that of its predecessor. This theorem will, after a few more formalities, supply us with an easy method for finding the general solution of $\dot{\mathbf{X}}=\mathbf{AX}$ in the case where some or all of the eigenvalues of $\mathbf{A}$ are coincident.

Theorem 3.4 *Suppose λ^* is an eigenvalue of $\mathbf{A}$. Let the sequence of vectors $\{\mathbf{V}_{k-j}\}$, $j=0,\ldots,k-1$, be defined by*

$$(\mathbf{A}-\lambda^*\mathbf{I})\,\mathbf{V}_{k-j}=\mathbf{V}_{k-(j+1)},\quad j=0,\ldots,k-2. \tag{3.7}$$

Then $\mathbf{V}_k$, $\mathbf{V}_{k-1},\ldots,\mathbf{V}_1$ are root vectors of order k, $k-1,\ldots,1$ respectively if and only if $\mathbf{V}_1$ is an eigenvector of $\mathbf{A}$ corresponding to λ^*.

If the root vectors correspond to different eigenvalues, one must distinguish between the different sequences of root vectors.

Definition 3.5 *If $\mathbf{U}_{i,k}$ is a root vector of order k, corresponding to the eigenvalue λ_i of $\mathbf{A}$, the* **sequence of root vectors** $\{\mathbf{U}_{i,k-j}\}$, $j=0,\ldots,k-1$, *defined by*

$$(\mathbf{A}-\lambda_i\mathbf{I})\,\mathbf{U}_{i,k-j}=\mathbf{U}_{i,k-(j+1)},\quad j=0,\ldots,k-2$$

is known as a **chain of length** k.

Example 3.6 Find a chain of length 3 for the matrix

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \end{bmatrix}$$

Solution From

$$C(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 8 & -12 & 6-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 12\lambda - 8 = -(\lambda-2)^3 = 0$$

we see that $\lambda = 2$ is the only eigenvalue of $\mathbf{A}$. From $(\mathbf{A} - \lambda\mathbf{I})\mathbf{U} = \mathbf{0}$ we get

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which easily reduces to

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hence $[1, 2, 4]^T$ is an eigenvector associated with $\lambda = 2$. Now from $(\mathbf{A} - 2\mathbf{I})\mathbf{V} = \mathbf{U}$:

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

we get

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -1 \\ 0 \end{bmatrix}$$

so that $\mathbf{V} = [-1, -1, 0]^T$.
Finally, from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{W} = \mathbf{V}$ we get

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

which can be reduced to

$$\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

so that $\mathbf{W} = [\frac{3}{4}, \frac{1}{2}, 0]^T$.
It is easy to check that $(\mathbf{A} - 2\mathbf{I})\mathbf{W} = \mathbf{V}$; $(\mathbf{A} - 2\mathbf{I})\mathbf{V} = \mathbf{U}$ and $(\mathbf{A} - 2\mathbf{I})\mathbf{V} = \mathbf{0}$, yielding the required chain.

Exercise 3.7

1. *Show that the matrices below have no root vectors of order 2.*

(a) $\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. *For each of the following matrices, find at least one chain of length* 2.

(a) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & -2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -8 & 5 \end{bmatrix}$

3. *For each of the following matrices, find at least one chain of length* 3.

(a) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}$

3.2 ROOT VECTORS AND SOLUTIONS OF $\dot{\mathbf{X}} = \mathbf{AX}$

We have, from the previous section, that if λ is a triple root of the characteristic equation, with $\mathbf{C}$ the corresponding eigenvector and $\mathbf{B}$ and $\mathbf{D}$ root vectors of order 2 and 3 respectively, then

$$\mathbf{X}_1(t) = e^{\lambda t}\mathbf{C}, \;\; \mathbf{X}_2(t) = e^{\lambda t}(\mathbf{B} + \mathbf{C}t) \;\text{ and }\; \mathbf{X}_3 = e^{\lambda t}\left(\frac{1}{2}\mathbf{C}t^2 + \mathbf{B}t + \mathbf{D}\right),$$

are solutions of $\dot{\mathbf{X}} = \mathbf{AX}$.
The following theorem contains the general result:

Theorem 3.8 *Suppose $\mathbf{U}_k$ is a root vector of order k, corresponding to the eigenvalue λ of $\mathbf{A}$. Then*

$$\mathbf{X}(t) = e^{\lambda t}\left(\mathbf{U}_k + t\mathbf{U}_{k-1} + \ldots + \frac{t^{k-1}}{(k-1)!}\mathbf{U}_1\right)$$

is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$.

Proof. We have to proof that $\dot{\mathbf{X}} = \mathbf{AX}$ for all t.
Apply $\mathbf{A} - \lambda\mathbf{I}$to $\mathbf{X}(t)$. Since $(\mathbf{A} - \lambda\mathbf{I})\mathbf{U}_i = \mathbf{U}_{i-1}$, we have

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X}(t) = e^{\lambda t}\left(\mathbf{U}_{k-1} + t\mathbf{U}_{k-2} + \ldots + \frac{t^{k-2}}{(k-2)!}\mathbf{U}_1\right).$$

(Note that $(\mathbf{A} - \lambda\mathbf{I})\mathbf{U}_1 = 0$.)
Hence

$$\begin{aligned}
\dot{\mathbf{X}}(t) &= \lambda e^{\lambda t}\left(\mathbf{U}_k + t\mathbf{U}_{k-1} + \ldots + \frac{t^{k-1}}{(k-1)!}\mathbf{U}_1\right) \\
&\quad + e^{\lambda t}\left(\mathbf{U}_{k-1} + \ldots + \frac{t^{k-2}}{(k-2)!}\mathbf{U}_1\right) \\
&= \lambda\mathbf{X}(t) + (\mathbf{A} - \lambda\mathbf{I})\mathbf{X}(t) \\
&= \mathbf{AX}(t).
\end{aligned}$$

■

Exercise 3.9 Prove that if $\mathbf{U}_k$ are root vectors of order k, corresponding to the eigenvalue λ of $\mathbf{A}$, then

$$\begin{aligned}
&\mathbf{X}_1(t) = e^{\lambda t}\mathbf{U}_1\\
&\mathbf{X}_2(t) = e^{\lambda t}(\mathbf{U}_2 + t\mathbf{U}_1)\\
&\quad\vdots\\
&\mathbf{X}_{k-1}(t) = e^{\lambda t}\left(\mathbf{U}_{k-1} + t\mathbf{U}_{k-2} + \ldots + \frac{t^{k-2}}{(k-2)!}\mathbf{U}_1\right)
\end{aligned}$$

are solutions of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$.

From this, and from Theorem 3.8 (where we proved that

$$\mathbf{X}_k(t) = e^{\lambda t}\left(\mathbf{U}_k + t\mathbf{U}_{k-1} + \ldots + \frac{t^{k-2}}{(k-2)!}\mathbf{U}_1\right)$$

is a solution of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$), we have that k solutions can be found from k root vectors of orders $1, \ldots, k$. If we can prove that these solutions are linearly independent, we can find the general solution of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ when the zeros of the characteristic polynomial are coincident, i.e. we now have at our disposal sufficient methods for finding the solution of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ if the roots of the characteristic equation are different and if the roots of $C(\lambda) = 0$ are coincident.

We first prove

Lemma 3.10 *If* $\mathbf{U}_1, \ldots, \mathbf{U}_k$ *are root vectors of order* $1, \ldots, k$ *corresponding to the eigenvalue* λ *of* $\mathbf{A}$, *then* $\mathbf{U}_1, \ldots, \mathbf{U}_k$ **are linearly independent**.

Proof. Put $c_1\mathbf{U}_1 + \ldots + c_k\mathbf{U}_k = 0$. We must show that $c_i = 0, \;\; i = 1, \ldots, k$. From (3.7) we have

$$\begin{aligned}
c_1(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{U}_k + c_2(\mathbf{A}-\lambda\mathbf{I})^{k-2}\mathbf{U}_k + \ldots + c_{k-2}(\mathbf{A}-\lambda\mathbf{I})^{k-(k-2)}\mathbf{U}_k&\\
+c_{k-1}(\mathbf{A}-\lambda\mathbf{I})^{k-(k-1)}\mathbf{U}_k + c_k\mathbf{U}_k = \mathbf{0}.&
\end{aligned} \tag{3.8}$$

From Definition 3.1 we have that $(\mathbf{A}-\lambda\mathbf{I})^{k-i}\mathbf{U}_k \neq \mathbf{0}, \; i \geq 1$. Apply $(\mathbf{A}-\lambda\mathbf{I})^{k-1}$ to (3.8). Then

$$\mathbf{0} = c_k(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{U}_k \tag{3.9}$$

since all the other terms vanish in view of the index of $\mathbf{A}-\lambda\mathbf{I}$ being equal to or greater than k. Since $(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{U}_k \neq \mathbf{0}$, (3.9) can only hold if $c_k = 0$.

Next apply $(\mathbf{A}-\lambda\mathbf{I})^{k-2}$ to (3.8). This yields

$$\begin{aligned}
\mathbf{0} &= c_{k-1}(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{U}_k + c_k(\mathbf{A}-\lambda\mathbf{I})^{k-2}\mathbf{U}_k\\
&= c_{k-1}(\mathbf{A}-\lambda\mathbf{I})^{k-1}\mathbf{U}_k
\end{aligned} \tag{3.10}$$

since $c_k = 0$. As before, equation (3.10) is valid only if $c_{k-1} = 0$.

By repeating the process, we obtain

$$c_{k-2} = 0, \ldots, c_1 = 0. \qquad \blacksquare$$

Next we prove

Lemma 3.11 *The solutions* $\mathbf{X}_1(t), \ldots, \mathbf{X}_k(t)$ *of* $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$, *with*

$$\mathbf{X}_k(t) = e^{\lambda t}\left(\mathbf{U}_k + t\mathbf{U}_{k-1} + \ldots + \frac{t^{k-1}}{(k-1)!}\mathbf{U}_1\right)$$

(*compare Theorem* 3.8 *and Exercise* 3.9) *are linearly independent.*

Proof. Put

$$c_1\mathbf{X}_1(t) + \ldots + c_k\mathbf{X}_k(t) = \mathbf{0},$$

i.e.

$$e^{\lambda t}\left\{c_1\mathbf{U}_1 + c_2(\mathbf{U}_2 + t\mathbf{U}_1) + c_3\left(\mathbf{U}_3 + t\mathbf{U}_2 + \tfrac{t^2}{2!}\mathbf{U}_1\right) + \ldots\right.$$

$$\ldots + c_{k-1}\left(\mathbf{U}_{k-1} + t\mathbf{U}_{k-2} + \ldots + \tfrac{t^{k-2}}{(k-2)!}\mathbf{U}_1\right) + c_k\left(\mathbf{U}_k + t\mathbf{U}_{k-1} + \ldots\right.$$

$$\left.\left.\ldots + \tfrac{t^{k-1}}{(k-1)!}\mathbf{U}_1\right)\right\} = \mathbf{0}.$$

Rearrangement of the terms yields

$$e^{\lambda t}\left\{\mathbf{U}_1\left(c_1 + c_2 t + \ldots + c_k\frac{t^{k-1}}{(k-1)!}\right) + \mathbf{U}_2\left(c_2 + c_3 t + \ldots + c_k\frac{t^{k-2}}{(k-2)!}\right) + \ldots\right.$$

$$\left.\ldots + \mathbf{U}_{k-2}\left(c_{k-2} + c_{k-1}t + c_k\frac{t^2}{2!}\right) + \mathbf{U}_{k-1}(c_{k-1} + c_k t) + \mathbf{U}_k c_k\right\} = \mathbf{0}.$$

This yields

$$\begin{aligned} c_1 + c_2 t + \ldots + c_k\frac{t^{k-1}}{(k-1)!} &= 0 \;\; \forall t, \\ c_2 + c_3 t + \ldots + c_k\frac{t^{k-2}}{(k-2)!} &= 0 \;\; \forall t, \\ &\vdots \\ c_{k-1} + c_k t &= 0 \;\; \forall t, \\ c_k &= 0, \end{aligned}$$

since $e^{\lambda t} \neq 0$ and $\mathbf{U}_1, \ldots, \mathbf{U}_k$ are linearly independent. Consequently $c_i = 0, \;\; i = 1, \ldots, k.$ ■

Example 3.12 Find a general solution of

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & -12 & 6 \end{bmatrix}\mathbf{X} \equiv \mathbf{A}\mathbf{X}$$

by using root vectors.

Solution

Put

$$\det\begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 8 & -12 & 6-\lambda \end{bmatrix} = 0,$$

i.e.

$$-\lambda\left\{(6-\lambda)(-\lambda)+12\right\}+8=0.$$

Then $-\lambda^3+6\lambda^2-12\lambda+8=0$ from which we obtain $\lambda=2,2,2$. Next we determine the root vectors $\mathbf{U},\mathbf{V},\mathbf{W}$ of order $1,2,3$ respectively.

From $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{U}=\mathbf{0}$, we have

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned} -2u_1+u_2 \qquad &= 0 && (1)\\ -2u_2\ +u_3 &= 0 && (2)\\ 8u_1-12u_2+4u_3 &= 0 && (3)\\ (3)\to(3)+4(1):\qquad -8u_2+4u_3 &= 0 && (4) \end{aligned}$$

Note that (4) is the same as (2).
A solution of this system is

$$\mathbf{U}=\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$

Therefore

$$\mathbf{X}_1(t)=e^{2t}\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

is a solution of $\dot{\mathbf{X}}=\mathbf{AX}$.

From $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{V}=\mathbf{U}$, we have

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix},$$

from which we obtain

$$\mathbf{V}=\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix},$$

so that according to Exercise 3.9

$$\mathbf{X}_2(t)=\left\{\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}+t\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}\right\}\in^{2t}$$

is a second solution of $\dot{\mathbf{X}}=\mathbf{AX}$.

Finally, we have from $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{W}=\mathbf{V}$,

$$\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & -12 & 4 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}=\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix},$$

from which

$$\mathbf{W}=\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}=\begin{bmatrix} 3/4 \\ 1/2 \\ 0 \end{bmatrix}.$$

A third solution of $\dot{\mathbf{X}}=\mathbf{A}\mathbf{X}$is therefore

$$\mathbf{X}_3\,(t)=e^{2t}\left(\begin{bmatrix} 3/4 \\ 1/2 \\ 0 \end{bmatrix}+t\begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}+\frac{t^2}{2}\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}\right).$$

It is easy to verify that we have three linearly independent solutions, by determining their Wronskian.

A general solution is then

$$\mathbf{X}\,(t)=k_1\mathbf{X}_1\,(t)+k_2\mathbf{X}_2\,(t)+k_3\mathbf{X}_3\,(t)$$

with k_1, k_2 and k_3 arbitrary constants.

Exercise 3.13 Do problems 19 – 31 of Exercise 8.2.2 on p. 347 of Z&W.

Remark

In the above example we have constructed the solution of an initial value problem of the form $\dot{\mathbf{X}}=\mathbf{A}\mathbf{X}$ if $\mathbf{A}$ has *constant entries, i.e. the entries are* (*real*) *numbers.*

3.3 TWO IMPORTANT RESULTS (AND KEEPING OUR PROMISES!)

The aim of this section is to present a theorem on the solvability of the initial value problem $\dot{\mathbf{X}}=\mathbf{A}\mathbf{X}$, $\mathbf{X}\,(0)=\mathbf{X}_0$ and to "revisit" symmetric matrices.

The result of the next theorem is stronger than the results in Theorems 2.27 and 2.29 in the sense that the conditions "If ...." are now no longer necessary.

We state without proof

Theorem 3.14 (Existence Theorem for Linear Systems with Constant Coefficients) *For any choice of* $\mathbf{X}_0\in R^n$ *and any* $n\times n$ matrix $\mathbf{A}$ *with* **constant** *entries, a vector function* $\mathbf{X}\,(t)$ *exist such that* $\dot{\mathbf{X}}=\mathbf{A}\mathbf{X}\,(t)\,,\ \mathbf{X}\,(0)=\mathbf{X}_0$.

Remark

(i) Note that the only condition is that the entries of $\mathbf{A}$ must be real — by virtue of the solution techniques developed in the preceding sections, the conditions imposed in Theorems 2.27 and 2.29 can be removed.

(ii) The solution whose existence is guaranteed by this theorem, is unique. This will be proved in the next chapter. In Chapter 7 we will prove a more general result in the sense that the matrix $\mathbf{A}$ can be dependent on t, i.e. the entries of $\mathbf{A}$ may be functions of t. *Students should take care to distinguish between the two cases.*

In case the reader doubts the value of existence and uniqueness theorems, he should read an article by AD Snider on this subject: "Motivating Existence–Uniqueness Theory for Applications Oriented Student"[1]. In this article the writer applies the fact that the one–dimensional wave problem

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad x > 0, \quad t > 0,$$

$$u(x,0) = f(x)$$

$$\frac{\partial u}{\partial t}(x,0) = g(x)$$

$$u(0,t) = 0$$

has one and only one solution, to answer the following question: if a moving wave meets the boundary, does the reflected wave travel below or above the position of equilibrium? The uniqueness of the solution of the wave equation enables one to deduce that the wave travels below the position of equilibrium.

The last result of this chapter pertains to symmetric matrices. We already know that symmetric matrices have only real eigenvalues. In view of our knowledge of root vectors, we can prove the following result:

Theorem 3.15 *If $\mathbf{A}$ is a real symmetric matrix, then $\mathbf{A}$ has no root vectors of order $n \geq 2$.*

Proof Suppose to the contrary that there exists a vector $\mathbf{b} \neq \mathbf{0}$ such that

$$\begin{aligned} (\mathbf{A} - \lambda \mathbf{I})\mathbf{b} &= \mathbf{a} \neq \mathbf{0} \\ (\mathbf{A} - \lambda \mathbf{I})\mathbf{a} &= \mathbf{0}. \end{aligned}$$

Hence

$$(\mathbf{A} - \lambda \mathbf{I})^2 \mathbf{b} = (\mathbf{A} - \lambda \mathbf{I})[(\mathbf{A} - \lambda \mathbf{I})\mathbf{b}] = (\mathbf{A} - \lambda \mathbf{I})\mathbf{a} = \mathbf{0}$$

and

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{b} \neq \mathbf{0}.$$

Since $\mathbf{a} \neq \mathbf{0}$, the matrix

$$\mathbf{a}^T \overline{\mathbf{a}} = \left[\sum_{i=1}^{n} a_i \overline{a}_i \right]$$

is a non-zero 1×1 matrix, i.e. $\mathbf{a}^T \overline{\mathbf{a}} \neq [0]$.
From Lemma A.3 of Appendix A of the study guide, we know that λ is real and since $\mathbf{A}$ and $\lambda \mathbf{I}$ are real, it follows that $\overline{(\mathbf{A} - \lambda \mathbf{I})\mathbf{b}} = (\mathbf{A} - \lambda \mathbf{I})\overline{\mathbf{b}}$.

[1] American Mathematical Monthly, **83**(1976), 805–807.

By the symmetry of $\mathbf{A} - \lambda\mathbf{I}$ and Lemma A.1 of Appendix A of the study guide we have that

$$\begin{aligned}
\mathbf{a}^T\overline{\mathbf{a}} &= [\{(\mathbf{A} - \lambda\mathbf{I})\mathbf{b}\}^T\{\overline{(\mathbf{A} - \lambda\mathbf{I})\mathbf{b}}\}] \\
&= [\{(\mathbf{A} - \lambda\mathbf{I})\mathbf{b}\}^T(\mathbf{A} - \lambda\mathbf{I})\overline{\mathbf{b}}] \\
&= [\overline{\mathbf{b}}^T(\mathbf{A} - \lambda\mathbf{I})(\mathbf{A} - \lambda\mathbf{I})\mathbf{b}] \quad \text{(by Lemma A.1)} \\
&= [\overline{\mathbf{b}}^T(\mathbf{A} - \lambda\mathbf{I})^2\mathbf{b}] \\
&= [\overline{\mathbf{b}}^T\mathbf{0}] \quad (\text{since} \;\; (\mathbf{A} - \lambda\mathbf{I})^2\mathbf{b} = \mathbf{0}) \\
&= 0
\end{aligned}$$

which yields the contradiction. ■

Exercise 3.16 Prove that if $(\mathbf{A} - \mu\mathbf{I})^k\mathbf{u} = \mathbf{0}$, then $\mathbf{u} \neq \mathbf{0}$ is an eigenvalue of $\mathbf{A}$.

Exercise 3.17 Solve the following initial value problems:

(a) $$\dot{\mathbf{X}} = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}\mathbf{X}, \qquad \mathbf{X}(0) = \begin{bmatrix} 4 \\ 5 \\ -9 \end{bmatrix}$$

(b) $$\dot{\mathbf{X}} = \begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{bmatrix}\mathbf{X}, \qquad \mathbf{X}(0) = \begin{bmatrix} 11 \\ 3 \\ -7 \end{bmatrix}.$$

CHAPTER 4

Fundamental Matrices
Non–homogeneous Systems
The Inequality of Gronwall

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding fundamental matrices and non–homogeneous systems:

- a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0;
- the normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0;
- the uniqueness of normalized fundamental matrices;
- using a (normalized) fundamental matrix to solve the IVP $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$;
- the method of variation of parameters;
- the norm of a matrix;
- the inequality of Gronwall.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- determine a fundamental matrix for $\dot{\mathbf{X}} = \mathbf{AX}$;
- normalize the fundamental matrix;
- find a general solution for $\dot{\mathbf{X}} = \mathbf{AX}$ using fundamental matrices;
- solve the IVP $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$ using fundamental matrices;
- use the method of variation of parameters to solve the non–homogeneous problem $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$ (with or without the initial condition $\mathbf{X}(t_0) = \mathbf{X}_0$);
- apply the inequality of Gronwall to estimate the growth of solutions.

4.1 FUNDAMENTAL MATRICES

The concept of the fundamental matrix of a system of differential equations arises when the solutions (which are vectors) of the system are grouped together as the columns of a matrix. Fundamental matrices provide a convenient method of formulating the behaviour of the system $\dot{\mathbf{X}} = \mathbf{AX}$ in an elegant way.

Important results in this chapter are the *uniqueness theorems* for the *homogeneous initial value problem* $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$ and the *non–homogeneous initial value problem* $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$, $\mathbf{X}(t_0) = \mathbf{X}_0$.

Let us now define:

Definition 4.1 *An $n \times n$ matrix $\mathbf{\Phi}$ with the property that its columns are solutions of $\dot{\mathbf{X}} = \mathbf{AX}$, and linearly independent at t_0, is said to be a* **fundamental matrix** *of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0.*

Example 4.2 Find a fundamental matrix for the system $\dot{\mathbf{x}} = \mathbf{Ax}$ at $t_0 = 0$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix}.$$

Solution

It is easy to see that $\lambda = 1$ and $\lambda = 3$ are the two eigenvalues of $\mathbf{A}$ and that $[2,1]^T$ and $[0,1]^T$ are two linearly independent eigenvectors which correspond with them respectively. Hence $\mathbf{x}_1(t) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^t$ and $\mathbf{x}_2(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{3t}$ are two solutions of the given system. Thus

$$\begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}$$

is a fundamental matrix. However,

$$\mathbf{x}_3(t) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{3t} \qquad \text{and} \qquad \mathbf{x}_4(t) = \begin{bmatrix} 0 \\ 4 \end{bmatrix} e^{3t}$$

are also solutions and are linearly independent at $t = 0$. Hence

$$\begin{bmatrix} 2e^t & 0 \\ e^t + e^{3t} & 4e^{3t} \end{bmatrix}$$

is also a fundamental matrix at $t_0 = 0$.

This example illustrates the fact that a system may have two fundamental matrices at a point t_0, i.e. *the fundamental matrix of a system at a point t_0 is not unique.*

In order to construct fundamental matrices, it is necessary to find n solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ which are linearly independent at t_0. From the contents of the two previous chapters, we know that this is always possible. Therefore the following theorem (stated without proof) holds.

Theorem 4.3 *Every system $\dot{\mathbf{X}} = \mathbf{AX}$ has a fundamental matrix at t_0.*

For matrices the equation $\dot{\chi} = \mathbf{A}\chi$, with χ a matrix function of t, is analogous to the equation $\dot{\mathbf{X}} = \mathbf{AX}$ with $\mathbf{X}$ a vector function of t. A fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0 may be formulated in terms of this equation. To begin with, we define:

Definition 4.4 *If $\mathbf{\Phi}(t)$ is the matrix $[a_{ij}(t)]$, then $\dot{\mathbf{\Phi}}(t) = \frac{d}{dt}[\mathbf{\Phi}(t)]$ is the matrix $[\dot{a}_{ij}(t)]$, i.e. the matrix of which the entries are the derivatives of the corresponding entries of $\mathbf{\Phi}(t)$.*

Theorem 4.5 $\mathbf{\Phi}$ is a matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ iff

(a) $\mathbf{\Phi}$ is a solution of $\dot{\chi} = \mathbf{A}\chi$ i.e. $\dot{\mathbf{\Phi}}(t) = \mathbf{A\Phi}(t)$;

(b) $\det \mathbf{\Phi}(t_0) \neq 0$.

Remark

Condition (b) is equivalent to the condition: $(b')\, \mathbf{\Phi}^{-1}(t_0)$ exists.

Exercise 4.6

(1) Prove Theorem 4.5.

(2) Verify Theorem 4.5 for the system with $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$.

On the strength of Theorem 4.5 it is possible to give a representation of the solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ in terms of fundamental matrices.

Theorem 4.7 *If $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0, then $\mathbf{X}(t) = \mathbf{\Phi}(t)\mathbf{K}$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ for every constant vector $\mathbf{K}$.*

It will turn out that any solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is of this form.

Corollary 4.8 *If $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0, then $\mathbf{X}(t) = \mathbf{\Phi}(t)\mathbf{K}$ with $\mathbf{K} = \mathbf{\Phi}^{-1}(t_0)\mathbf{X}_0$, is a solution of the initial value problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$.*

In the expression $\mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(t_0)\mathbf{X}_0$, we are dealing with the product of two matrices of which the second is constant. From Theorem 4.5 we deduce

Corollary 4.9 *If $\mathbf{B}$ is non–singular (invertible) constant matrix and $\mathbf{\Phi}$ a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0, then $\mathbf{\Phi B}$ is also a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0.*

If $t = t_0$, then $\mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(t_0)$ is the identity matrix, so that at $t = t_0$ we have $\mathbf{\Phi}(t)\mathbf{\Phi}^{-1}(t_0)\mathbf{X}_0 = \mathbf{X}_0$. This leads to the concept of normalized fundamental matrix at t_0.

Definition 4.10 *A fundamental matrix $\mathbf{\Phi}(t)$ of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0 with the property that $\mathbf{\Phi}(t_0) = \mathbf{I}$, the identity matrix, is said to be a* **normalized fundamental matrix at** t_0.

Remark

A fundamental matrix $\mathbf{\Phi}(t)$ of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0 can therefore be normalized by multiplying by $\mathbf{\Phi}^{-1}(t_0)$ from the right as follows:

$$\mathbf{\Psi}(t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(t_0) \tag{4.1}$$

where $\mathbf{\Psi}(t)$ is then the *normalized fundamental matrix* of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0. (Clearly, from Corollary 4.9, the matrix $\mathbf{\Psi}(t)$ is a *fundamental* matrix and furthermore, at $t = t_0$, we have

$$\mathbf{\Psi}(t_0) = \mathbf{\Phi}(t_0)\,\mathbf{\Phi}^{-1}(t_0) = \mathbf{I},$$

so that it is also *normalized*.)

Corollary 4.11 can now be formulated:

Corollary 4.11 *If $\mathbf{\Psi}(t)$ is a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0, then $\mathbf{X}(t) = \mathbf{\Psi}(t)\,\mathbf{X}_0$ is a solution of the initial value problem $\dot{\mathbf{X}} = \mathbf{AX},\ \ \mathbf{X}(t_0) = \mathbf{X}_0$.*

The initial value problem $\dot{\mathbf{X}} = \mathbf{AX},\ \ \mathbf{X}(t_0) = \mathbf{X}_0$, can therefore be solved simply by finding a normalized fundamental matrix at t_0.

Exercise 4.12 Prove Theorem 4.7 and Corollaries 4.8, 4.9.

Example 4.13 Find the solution of the initial value problem

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}\mathbf{X}, \quad \mathbf{X}(0) = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Solution

The vector functions

$$\mathbf{X}_1(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{t} \quad \text{and} \quad \mathbf{X}_2(t) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} e^{3t}$$

are two linearly independent solutions of the given system. Therefore

$$\mathbf{\Phi}(t) = \begin{bmatrix} e^{t} & e^{3t} \\ e^{t} & 3e^{3t} \end{bmatrix}$$

is a fundamental matrix. Now

$$\mathbf{\Phi}^{-1}(0) = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

so that

$$\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(0) = \frac{1}{2}\begin{bmatrix} 3e^{t} - e^{3t} & e^{3t} - e^{t} \\ 3e^{t} - 3e^{3t} & 3e^{3t} - e^{t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{if } t = 0.$$

From (4.1) we have that $\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(0)$ is a normalized fundamental matrix at $t = 0$. Consequently from Corollary 4.11

$$\chi(t) = \begin{bmatrix} 3e^{t} - e^{3t} & e^{3t} - e^{t} \\ 3e^{t} - 3e^{3t} & 3e^{3t} - e^{t} \end{bmatrix}\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} e^{t} + e^{3t} \\ e^{t} + 3e^{3t} \end{bmatrix}$$

is a solution to the given problem.

Exercise 4.14 Find a normalized fundamental matrix of the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ at $t_0 = 0$ where $\mathbf{A}$ is given by

(a) $\begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

4.2 THE UNIQUENESS THEOREM FOR LINEAR SYSTEMS WITH CONSTANT COEFFICIENTS

In the proof of this theorem the fact that a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ at t_0 is invertible for every t, is used. We first prove the result for a normalized fundamental matrix.

Suppose that $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ at t_0. A matrix $\mathbf{B}(t)$ has to be determined such that

$$\mathbf{B}(t)\,\mathbf{\Phi}(t) = \mathbf{I} \quad \forall t. \tag{4.2}$$

In the next theorem it is shown that if we take $\mathbf{B}(t)$ as the transpose of a normalized fundamental matrix of

$$\dot{\mathbf{X}} = -\mathbf{A}^T\mathbf{X} \tag{4.3}$$

at t_0, then (4.2) is satisfied. Equation (4.3) is known as the *conjugate* of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$.

Theorem 4.15 *Assume that $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are normalized fundamental matrices of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ and $\dot{\mathbf{X}} = -\mathbf{A}^T\mathbf{X}$ at t_0 respectively. Then*

$$\mathbf{\Psi}^T(t)\,\mathbf{\Phi}(t) = \mathbf{I} \quad \forall t,$$

i.e.

$$(\mathbf{\Phi}(t))^{-1} = \mathbf{\Psi}^T(t).$$

In words the theorem reads: *The inverse of a normalized fundamental matrix $\mathbf{\Phi}(t)$ of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$ at t_0, is equal to the transpose of the normalized fundamental matrix of the conjugate equation $\dot{\mathbf{X}} = -\mathbf{A}^T\mathbf{X}$ at t_0.*

Proof. Since $\mathbf{\Psi}$ is differentiable, $\mathbf{\Psi}^T$ and $\mathbf{\Psi}^T\mathbf{\Phi}$ are likewise differentiable. From the rules for differentiating matrix functions we have

$$\begin{aligned}
\frac{d}{dt}\left[\mathbf{\Psi}^T\mathbf{\Phi}\right] &= \frac{d}{dt}\left[\mathbf{\Psi}^T\right]\mathbf{\Phi} + \mathbf{\Psi}^T\frac{d\mathbf{\Phi}}{dt} \\
&= \left\{\frac{d\mathbf{\Psi}}{dt}\right\}^T\mathbf{\Phi} + \mathbf{\Psi}^T\mathbf{A}\mathbf{\Phi} \\
&= \left\{-\mathbf{A}^T\mathbf{\Psi}\right\}^T\mathbf{\Phi} + \mathbf{\Psi}^T\mathbf{A}\mathbf{\Phi} \\
&= -\mathbf{\Psi}^T\mathbf{A}\mathbf{\Phi} + \mathbf{\Psi}^T\mathbf{A}\mathbf{\Phi} \\
&= \mathbf{0}
\end{aligned}$$

where in the second last step we have made use of $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ and $\left(\mathbf{A}^T\right)^T = \mathbf{A}$ for any matrices $\mathbf{A}$ and $\mathbf{B}$.

Consequently $\mathbf{\Psi}^T\mathbf{\Phi} = \mathbf{C}$, a constant matrix, i.e. $\mathbf{\Psi}^T\mathbf{\Phi}(t) = \mathbf{C}$ for all t. In particular, $\mathbf{\Psi}^T(t_0)\,\mathbf{\Phi}(t_0) = \mathbf{C}$ holds. Since $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are normalized fundamental matrices, we have $\mathbf{I} = \mathbf{\Psi}^T(t_0)\,\mathbf{\Phi}(t_0) = \mathbf{C}$, so that $\mathbf{C} = \mathbf{I}$ holds. This completes the proof. ∎

Corollary 4.16 *Fundamental matrices are non–singular for all t.*

Proof. Suppose $\mathbf{\Phi}$ is a fundamental matrix at t_0 as well as non–singular in t_0, i.e. $\mathbf{\Phi}^{-1}(t_0)$ exists. From the definition $\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(t_0)$ is then a normalized fundamental matrix at t_0, so that $\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(t_0)$ is non–singular for every t, i.e. a matrix $\mathbf{B}$ exists so that $\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(t_0)\,\mathbf{B}(t) = \mathbf{I}$ for every t. This implies that $\mathbf{\Phi}(t)$ itself is non–singular for every t. ∎

Corollary 4.17 *Solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ which are linearly independent at $t = t_0$, are linearly independent for all t.*

Proof. The proof follows immediately from the preceding corollary by recalling that solutions of $\dot{\mathbf{X}} = \mathbf{AX}$ are linearly independent at t_0 if $\det\left(\mathbf{\Phi}^{-1}(t_0)\right) \neq 0$, where $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0. This requirement is equivalent to that of the non–singularity of $\mathbf{\Phi}$ at t_0. ∎

Corollary 4.18 *Fundamental matrices at t_0 are fundamental for all t.*

Proof. From Definition 4.1, the result is merely a formulation of the preceding result. ∎

From the above discussion, it follows that vector functions satisfying $\dot{\mathbf{X}} = \mathbf{AX}$, cannot be linearly independent at one point and linearly dependent at another.

Exercise 4.19 Show by means of an example that arbitrary chosen vectors, linearly independent at a point t_0, are not necessarily linearly independent for all t. The property of Corollary 4.17 is, therefore, a specific property of solutions of $\dot{\mathbf{X}} = \mathbf{AX}$. The converse, however, is not true, i.e. a system of vector functions, linearly independent for all t, does not necessarily constitute a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ with $\mathbf{A}$ some or other matrix.

Theorem 4.20 *Any solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is of the form $\mathbf{X}(t) = \mathbf{\Phi}(t)\,\mathbf{X}(t_0)$ with $\mathbf{\Phi}$ a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0.*

Proof. We show that $\mathbf{\Phi}^{-1}(t)\,\mathbf{X}(t) = \mathbf{Z}(t)$ is a constant vector by showing that $\dot{\mathbf{Z}}(t) = \mathbf{0}$. From Theorem 4.15 we have $\mathbf{Z}(t) = \mathbf{\Psi}^T(t)\,\mathbf{X}(t)$ with $\mathbf{\Psi}(t)$ a solution of $\dot{\chi} = -\mathbf{A}^T\chi$. Hence

$$\begin{aligned}
\dot{\mathbf{Z}} &= \left(\dot{\mathbf{\Psi}}\right)^T\mathbf{X} + \left(\mathbf{\Psi}^T\right)\dot{\mathbf{X}} \\
&= \left(-\mathbf{A}^T\mathbf{\Psi}\right)^T\mathbf{X} + \mathbf{\Psi}^T\mathbf{AX} \\
&= -\mathbf{\Psi}^T\mathbf{AX} + \mathbf{\Psi}^T\mathbf{AX} \\
&= \mathbf{0}.
\end{aligned}$$

Consequently, $\mathbf{Z}(t) = \mathbf{\Phi}^{-1}(t)\,\mathbf{X}(t) = \mathbf{C}$, with $\mathbf{C}$ a constant vector. If we put $t = t_0$, it follows that $\mathbf{C} = \mathbf{X}(t_0)$, so that $\mathbf{\Phi}^{-1}(t)\,\mathbf{X}(t) = \mathbf{X}(t_0)$, i.e. $\mathbf{X}(t) = \mathbf{\Phi}(t)\,\mathbf{X}(t_0)$. ∎

We recall that if $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0, then $\mathbf{\Phi}(t)\mathbf{K}$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ for every choice of $\mathbf{K}$ (Theorem 4.7). From Theorem 4.20 we have that the system $\mathbf{\Phi}(t)\mathbf{K}$ contains all the solutions, i.e. each solution of $\dot{\mathbf{X}} = \mathbf{AX}$ can be obtained from $\mathbf{\Phi}(t)\mathbf{K}$ by a suitable choice of $\mathbf{K}$. If we choose $\mathbf{\Phi}$ a normalized fundamental matrix, we can deduce from Theorem 4.20 that every solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is some or other linear combination of a set of fundamental solutions. This leads to

Definition 4.21 *The vector function $\mathbf{\Phi}(t)\mathbf{K}$ is said to be a* **general solution** *of $\dot{\mathbf{X}} = \mathbf{AX}$ if $\mathbf{\Phi}$ is any fundamental matrix of the system $\dot{\mathbf{X}} = \mathbf{AX}$.*

From Theorem 4.20 we also have

Corollary 4.22 *Any solution of $\dot{\mathbf{X}} = \mathbf{AX}$ that vanishes at t_0, is identically zero. In particular, the problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(0) = \mathbf{0}$, has the unique solution $\mathbf{X}(t) \equiv \mathbf{0}$.*

We now have all the tools for proving the uniqueness theorem for linear systems with constant coefficients.

Theorem 4.23 [1] **(Uniqueness Theorem for Linear Systems with Constant Coefficients)** *The initial value problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$ in which $\mathbf{A}$ is a* **constant** *matrix, has a unique solution.*

Proof. We have to show that if $\mathbf{X}(t)$ and $\mathbf{Y}(t)$ are both solutions, then $\mathbf{X}(t) \equiv \mathbf{Y}(t)$. Put $\mathbf{Z}(t) = \mathbf{X}(t) - \mathbf{Y}(t)$. Then

$$\dot{\mathbf{Z}}(t) = \dot{\mathbf{X}}(t) - \dot{\mathbf{Y}}(t) \quad \text{and} \quad \mathbf{AZ}(t) = \mathbf{AX}(t) - \mathbf{AY}(t) = \mathbf{A}(\mathbf{X}(t) - \mathbf{Y}(t)).$$

Therefore $\mathbf{Z}(t) = \mathbf{AZ}(t)$. Also

$$\mathbf{Z}(t_0) = \mathbf{X}(t_0) - \mathbf{Y}(t_0) = \mathbf{0}.$$

Therefore $\mathbf{Z}(t)$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX}$ which vanishes at t_0. By Corollary 4.22 we conclude that $\mathbf{Z}(t) = \mathbf{0}\,\forall t$, i.e. $\mathbf{X}(t) \equiv \mathbf{Y}(t)$. ■

Corollary 4.24 *Let $\mathbf{\Phi}$ be a normalized fundamental matrix for the system $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0. Then $\chi(t) = \mathbf{\Phi}(t)\chi_0$ is the unique solution of the matrix initial value problem $\dot{\chi} = \mathbf{A}\chi$, $\chi(t_0) = \chi_0$.*

The uniqueness theorem for the initial value problem $\dot{\mathbf{X}} = \mathbf{AX}$, $\mathbf{X}(t_0) = \mathbf{X}_0$, may be reformulated in terms of fundamental matrices.

Theorem 4.25 *There exists one and only one normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0.*

Exercise 4.26 Prove Corollary 4.24 and Theorem 4.25.

[1] This theorem must not be confused with the Existence and Uniqueness Theorem for linear systems of the form $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$, i.e. the case where the entries of $\mathbf{A}$ vary with time t (Chapter 7).

4.3 APPLICATIONS OF THE UNIQUENESS THEOREM. SOLUTION OF THE NON–HOMOGENEOUS PROBLEM

The following theorems, which deal with the properties of fundamental matrices, which may be normalized, are proved by using the Uniqueness Theorem 4.23.

Theorem 4.27 *If $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at $t_0 = 0$, then*

$$\mathbf{\Phi}(s+t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}(s) \tag{4.4}$$

for all real numbers s and t.

Proof. Let s be fixed, say $s = s_0$. We show that both sides of (4.4) are solutions of the matrix initial value problem

$$\dot{\chi} = \mathbf{A}\chi, \quad \chi(0) = \mathbf{\Phi}(s_0). \tag{4.5}$$

From Corollary 4.24, we have that $\chi(t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}(s_0)$ satisfies (4.5), i.e.

$$\frac{d}{dt}\left[\mathbf{\Phi}(t)\,\mathbf{\Phi}(s_0)\right] = \mathbf{A}\left[\mathbf{\Phi}(t)\,\mathbf{\Phi}(s_0)\right] \text{ and } \chi(0) = \mathbf{\Phi}(s_0).$$

Since $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at $t_0 = 0$, we have, from Theorem 4.5,

$$\frac{d}{dt}\left[\mathbf{\Phi}(t)\right] = \mathbf{A\Phi}(t) \quad \forall t,$$

so that in particular

$$\frac{d}{dt}\left[\mathbf{\Phi}(s_0+t)\right] = \mathbf{A\Phi}(s_0+t) \quad \forall t.$$

Moreover $\mathbf{\Phi}(s_0+t) = \mathbf{\Phi}(s_0)$ at $t = 0$. Consequently $\chi(t) = \mathbf{\Phi}(s_0+t)$ is a solution of (4.5). Since the solution of (4.5) is unique, $\mathbf{\Phi}(s_0+t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}(s_0)$ holds for arbitrary s_0 and all real numbers t. Therefore

$$\mathbf{\Phi}(s+t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}(s) \text{ for all real numbers } s \text{ and } t.$$

■

Exercise 4.28

1. Show by means of an example that the hypothesis that $\mathbf{\Phi}$ is a normalized fundamental matrix at $t_0 = 0$, cannot be scrapped.

2. Prove that under the hypothesis of Theorem 4.27,

   $$\mathbf{\Phi}(t)\,\mathbf{\Phi}(s) = \mathbf{\Phi}(s)\,\mathbf{\Phi}(t) \text{ for all real numbers } s \text{ and } t.$$

   Note that $\mathbf{\Phi}(0) = \mathbf{I}$ follows from the definition, so that $\mathbf{\Phi}$ indeed displays the properties of the exponential function.

Corollary 4.29 *If $\mathbf{\Phi}$ is a normalized fundamental matrix at $t_0 = 0$, then $\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}(-t)$ for all t.*

Corollary 4.30 *For any fundamental matrix* $\mathbf{\Phi}$ of $\dot{\mathbf{X}} = \mathbf{AX}$

$$\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}^{-1}(0)\,\mathbf{\Phi}(-t)\,\mathbf{\Phi}^{-1}(0)$$

holds for all t.

Remark
This formula provides a method for determining the inverse of a fundamental matrix of a system of differential equations.

Exercise 4.31

1. Prove Corollaries 4.29 and 4.30.

2. Verify Corollary 4.30 for the fundamental matrix $\mathbf{\Phi}(t) = \begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}$.

Theorem 4.32 (Transition Property) *Suppose the system* $\dot{\mathbf{X}} = \mathbf{AX}$ *has the fundamental matrices* $\mathbf{\Phi}_0$ *and* $\mathbf{\Phi}_1$. *Suppose* $\mathbf{\Phi}_0$ *and* $\mathbf{\Phi}_1$ *are respectively normalized at* t_0 *and* t_1. *Then*

$$\mathbf{\Phi}_1(t)\,\mathbf{\Phi}_0(t_1) = \mathbf{\Phi}_0(t) \text{ for each value of } t.$$

Corollary 4.33 *With the notation the same as in the previous theorem we have*

$$\mathbf{\Phi}_1(t_0)\,\mathbf{\Phi}_0(t_1) = \mathbf{I}.$$

Theorem 4.34 *If* $\mathbf{\Phi}$ *is a normalized fundamental matrix of* $\dot{\mathbf{X}} = \mathbf{AX}$ *at* t_0, *then*

$$\mathbf{A\Phi}(t) = \mathbf{\Phi}(t)\,\mathbf{A}.$$

Exercise 4.35

1. Show by means of an example that the hypothesis that $\mathbf{\Phi}$ is a *normalized* fundamental matrix, is indispensable.

2. Let $\mathbf{y}(t) = -\mathbf{x}(t)$. Show that $\dot{\mathbf{y}}(t) = \mathbf{Ay}(t)$ if $\dot{\mathbf{x}}(t) = -\mathbf{Ax}(t)$.

3. Suppose $\mathbf{A}$ is symmetric and $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{x}} = \mathbf{Ax}$ at $t_0 = 0$. Prove that $\mathbf{\Phi}^T(t) = \mathbf{\Phi}(t)$. (Hint: Consider the system $\dot{\mathbf{x}} = -\mathbf{Ax}$ and use Problem 2 above and $\mathbf{\Psi}^t\mathbf{\Phi} = \mathbf{I}$ and $\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}(-t)$.)

4.4 THE NON–HOMOGENEOUS PROBLEM $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$: VARIATION OF PARA–METERS

In this section, we want to derive a formula which, under suitable conditions on $\mathbf{F}$, yields a general solution of the non–homogeneous system $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$, when a general solution of $\dot{\mathbf{X}} = \mathbf{AX}$ (known as the complementary homogeneous system), is known. The fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ plays an important role in this formula. The method used is known as the *Method of Variation of Parameters*. (See also pp

351 – 354 of Z&W. On pp 156 – 162 of Z&W see also how a one–dimensional non–homogeneous differential equation can be solved by the method of variation of parameters.)

In the case of the n–dimensional problem $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$, the part played by the exponential function in the one–dimensional problem, is played by a fundamental matrix of the complementary system $\dot{\mathbf{X}} = \mathbf{AX}$. We assume the existence of a vector $\mathbf{U}$ such that

$$\mathbf{X}_p(t) = \mathbf{\Phi}(t)\,\mathbf{U}(t). \tag{4.6}$$

is a solution of

$$\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t).$$

Then

$$\begin{aligned}\dot{\mathbf{X}}_p(t) &= \dot{\mathbf{\Phi}}(t)\,\mathbf{U}(t) + \mathbf{\Phi}(t)\,\dot{\mathbf{U}}(t)\\ &= \mathbf{A\Phi}(t)\,\mathbf{U}(t) + \mathbf{\Phi}(t)\,\dot{\mathbf{U}}(t)\end{aligned}$$

since $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$.
Also

$$\dot{\mathbf{X}}_p(t) = \mathbf{A}\left[\mathbf{\Phi}(t)\,\mathbf{U}(t)\right] + \mathbf{F}(t).$$

Consequently

$$\mathbf{\Phi}(t)\,\dot{\mathbf{U}}(t) = \mathbf{F}(t),$$

so that

$$\dot{\mathbf{U}}(t) = \mathbf{\Phi}^{-1}(t)\,\mathbf{F}(t)$$

since $\mathbf{\Phi}^{-1}$ exists. We therefore have

$$\mathbf{X}_p(t) = \mathbf{\Phi}(t)\,\mathbf{U}(t) = \mathbf{\Phi}(t)\left\{\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds\right\}.$$

The above gives rise to

Theorem 4.36 *If $\mathbf{F}(t)$ is continuous and $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$, then the inhomogeneous system $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$ has solution*

$$\mathbf{X}_p(t) = \mathbf{\Phi}(t)\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds. \tag{4.7}$$

This solution is called a *particular* solution.
Proof. We have to show that $\dot{\mathbf{X}}_p(t) = \mathbf{AX}_p(t) + \mathbf{F}(t)$. Since $\mathbf{\Phi}$ is a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$, we have from Theorem 4.5 that $\dot{\mathbf{\Phi}}(t) = \mathbf{A\Phi}(t)$. We therefore have

$$\begin{aligned}\dot{\mathbf{X}}_p(t) &= \mathbf{A\Phi}\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds + \mathbf{\Phi}(t)\frac{d}{dt}\left(\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds\right)\\ &= \mathbf{A\Phi}(t)\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds + \mathbf{\Phi}(t)\left(\mathbf{\Phi}^{-1}(t)\,\mathbf{F}(t)\right)\\ &= \mathbf{A\Phi}(t)\int_{t_0}^{t}\mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds + \mathbf{F}(t)\\ &= \mathbf{AX}_p(t) + \mathbf{F}(t)\end{aligned}$$

by making use of the differentiation formula

$$\frac{\partial}{\partial t}\left(\int_{H(t)}^{F(t)} G(t,s)\,ds\right) = G(t,F(t))\,F'(t) - G(t,H(t))\,H'(t) + \int_{H(t)}^{F(t)}\left\{\frac{\partial}{\partial t}(G(t,s))\right\}ds \tag{4.8}$$

where F, G and H satisfy suitable conditions. ■

Example 4.37 Determine a solution of

$$\dot{\mathbf{X}} = \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix}\mathbf{X} + \begin{bmatrix} e^t \\ 1 \end{bmatrix} \equiv \mathbf{AX} + \mathbf{F}.$$

Solution

From (4.6), $\mathbf{X}_p(t) = \mathbf{\Phi}(t)\,\mathbf{U}(t)$ is a solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$, where $\mathbf{\Phi}(t)\,\dot{\mathbf{U}}(t) = \mathbf{F}(t)$ with $\mathbf{\Phi}$ a fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$. A fundamental matrix is

$$\begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}$$

so that we must solve for $\mathbf{U}$ from

$$\begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}\dot{\mathbf{U}} = \begin{bmatrix} e^t \\ 1 \end{bmatrix}.$$

If

$$\dot{\mathbf{U}} = \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix},$$

we have

$$\begin{aligned} 2e^t\dot{u}_1 &= e^t && (1) \\ e^t\dot{u}_1 + e^{3t}\dot{u}_2 &= 1. && (2) \end{aligned}$$

Now (1) is satisfied by $u_1 = \dfrac{t}{2}$ while (2), which reduces to

$$\dot{u}_2 = \left(1 - \frac{e^t}{2}\right)e^{-3t},$$

is satisfied by

$$u_2 = \frac{e^{-3t}}{-3} + \frac{e^{-2t}}{4}.$$

Consequently

$$\mathbf{U} = \begin{bmatrix} \frac{t}{2} \\ \frac{e^{-3t}}{-3} + \frac{e^{-2t}}{4} \end{bmatrix}$$

and

$$\mathbf{\Phi}(t)\,\mathbf{U}(t) = \begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}\begin{bmatrix} \frac{t}{2} \\ \frac{e^{-3t}}{-3} + \frac{e^{-2t}}{4} \end{bmatrix} = \begin{bmatrix} te^t \\ \left(\frac{t}{2} + \frac{1}{4}\right)e^t - \frac{1}{3} \end{bmatrix}$$

is a solution of the given problem.

Example 4.38 Use Theorem 4.36 to find the general solution of the inhomogeneous system

$$\dot{\mathbf{X}} = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 1 \\ \cot t \end{bmatrix}$$

where $t_0 = \dfrac{\pi}{2}$.

Solution

The characteristic equation $C(\lambda) = 0$ has roots $\lambda = \pm i$. An eigenvector corresponding to $\lambda = i$ is

$$\mathbf{U} = \begin{bmatrix} 1 \\ 1+i \end{bmatrix}.$$

From this we obtain, with the aid of Theorem 2.19, the two real solutions

$$\begin{bmatrix} \cos t \\ \cos t - \sin t \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \sin t \\ \cos t + \sin t \end{bmatrix}.$$

A fundamental matrix is therefore

$$\mathbf{\Phi}(t) = \begin{bmatrix} \cos t & \sin t \\ \cos t - \sin t & \cos t + \sin t \end{bmatrix}.$$

Now

$$\mathbf{\Phi}^{-1}(t) = \begin{bmatrix} \cos t + \sin t & -\sin t \\ \sin t - \cos t & \cos t \end{bmatrix}.$$

Therefore, by choosing $t_0 = \dfrac{\pi}{2}$, we have from Theorem 4.36

$$\begin{aligned} \mathbf{X}_p(t) &= \mathbf{\Phi}(t) \int_{\frac{\pi}{2}}^{t} \begin{bmatrix} \cos s + \sin s & -\sin s \\ \sin s - \cos s & \cos s \end{bmatrix} \begin{bmatrix} 1 \\ \cot s \end{bmatrix} ds \\ &= \mathbf{\Phi}(t)\,\mathbf{J} \end{aligned}$$

where

$$\begin{aligned} \mathbf{J} &= \int_{\frac{\pi}{2}}^{t} \begin{bmatrix} \sin s \\ -\cos s + \operatorname{cosec} s \end{bmatrix} ds \\ \left(\text{recall } \frac{\cos^2 s}{\sin s}\right. &= \left.\frac{1-\sin^2 s}{\sin s} = \operatorname{cosec} s - \sin s\right) \\ &= \begin{bmatrix} -\cos s \\ -\sin s - \ln|\operatorname{cosec} s + \cot s| \end{bmatrix}_{\frac{\pi}{2}}^{t} \\ &= \begin{bmatrix} -\cos t \\ -\sin t - \ln|\operatorname{cosec} t + \cot t| + 1 \end{bmatrix}. \end{aligned}$$

This yields

$$\begin{aligned} \mathbf{X}_p(t) &= \mathbf{\Phi}(t)\,\mathbf{J} \\ &= \begin{bmatrix} -1 + \sin t\,(1 - \ln|\operatorname{cosec} t + \cot t|) \\ -1 + (\cos t + \sin t)(1 - \ln|\operatorname{cosec} t + \cot t|) \end{bmatrix}. \end{aligned}$$

Remark
The inverse $\mathbf{\Phi}^{-1}(t)$ of the matrix

$$\mathbf{\Phi}(t)=\begin{bmatrix}\cos t & \sin t\\ \cos t-\sin t & \cos t+\sin t\end{bmatrix}$$

could quite easily be determined by recalling that the inverse of the matrix

$$\mathbf{S}=\begin{bmatrix}A & B\\ C & D\end{bmatrix}$$

is given by

$$\mathbf{S}^{-1}=\frac{1}{AD-BC}\begin{bmatrix}D & -B\\ -C & A\end{bmatrix},\text{ provided that } AD-BC\neq 0.$$

If $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{X}}=\mathbf{AX}$ at $t=0$, the formula in Theorem 4.36 can be changed slightly by applying Corollary 4.29.

Corollary 4.39 *If $\mathbf{\Phi}$ is a normalized fundamental matrix of $\dot{\mathbf{X}}=\mathbf{AX}$ at $t=0$, then*

$$\mathbf{X}_p(t)=\int_0^t\mathbf{\Phi}(t-s)\,\mathbf{F}(s)\,ds \tag{4.9}$$

is a solution of $\dot{\mathbf{X}}=\mathbf{AX}+\mathbf{F}$.

Remark
Note that the solution $\mathbf{X}_p(t)$ of $\dot{\mathbf{X}}=\mathbf{AX}+\mathbf{F}$, given by (4.7) or (4.9) vanishes at t_0, since the integral becomes zero at $t_0=0$.

Example 4.40 Use Corollary 4.39 to solve the inhomogeneous system

$$\dot{\mathbf{X}}=\begin{bmatrix}4 & 5\\ -2 & -2\end{bmatrix}\mathbf{X}+\begin{bmatrix}4e^t\cos t\\ 0\end{bmatrix}$$

by taking $t_0=0$.

Solution
The corresponding homogeneous system has the characteristic equation

$$C(\lambda)=\lambda^2-2\lambda+2=0$$

which yields the roots $\lambda=1\pm i$.

$1+i$: Solve

$$\begin{bmatrix}3-i & 5\\ -2 & -3-i\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

or equivalently

$$\begin{aligned}(3-i)\,v_1+\qquad 5v_2&=0 \qquad (1)\\ -2v_1+(3+i)\,v_2&=0. \qquad (2)\end{aligned}$$

(Note that the second equation is identical to the first one, i.e. $-(3-i)/2\times(2)=(1)$.) Choose, for example, $v_1=5$, then $v_2=-3+i$, so that an eigenvector corresponding to $1+i$ is

$$\mathbf{V}=\begin{bmatrix}5\\-3+i\end{bmatrix}=\begin{bmatrix}5\\-3\end{bmatrix}+i\begin{bmatrix}0\\1\end{bmatrix}.$$

Thus we get the two real solutions of the corresponding homogeneous problem

$$\begin{aligned}\mathbf{X}_1(t)&=e^t\left(\begin{bmatrix}5\\-3\end{bmatrix}\cos t-\begin{bmatrix}0\\1\end{bmatrix}\sin t\right)\\&=e^t\begin{bmatrix}5\cos t\\-3\cos t-\sin t\end{bmatrix}\end{aligned}$$

and

$$\begin{aligned}\mathbf{X}_2(t)&=e^t\left(\begin{bmatrix}5\\-3\end{bmatrix}\sin t+\begin{bmatrix}0\\1\end{bmatrix}\cos t\right)\\&=e^t\begin{bmatrix}5\sin t\\-3\sin t+\cos t\end{bmatrix}.\end{aligned}$$

A fundamental matrix is therefore

$$\mathbf{\Phi}(t)=e^t\begin{bmatrix}5\cos t&5\sin t\\-3\cos t-\sin t&-3\sin t+\cos t\end{bmatrix}.$$

Hence

$$\mathbf{\Phi}(0)=\begin{bmatrix}5&0\\-3&1\end{bmatrix}$$

and thus

$$\mathbf{\Phi}^{-1}(0)=\frac{1}{5}\begin{bmatrix}1&0\\3&5\end{bmatrix}$$

so that the *normalized* fundamental matrix is given by

$$\mathbf{\Psi}(t)=\mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(0)=e^t\begin{bmatrix}\cos t+3\sin t&5\sin t\\-2\sin t&-3\sin t+\cos t\end{bmatrix}.$$

From Corollary 4.39 it therefore follows that

$$
\begin{aligned}
\mathbf{X}_p(t) &= \int_0^t \mathbf{\Psi}(t-s)\,\mathbf{F}(s)\,ds \\
&= \int_0^t e^{t-s}\begin{bmatrix} \cos(t-s)+3\sin(t-s) & 5\sin(t-s) \\ -2\sin(t-s) & -3\sin(t-s)+\cos(t-s) \end{bmatrix}\begin{bmatrix} 4e^s\cos s \\ 0 \end{bmatrix} ds \\
&= 4e^t\int_0^t \begin{bmatrix} \cos(t-s)\cos s+3\sin(t-s)\cos s \\ -2\sin(t-s)\cos \end{bmatrix} ds \\
&= 2e^t\int_0^t \begin{bmatrix} \cos t+\cos(t-2s)+3\sin t+3\sin(t-2s) \\ -2\sin t-2\sin(t-2s) \end{bmatrix} ds \\
&= 2e^t\begin{bmatrix} s\cos t+3s\sin t-\frac{1}{2}\sin(t-2s)+\frac{3}{2}\cos(t-2s) \\ -2s\sin t-\cos(t-2s) \end{bmatrix}_0^t \\
&= 2e^t\begin{bmatrix} t\cos t+3t\sin t+\sin t \\ -2t\sin t \end{bmatrix}.
\end{aligned}
$$

Remark

Note that the fundamental matrix $\mathbf{\Phi}(t)$ in Corollary 4.39 is *normalized* while the one in Theorem 4.36 is *not necessarily* normalized. Also, in Theorem 4.36 the point t_0 is arbitrary, while in Corollary 4.39, $t_0 = 0$.

Exercise 4.41

1. Find a particular solution of the problem in Example 4.37 by using formula (4.7). Show that the answer obtained and the solution already found, differ by $\left[0, \frac{e^{3t}}{12}\right]^T$ — a solution of the complementary system $\dot{\mathbf{X}} = \mathbf{AX}$.
2. Prove that the difference between any two particular solutions of a non–homogeneous system, is a solution of the complementary homogeneous system.

From this it follows that the solutions of the non–homogeneous system $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$ may be expressed as $\mathbf{X}(t) = \mathbf{X}_c(t) + \mathbf{X}_p(t)$ with $\mathbf{X}_p(t)$ a particular solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$ and $\mathbf{X}_c(t)$ a general solution of $\dot{\mathbf{X}} = \mathbf{AX}$. As $\mathbf{X}_c(t)$ assumes all solutions of $\dot{\mathbf{X}} = \mathbf{AX}$, it follows that $\mathbf{X}(t)$ assumes all solutions of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$. As a result of Definition 4.21, we define a general solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$ as follows:

Definition 4.42 *Suppose that* $\mathbf{X}_p(t)$ *is any particular solution of the non–homogeneous system* $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$ *and* $\mathbf{\Phi}$ *a fundamental matrix of the complementary system* $\dot{\mathbf{X}} = \mathbf{AX}$. *Then the vector function*

$$\mathbf{X}(t) = \mathbf{X}_p(t) + \mathbf{\Phi}(t)\,\mathbf{K} \tag{4.10}$$

with $\mathbf{K}$ *a constant vector, is called a* **general solution** *of* $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$.

Theorem 4.43 (Uniqueness Theorem for the Non–Homogeneous Initial Value Problem $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F},\ \mathbf{X}(t_0) = \mathbf{X}_0$) *The problem*

$$\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F},\ \ \mathbf{X}(t_0) = \mathbf{X}_0 \tag{4.11}$$

has one and only one solution

$$\mathbf{X}(t) = \mathbf{X}_p(t) + \mathbf{\Phi}(t)\,\mathbf{X}_0 \tag{4.12}$$

with $\mathbf{\Phi}$ *a normalized fundamental matrix of the complementary system* $\dot{\mathbf{X}} = \mathbf{AX}$ at t_0 *and*

$$\mathbf{X}_p(t) = \mathbf{\Phi}(t)\int_{t_0}^{t} \mathbf{\Phi}^{-1}(s)\,\mathbf{F}(s)\,ds.$$

Proof. By differentiation it follows that the function given by equation (4.12), is a solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$, which also satisfies the initial condition. Suppose $\mathbf{Y}(t)$ is any solution of (4.11). We then have to prove that $\mathbf{X}(t) \equiv \mathbf{Y}(t)$. Define $\mathbf{Z}(t) = \mathbf{X}(t) - \mathbf{Y}(t)$. Then

$$\begin{aligned}
\dot{\mathbf{Z}}(t) &= \dot{\mathbf{X}}(t) - \dot{\mathbf{Y}}(t) \\
&= \mathbf{AX}(t) + \mathbf{F}(t) - \mathbf{AY}(t) - \mathbf{F}(t) \\
&= \mathbf{A}(\mathbf{X}(t) - \mathbf{Y}(t)) \\
&= \mathbf{AZ}(t)
\end{aligned}$$

and

$$\mathbf{Z}(t_0) = \mathbf{X}(t_0) - \mathbf{Y}(t_0) = \mathbf{0}.$$

The vector function $\mathbf{Z}(t)$ is, therefore, a solution of a homogeneous system which vanishes at t_0. From Corollary 4.22, $\mathbf{Z}(t)$ is identically zero, so that $\mathbf{X}(t) \equiv \mathbf{Y}(t)$. This completes the proof. ∎

Example 4.44 Find a general solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 3 \end{bmatrix}, \quad \mathbf{F}(t) = \begin{bmatrix} e^t \\ 1 \end{bmatrix}.$$

Solution
Take $t_0 = 0$.
A general solution of $\dot{\mathbf{X}} = \mathbf{AX}$ is

$$\mathbf{X}(t) = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^t + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{3t},$$

with c_1, c_2 arbitrary constants. A fundamental matrix is therefore given by

$$\mathbf{\Phi}(t) = \begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix}.$$

We find a particular solution by using Corollary 4.39. A normalized fundamental matrix at $t_0 = 0$ for the homogeneous system is given by

$$\begin{aligned}
\mathbf{\Psi}(t) &= \mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(0) \\
&= \begin{bmatrix} 2e^t & 0 \\ e^t & e^{3t} \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} e^t & 0 \\ \dfrac{e^t - e^{3t}}{2} & e^{3t} \end{bmatrix}.
\end{aligned}$$

From equation (4.9), a particular solution is

$$
\begin{aligned}
\mathbf{X}_p(t) &= \int_0^t \mathbf{\Psi}(t-s)\,\mathbf{F}(s)\,ds \\
&= \int_0^t \begin{bmatrix} e^{t-s} & 0 \\ \frac{e^{t-s}-e^{3(t-s)}}{2} & e^{3(t-s)} \end{bmatrix} \begin{bmatrix} e^{s} \\ 1 \end{bmatrix} ds \\
&= \int_0^t \begin{bmatrix} e^{t} \\ \frac{e^{t}-e^{3t-2s}}{2} + e^{3(t-s)} \end{bmatrix} ds \\
&= \begin{bmatrix} se^{t} \\ \frac{se^{t}}{2} - \frac{e^{3t-2s}}{-4} + \frac{e^{3(t-s)}}{-3} \end{bmatrix}_0^t .
\end{aligned}
$$

We therefore have the particular solution

$$
\mathbf{X}_p(t) = \begin{bmatrix} te^{t} \\ \frac{te^{t}}{2} + \frac{e^{t}}{4} - \frac{1}{3} - \frac{e^{3t}}{4} + \frac{e^{3t}}{3} \end{bmatrix} = \begin{bmatrix} te^{t} \\ \left(\frac{t}{2}+\frac{1}{4}\right)e^{t} + \left(\frac{e^{3t}}{12} - \frac{1}{3}\right) \end{bmatrix}.
$$

A general solution is, therefore,

$$
\mathbf{X}(t) = \begin{bmatrix} te^{t} \\ \left(\frac{t}{2}+\frac{1}{4}\right)e^{t} + \left(\frac{e^{3t}}{12} - \frac{1}{3}\right) \end{bmatrix} + c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{3t},
$$

with c_1, c_2 arbitrary constants.

Exercise 4.45 Do problems 11 – 33 of Exercise 8.3.2 on pp. 354 – 355 of Z&W.

Exercise 4.46

1. Use Theorem 4.36 to find a ***general solution*** of the inhomogeneous problem

$$
\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 0 \\ 2\cos t \end{bmatrix}. \text{ (Take } t_0 = 0.\text{)}
$$

2. Use Corollary 4.39 to solve the inhomogeneous initial value problem

$$
\dot{\mathbf{X}} = \begin{bmatrix} 2 & 1 \\ -4 & 2 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 3 \\ t \end{bmatrix} e^{2t}, \qquad \mathbf{X}(0) = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.
$$

4.5 THE INEQUALITY OF GRONWALL

In applications of differential equations, it is not always possible to obtain an exact solution of the differential equation. In such cases it is sometimes necessary to find, instead, an estimate for, or an upper bound of, the so–called *norm of the solution*. The latter quantity, as we will see below, measures the "size" of the solution.

In this section we treat the *inequality of Gronwall*, which is a useful tool in obtaining estimates for the norm of a solution of a differential equation of the form $\dot{\mathbf{X}} = \mathbf{AX}$.

Firstly we recall that if $\mathbf{X}(t)$ is a vector function, e.g.

$$\mathbf{X}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T,$$

then

$$\begin{aligned} ||\mathbf{X}(t)|| &= \sqrt{x_1^2(t) + x_2^2(t) + \ldots + x_n^2(t)} \\ &= \left(\sum_{i=1}^{n} x_i^2(t)\right)^{\frac{1}{2}}. \end{aligned}$$

Next we recall that for an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ the number $||\mathbf{A}||$ is defined [2] by

$$||\mathbf{A}|| = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}.$$

We show that

$$||\mathbf{A}|| = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$$

is indeed a norm. The system of all $n \times n$ matrices is then a normed linear space. If $\mathbf{A}$ denotes $[a_{ij}]$ and $\mathbf{B} = [b_{ij}]$, we have

(i) $||\mathbf{A}|| = 0 \Longleftrightarrow a_{ij} = 0 \quad \forall i = 1, \ldots, n; \quad j = 1, \ldots, n$
$\Longleftrightarrow \mathbf{A}$ is the zero matrix.

(ii) $$\begin{aligned} ||\lambda\mathbf{A}|| &= \left(\sum_{i,j=1}^{n} (\lambda a_{ij})^2\right)^{\frac{1}{2}} \\ &= |\lambda|\, ||\mathbf{A}|| \text{ for any scalar } \lambda. \end{aligned}$$

(iii) $$\begin{aligned} ||\mathbf{A} + \mathbf{B}|| &= \left(\sum_{i,j=1}^{n} (a_{ij} + b_{ij})^2\right)^{\frac{1}{2}} \\ &\leq \left(\sum_{i,j=1}^{n} a_{i,j}^2\right)^{\frac{1}{2}} + \left(\sum_{i,j=1}^{n} b_{ij}^2\right)^{\frac{1}{2}} \\ &= ||\mathbf{A}|| + ||\mathbf{B}|| \end{aligned}$$

where in the second last step we used the inequality of Minkowski.

Before formulating the inequality of Gronwall, we prove the following auxiliary result, which will be needed in proving Gronwall's inequality.

[2] The norm of $\mathbf{A}$ may also be defined in another way, as long as the properties of a norm are satisfied — see (i)–(iii) further on.

Lemma 4.47 *Suppose that f and g are continuous, real–valued functions in $a \leq t \leq b$ and that $\dot{f}$ exists in this interval. If, in addition, $\dot{f}(t) \leq f(t)\, g(t)$ in $a \leq t \leq b$, then*

$$f(t) \leq f(a) \exp\left[\int_a^t g(s)\, ds\right] \quad \forall t \in [a,b]. \tag{4.13}$$

ProofLet $p(t) = \exp\left(-\int_a^t g(s)\, ds\right)$. Then

$$\begin{aligned} \dot{p}(t) &= \exp\left(-\int_a^t g(s)\, ds\right) \cdot \frac{d}{dt}\left(-\int_a^t g(s)\, ds\right) \\ &= \exp\left(-\int_a^t g(s)\, ds\right)[-g(t)] \\ &= -p(t)\, g(t) \end{aligned}$$

and $p(a) = 1$. Now let $F(t) = f(t)p(t)$. Then

$$\begin{aligned} F' &= f'p + fp' \\ &= f'p - fpg \\ &= p(f' - gf) \leq 0 \end{aligned}$$

since $p > 0$ and $f' - fg \leq 0$. Hence F is nonincreasing for $a \leq t \leq b$ and therefore $F(a) \geq F(t)$. But $F(a) = f(a)$ since $p(a) = 1$. Hence we have

$$F(a) = f(a) \geq f(t)p(t) = F(t) = f(t)exp\left[-\int_a^t g(s)ds\right]$$

which yields the desired result. ■

We now proceed to

Theorem 4.48 (The Inequality of Gronwall) *If f and g $(g \geq 0)$ are continuous, real–valued functions in $a \leq t \leq b$, and K a real constant such that*

$$f(t) \leq K + \int_a^t f(s)\, g(s)\, ds, \quad a \leq t \leq b, \tag{4.14}$$

then

$$f(t) \leq K \exp\left[\int_a^t g(s)\, ds\right], \quad a \leq t \leq b. \tag{4.15}$$

Proof. Put

$$G(t) = K + \int_a^t f(s)\, g(s)\, ds.$$

Then $f(t) \leq G(t)$ in $[a,b]$. G is now differentiable in $[a,b]$. Indeed $G'(t) = f(t)\, g(t)$. Since $g \geq 0$ and $f \leq G$, we have $fg \leq Gg$; therefore $G' \leq Gg$. All the requirements of Lemma 4.47 are satisfied with the role of f assumed by G. Consequently we have, according to inequality (4.13),

$$G(t) \leq G(a) \exp\left(\int_a^t g(s)\, ds\right), \quad a \leq t \leq b.$$

Now

$$G(a) = K + \int_a^a f(s)\, g(s)\, ds = K.$$

Therefore

$$G(t) \leq K \exp\left[\int_a^t g(s)\, ds\right], \quad a \leq t \leq b.$$

The result now follows since $f(t) \leq G(t)$ in $[a, b]$. ■

Example 4.49

(a) Show that the only continuous function, f, satisfying

$$0 \leq f(t) \leq \int_a^t f(s) ds, \quad t_0 \leq t$$

is the identically zero function.

Solution Set $g(s) = 1$ and $K = 0$ in equation (4.14). Since all the conditions of Theorem 4.48 are satisfied, the inequality in equation (4.15) yields the desired result.

(b) Use the Gronwall inequality to prove that only the identically zero function is a solution of the initial value problem

$$\dot{x} = g(t)x, \quad x(t_0) = 0$$

with g continuous in $t \geq t_0$.

Solution By integrating the above equation we get

$$x(t) = \int_{t_0}^t g(s)x(s)ds$$

for any function $x(t)$ which satisfies the initial condition. Hence we get

$$0 \leq |x(t)| = \left|\int_{t_0}^t g(s)x(s)ds\right| \leq \int_{t_0}^t |g(s)||x(s)|ds.$$

Now set $f(t) = |x(t)|$ and $g(t) = |g(t)|$ and $K = 0$ in Theorem 4.48. Then the inequality in equation (4.15) implies that $x(t) = 0$.

The technique used in these examples is applied to systems of differential equations in the following section. As in the case of the one–dimensional equation, the procedure is to replace the differentiable equation by an equation containing an integral, and then to take the norm.

4.6 THE GROWTH OF SOLUTIONS

The Gronwall Inequality may be used to obtain an *estimate of the norm of solutions* of the system $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$, where $\mathbf{A}$ is a constant matrix.

Theorem 4.50 *If $\mathbf{X}(t)$ is any solution of $\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}$, then the following inequality holds for all t and t_0:*

$$||\mathbf{X}(t)|| \leq ||\mathbf{X}(t_0)|| \exp(||\mathbf{A}||\,|t - t_0|) \tag{4.16}$$

with equality for $t = t_0$.

Proof. If $t=t_0$, the assertion holds. If $t>t_0$, we have $|t-t_0|=t-t_0$, so that the inequality (4.16) reduces to

$$||\mathbf{X}(t)|| \leq ||\mathbf{X}(t_0)|| \exp\left[||\mathbf{A}||(t-t_0)\right].$$

By integrating $\dot{\mathbf{X}}(t)=\mathbf{A}\mathbf{X}(t)$, we obtain

$$\mathbf{X}(t)-\mathbf{X}(t_0)=\int_{t_0}^{t}\mathbf{A}\mathbf{X}(s)\,ds.$$

From the properties of norms we have

$$||\mathbf{X}(t)|| \leq ||\mathbf{X}(t_0)|| + \int_{t_0}^{t}||\mathbf{A}||\;||\mathbf{X}(s)||\,ds. \tag{4.17}$$

By putting $f(t)=||\mathbf{X}(t)||$, $g(t)=||\mathbf{A}||$ and $K=||\mathbf{X}(t_0)||$, it follows from Theorem 4.48 that

$$||\mathbf{X}(t)|| \leq ||\mathbf{X}(t_0)|| \exp\left(\int_{t_0}^{t}||\mathbf{A}||\,ds\right) = ||\mathbf{X}(t_0)|| \exp\left[||\mathbf{A}||(t-t_0)\right].$$

We still have to prove the theorem for $t<t_0$. In this case, put $\mathbf{Y}(t)=\mathbf{X}(-t)$. If we assume that $\mathbf{X}(t)$ is a solution of $\dot{\mathbf{X}}=\mathbf{A}\mathbf{X}$, it follows that

$$\dot{\mathbf{Y}}(t)=\dot{\mathbf{X}}(-t)=-\mathbf{A}\mathbf{Y}(t).$$

Therefore $\mathbf{Y}(t)$ is a solution of $\dot{\mathbf{Y}}=-\mathbf{A}\mathbf{Y}$. Moreover $-t>-t_0$. Therefore, by applying (4.17) to the system $\dot{\mathbf{Y}}=-\mathbf{A}\mathbf{Y}$, we obtain

$$\begin{aligned}
||\mathbf{X}(t)|| &= ||\mathbf{Y}(-t)|| \\
&\leq ||\mathbf{Y}(-t_0)|| \exp\left[||-\mathbf{A}||(-t+t_0)\right] \\
&= ||\mathbf{X}(t_0)|| \exp\left[||\mathbf{A}||(-t+t_0)\right] \\
&= ||\mathbf{X}(t_0)|| \exp\left[||\mathbf{A}||\;|t-t_0|\right],
\end{aligned}$$

since if $t<t_0$, then $|t-t_0|=-t+t_0$. This completes the proof. ■

Remark

If $t-t_0>0$, the inequality (4.16) shows that $||\mathbf{X}(t)||$ does not increase more rapidly than a certain exponential function. Such functions $\mathbf{X}(t)$ are known as functions of exponential type.

Example 4.51

The norms of a vector

$$\mathbf{X}(t)=(x_1(t),x_2(t),\ldots,x_n(t))^T$$

and an $n\times n$ matrix $\mathbf{A}=[a_{i,j}]$ were previously defined by

$$||\mathbf{X}||=\sqrt{\sum_{i=1}^{n}x_i^2} \quad \text{and} \quad ||\mathbf{A}||=\sqrt{\sum_{i,j}^{n}a_{ij}^2}.$$

Alternatively we can define

$$||\mathbf{X}||=\sum_{i=1}^{n}|x_i| \quad \text{and} \quad ||\mathbf{A}||=\max_{j=1,\ldots,n}\sum_{i=1}^{n}|a_{ij}|.$$

The latter definition means that the norm of a matrix $\mathbf{A}$ is defined as the maximum of the sums obtained by addition of the absolute values of the entries in each column of $\mathbf{A}$.

Now consider the system

$$\begin{aligned}\dot{x}_1 &= -(\sin t)\,x_2 + 4\\ \dot{x}_2 &= -x_1 + 2tx_2 - x_3 + e^t\\ \dot{x}_3 &= 3(\cos t)\,x_1 + x_2 + \frac{1}{t}x_3 - 5t^2.\end{aligned}$$

Assume that there exists solutions $\mathbf{X}$ and $\mathbf{Y}$ on the interval $(1, 3)$ with

$$\mathbf{X}(2) = [7;\ 3;\ -2]^T \quad \text{and} \quad \mathbf{Y}(2) = [6.7;\ 3.2;\ -1,9]^T.$$

Use the alternative definitions given above and Theorem 4.50 to estimate the error

$$||\mathbf{X}(t) - \mathbf{Y}(t)|| \quad \text{for } 1 < t < 3.$$

Solution

For the given system we have

$$\mathbf{A}(t) = \begin{bmatrix} 0 & -\sin t & 0\\ -1 & 2t & -1\\ 3\cos t & 1 & \dfrac{1}{t}\end{bmatrix}$$

so that

$$\begin{aligned}||\mathbf{A}(t)|| &= \max\left\{1 + 3|\cos t|,\ |\sin t| + 2t + 1,\ 1 + \frac{1}{t}\right\}\\ &\le \max\{4,\ 2 + 2t,\ 2\}\\ &= 8, \quad \text{since } 1 < t < 3.\end{aligned}$$

By Theorem 4.50 we have

$$||\mathbf{X}(t) - \mathbf{Y}(t)|| \le ||\mathbf{X}(2) - \mathbf{Y}(2)||\exp(||\mathbf{A}||\ \ |t-2|) < (0.3 + 0.2 + 0.1)\,e^8.$$

CHAPTER 5

Higher Order One–dimensional Equations as Systems of First Order Equations

Objectives of this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding higher order one–dimensional equations:

- the link between higher order one–dimensional equations and linear systems of first order differential equations;
- companion matrix;
- companion system for the n–th order equation with constant coefficients as given by (5.1).

Outcomes of this Chapter

After studying this chapter the learner should be able to

- rewrite the n–th order equation (5.1) as a linear system of first order differential equations;
- solve the corresponding linear system of first order differential equations and interpret the solution in terms of the original higher order one dimensional equation.

5.1 INTRODUCTION

In Chapter 4 of the Z&W the link that exists between higher order one–dimensional linear differential equations and linear systems of first order differential equations is explored. The student can revise Sections 4.1 through 4.6 of Z&W as it was studied in APM2611.

5.2 COMPANION SYSTEMS FOR HIGHER–ORDER ONE–DIMENSIONAL DIFFERENTIAL EQUATIONS

The n–th order equation with constant coefficients

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y^{(1)} + a_0 y = f(t) \tag{5.1}$$

has been treated in APM2611. This equation can be reduced to a system of n first order equations. We illustrate by means of an example:

Consider the equation

$$y^{(3)} - 6y^{(2)} + 5y^{(1)} = \sin t \tag{5.2}$$

with initial conditions $y(0) = y'(0) = y''(0) = 0$.

If $y(t)$ represents the distance of a body from a fixed point on a line, this equation describes the movement of that body along the line, subject to different forces. Since y measures distance, $y^{(1)}$ measures velocity and $y^{(2)}$ acceleration. By putting $y^{(1)} = v$ and $y^{(2)} = a$, equation (5.2) may be expressed as

$$\begin{aligned} y^{(1)} &= v \\ y^{(2)} &= a \\ y^{(3)} &= -5v + 6a + \sin t. \end{aligned} \tag{5.3}$$

This, in turn, may be expressed as

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{y} \\ \dot{v} \\ \dot{a} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & 6 \end{bmatrix} \begin{bmatrix} y \\ v \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \sin t \end{bmatrix} \tag{5.4}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & 6 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 0 \\ 0 \\ \sin t \end{bmatrix}. \tag{5.5}$$

Equation (5.5) is now merely a special case of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$ with the initial condition

$$\mathbf{X}(0) = \begin{bmatrix} y(0) \\ v(0) \\ a(0) \end{bmatrix} = \begin{bmatrix} y(0) \\ y^{(1)}(0) \\ y^{(2)}(0) \end{bmatrix} = \mathbf{0}.$$

More generally, the third order equation

$$y^{(3)} + a_2 y^{(2)} + a_1 y^{(1)} + a_0 y = f(t) \tag{5.6}$$

from which we obtain

$$y^{(3)} = -a_0 y - a_1 y^{(1)} - a_2 y^{(2)} + f(t),$$

can be expressed in matrix form as three first order equations as follows:

Let

$$\mathbf{X} = \begin{bmatrix} y \\ y^{(1)} \\ y^{(2)} \end{bmatrix}.$$

Then

$$\dot{\mathbf{X}} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \end{bmatrix}$$

so that

$$\dot{\mathbf{X}} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ y^{(3)} \end{bmatrix} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ -a_0 y - a_1 y^{(1)} - a_2 y^{(2)} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ f(t) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \mathbf{X} + f(t) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{5.7}$$

$$= \mathbf{AX} + \mathbf{F}(t)$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_2 & -a_3 \end{bmatrix}$$

and

$$\mathbf{F}(t) = f(t) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The system (5.7) is known as the *companion system* for the third order equation (5.6). The companion system for the n–th order equation (5.1) may be defined in a similar way.

Definition 5.1 *A matrix of the form*

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}$$

is known as a **companion matrix***.*

Definition 5.2 *The system*

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix} \mathbf{X} + f(t) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{5.8}$$

is known as the **companion system** *for the n–th order equation* (5.1)*.*

Theorem 5.3 *If y is a solution of equation* (5.1)*, then the vector*

$$\mathbf{X} = \begin{bmatrix} y \\ y^{(1)} \\ \vdots \\ y^{(n-1)} \end{bmatrix}$$

is a solution of the companion system (5.8)*. Conversely: if*

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is a solution of the system (5.8)*, then* x_1*, the first component of* $\mathbf{X}$*, is a solution of* (5.1) *and*

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_1^{(1)} \\ \vdots \\ x_1^{(n-1)} \end{bmatrix},$$

i.e.

$$x_2 = x_1^{(1)}, \;\; x_3 = x_1^{(2)}, \ldots, x_n = x_1^{(n-1)}.$$

Exercise 5.4

1. Prove Theorem 5.3 for the case $n = 3$.

2. Prove, by applying Theorem 5.3, that the initial value problem

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y^{(1)} + a_0 y = f(t),$$

$$y(t_0) = y_1, \; y^{(t)}(t_0) = y_2, \ldots, y^{(n-1)}(t_0) = y_n,$$

   has a *unique solution*.

We summarize: *Equations of order* n *in one dimension may be solved by determining a general solution*

$$\mathbf{X}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

of the companion system. The function $x_1(t)$ *is then a solution of the given* n*–th order equation.*

As the companion system is but a special case of the non–homogeneous linear system with constant coefficients, one needs no new techniques.

Example 5.5 Determine the general solution of

$$\frac{d^2y}{dt^2} + 4y = \sin 3t \tag{5.9}$$

by using the companion system.

Solution

The companion system is found as follows:

Rewrite (5.9) as

$$y^{(2)} = -4y + \sin 3t.$$

Let

$$\mathbf{X}=\begin{bmatrix} y \\ y^{(1)} \end{bmatrix},$$

then

$$\dot{\mathbf{X}}=\begin{bmatrix} y^{(1)} \\ y^{(2)} \end{bmatrix}$$

so that

$$\begin{aligned}
\dot{\mathbf{X}} &= \begin{bmatrix} y^{(1)} \\ y^{(2)} \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}\begin{bmatrix} y \\ y^{(1)} \end{bmatrix}+\begin{bmatrix} 0 \\ \sin 3t \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}\begin{bmatrix} y \\ y^{(1)} \end{bmatrix}+\sin 3t\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \mathbf{AX}+\mathbf{F}(t). \qquad (5.10)
\end{aligned}$$

We first determine a general solution of $\dot{\mathbf{X}}=\mathbf{AX}$. Now

$$\det\begin{bmatrix} -\lambda & 1 \\ -4 & -\lambda \end{bmatrix}=0\Rightarrow\lambda=\pm 2i.$$

Choose $\lambda=2i$. (Remember $\lambda=-2i$ yields identical solutions up to a constant). Now put

$$\begin{bmatrix} -2i & 1 \\ -4 & -2i \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then

$$\begin{aligned}
-2iu_1+u_2 &= 0 \\
-4u_1-2iu_2 &= 0.
\end{aligned}$$

These two equations are identical. Choose $u_1=1$, then $u_2=2i$. The vector

$$\mathbf{U}\equiv\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}=\begin{bmatrix} 1 \\ 2i \end{bmatrix}$$

is then an eigenvector corresponding to the eigenvalue $\lambda=2i$. Consequently

$$\mathbf{X}_1(t)=\left\{\cos 2t\begin{bmatrix} 1 \\ 0 \end{bmatrix}-\sin 2t\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right\}=\begin{bmatrix} \cos 2t \\ -2\sin 2t \end{bmatrix}$$

and

$$\mathbf{X}_2(t)=\left\{\sin 2t\begin{bmatrix} 1 \\ 0 \end{bmatrix}+\cos 2t\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right\}=\begin{bmatrix} \sin 2t \\ 2\cos 2t \end{bmatrix}$$

are solutions of $\dot{\mathbf{X}}=\mathbf{AX}$ and a general solution is

$$\mathbf{X}(t)=c_1\mathbf{X}_1(t)+c_2\mathbf{X}_2(t)$$

with c_1 and c_2 arbitrary constants.

A fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ is therefore

$$\mathbf{\Phi}(t) = \begin{bmatrix} \cos 2t & \sin 2t \\ -2\sin 2t & 2\cos 2t \end{bmatrix}$$

and a normalized fundamental matrix of $\dot{\mathbf{X}} = \mathbf{AX}$ at $t_0 = 0$ is

$$\mathbf{\Psi}(t) = \mathbf{\Phi}(t)\,\mathbf{\Phi}^{-1}(0) = \frac{1}{2}\begin{bmatrix} 2\cos 2t & \sin 2t \\ -4\sin 2t & 2\cos 2t \end{bmatrix}.$$

Therefore, a particular solution of the system (5.10) is

$$\begin{aligned}
\mathbf{X}_p(t) &= \frac{1}{2}\int_0^t \left\{ \begin{bmatrix} 2\cos 2(t-s) & \sin 2(t-s) \\ -4\sin 2(t-s) & 2\cos 2(t-s) \end{bmatrix} \begin{bmatrix} 0 \\ \sin 3s \end{bmatrix} \right\} ds \\
&= \frac{1}{2}\int_0^t \begin{bmatrix} \sin 2(t-s)\sin 3s \\ 2\cos 2(t-s)\sin 3s \end{bmatrix} ds \\
&= \frac{1}{2}\int_0^t \begin{bmatrix} \frac{1}{2}\cos(2t-5s) - \frac{1}{2}\cos(2t+s) \\ \sin(2t+s) + \sin(5s-2t) \end{bmatrix} ds \\
&= \frac{1}{2}\begin{bmatrix} \left(\frac{-\sin(2t-5s)}{10} - \frac{\sin(2t+s)}{2}\right) \\ \left(-\cos(2t+s) - \frac{\cos(5s-2t)}{5}\right) \end{bmatrix}_0^t \\
&= \frac{1}{2}\begin{bmatrix} \left(-\frac{2\sin 3t}{5} + \frac{3\sin 2t}{5}\right) \\ \left(-\frac{6\cos 3t}{5} + \frac{6\cos 2t}{5}\right) \end{bmatrix}.
\end{aligned}$$

Consequently the general solution of $\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{F}(t)$ is

$$\mathbf{X}(t) = c_1\begin{bmatrix} \cos 2t \\ -2\sin 2t \end{bmatrix} + c_2\begin{bmatrix} \sin 2t \\ 2\cos 2t \end{bmatrix} + \frac{1}{10}\begin{bmatrix} -2\sin 3t + 3\sin 2t \\ -6\cos 3t + 6\cos 2t \end{bmatrix}.$$

According to Theorem 5.3, a general solution of (5.9) is given by the first component of $\mathbf{X}(t)$. Therefore

$$y(t) = c_1\cos 2t + c_2\sin 2t - \frac{\sin 3t}{5} + \frac{3\sin 2t}{10}. \tag{5.11}$$

Note that the second component of $\mathbf{X}(t)$ is the derivative of $y(t)$.

Remark

This solution can be obtained more readily by applying the standard techniques for solving higher order one–dimensional equations: For the equation $\left(D^2+4\right)y = \sin 3t$, we have the auxiliary equation $m^2+4=0$ with roots $\pm 2i$. Hence

$$y_{\text{C.F.}}(t) = d_1\cos 2t + d_2\sin 2t.$$

Furthermore

$$y_{\text{P.I.}}(t) = \frac{\sin 3t}{-3^2+4} = -\frac{\sin 3t}{5}.$$

Hence

$$y(t) = d_1\cos 2t + d_2\sin 2t - \frac{\sin 3t}{5}.$$

This is equivalent to the solution (5.11).

Exercise 5.6

1. Write the companion systems for the equations given below.

   (a) $y^{(2)} - y^{(1)} = 0$ (b) $y^{(2)} - y = 0$ (c) $y^{(2)} = 0$ (d) $y^{(4)} - y^{(3)} = 0$

   (e) $y^{(4)} - y = 0$ (f) $y^{(3)} - (\sin t)y = f(t)$

2. Each of the matrices below are system's matrices for the companion system of an nth-order equation. Determine the nth-order equation in each case.

   (a) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$

CHAPTER 6

Analytic Matrices and Power Series Solutions of Systems of Differential Equations

Objectives for this Chapter

The main objectives of this chapter is to gain an understanding of the following concepts regarding power series solutions of differential equations:

- analytic matrices;
- power series expansions of analytic functions;
- power series solutions for $\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X}$;
- the exponential of a matrix.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- determine whether a matrix is analytic or not;
- solve the system $\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X}$ using power series methods.

6.1 INTRODUCTION

The system $\dot{\mathbf{X}} = \mathbf{AX}$, with $\mathbf{A}$ an $n \times n$ matrix with constant entries, is a special case of the more general system

$$\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X}, \quad a < t < b, \tag{6.1}$$

with $\mathbf{A}(t) = [a_{ij}(t)]$ a matrix whose entries are functions of time and which we will call continuous and real–valued if all the functions a_{ij} are continuous and real–valued. Everything that holds for the system $\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X}$, is therefore applicable to the case where $\mathbf{A}$ is constant. The converse, is, however, not easy. Whereas we have to our disposal standard techniques for finding n linearly independent solutions of $\dot{\mathbf{X}} = \mathbf{AX}$, each of which is usually a linear combination of known functions, while we also know that exactly n linearly independent solutions exist, the construction of only a single solution of $\dot{\mathbf{X}} = \mathbf{A}(t)\,\mathbf{X}$, usually entails an enormous amount of work. Moreover, it requires advanced mathematics to prove theoretically that n linearly independent solutions always exist.

A discussion of this theory, which is outside the scope of this course, is found, amongst others, in "Lectures on Ordinary Differential Equations" by W. Hurewicz[1]. It is sufficient to say that, as in the case of the equation $\dot{\mathbf{X}} = \mathbf{AX}$, fundamental and normalized fundamental matrices $\mathbf{\Phi}(t)$ of $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$, are used in the development of this theory.

Whereas the solution of $\dot{\mathbf{X}} = \mathbf{AX}$ are usually linear combinations of known functions such as $\sin t$ and $\cos t$, such linear combinations are rarely solutions of $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$. Indeed, systems $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$ are often used to define and to study new functions. For instance, the functions of Bessel, Legendre and Laguerre and the Hermite Functions (some of which you have encountered in APM2611), are examples of functions defined as solutions of certain differential equations.

The main result of this chapter is that, under certain conditions on the matrix $\mathbf{A}(t)$, viz that it is possible to write the entries of $\mathbf{A}(t)$ as power series, a solution of $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$ can be expressed as a power series

$$\mathbf{X}(t) = \mathbf{U}_0 + \mathbf{U}_1(t-t_0) + \ldots + \mathbf{U}_k(t-t_0)^k + \ldots \tag{6.2}$$

about t_0. If we wish to find a power series solution of $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$, the vectors $\mathbf{U}_0, \mathbf{U}_1, \ldots$ have to be determined, while we also have to show that the infinite sum (6.2) converges to a solution of $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}$. The property which $\mathbf{A}(t)$ must satisfy, is that of *analyticity*, a powerful property, since continuity and even the existence of derivatives of all orders, are not sufficient to ensure this property. However, in most cases, the coefficient matrices of differential equations, as well as of systems that are important for all practical purposes, do have the property of analyticity.

6.2 POWER SERIES EXPANSIONS OF ANALYTIC FUNCTIONS

Although the property of analyticity is usually associated with functions of a complex variable, we can define the property of analyticity for real functions, in view of the fact that a function of a real variable is only a special case of a function of a complex variable.

Definition 6.1 *A function F of a real variable is* **analytic at a real number** t_0 *if a positive number r and a sequence of numbers $a_0, a_1, a_2, \ldots$ exist such that*

$$F(t) = \lim_{N\to\infty} \sum_{k=0}^{N} a_k (t-t_0)^k = \sum_{k=0}^{\infty} a_k (t-t_0)^k \quad \textit{for } |t-t_0| < r. \tag{6.3}$$

The following theorem deals with the differentiation of power series. *This is no trivial matter, since infinite sums are involved.*

Theorem 6.2 (Taylor's Theorem) *If F is analytic at t_0, and has series expansion*

$$F(t) = \sum_{k=0}^{\infty} a_k (t-t_0)^k \quad \text{for } |t-t_0| < r,$$

then

$$F^{(j)}(t) = \sum_{k=0}^{\infty} \frac{(k+j)!}{k!} a_{k+j} (t-t_0)^k \quad \text{for } |t-t_0| < r,$$

for every positive integer j.

[1]W. Hurewicz, Lectures on Ordinary Differential Equations, The Technology Press, Cambridge, Mass., 1958.

From the above it is clear that the j–th derivative $F^{(j)}$ is also analytic at t_0 and $F^{(j)}(t_0) = j!a_j$. The constants

$$a_k = \frac{1}{k!}F^{(k)}(t_0)$$

are known as the *Taylor coefficients* of F at t_0, and the expansion

$$\sum_{k=0}^{\infty} a_k (t - t_0)^k = \sum_{k=0}^{\infty} \frac{1}{k!}F^{(k)}(t_0)(t - t_0)^k$$

is the *Taylor series of F about t_0.*

Remark

(1) It is worthwhile selecting analytic functions to form a class of functions for the following reasons:

(a) It is possible that a function may not possess derivatives of all orders. (Can you think of a good example?)

(b) Functions F exist which have derivatives of all orders, but of which the Taylor series

$$\sum_{k=0}^{\infty} \frac{1}{k!}F^{(k)}(t_0)(t - t_0)^k$$

converges only at $t = t_0$. It may happen that a function F has a Taylor series expansion

$$\sum_{k=0}^{\infty} \frac{1}{k!}F^{(k)}(t_0)(t - t_0)^k$$

converging for $|t - t_0| < r$ for some $r > 0$, but not to $F(t)$, i.e.

$$F(t) \neq \sum_{k=0}^{\infty} \frac{1}{k!}F^{(k)}(t_0)(t - t_0)^k,$$

unless $t = t_0$.

(2) If F is a function of the complex variable z and F is analytic at $z = z_0$ (with the concept of analyticity defined as in Complex Analysis), a power series expansion of the form

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

exists for $F(z)$, viz the Laurent Series

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k + \sum_{k=1}^{\infty} b_k (z - z_0)^{-k},$$

with all the coefficients b_k of the "principal part"

$$\sum_{k=0}^{\infty} b_k (z - z_0)^{-k}$$

equal to zero.

In this case the coefficients a_k $(k = 0, 1, 2, \ldots)$ are given in terms of contour integrals by means of the formula

$$a_k = \frac{1}{2\pi i}\oint_C \frac{F(z)}{(z - z_0)^{k+1}}dz$$

with C a circle with centre z_0. This is, of course, a result from Complex Analysis, quoted here only for the sake of interest (as this result is dealt with in the module on Complex Analysis).

A list of Taylor series for some elementary functions is given below.

(1) $$e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!} \qquad \text{for all } t$$

(2) $$\sin t = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} \qquad \text{for all } t$$

(3) $$\cos t = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \qquad \text{for all } t$$

(4) $$\frac{1}{1-t} = \sum_{k=0}^{\infty} t^k \qquad -1 < t < 1$$

(5) $$\frac{1}{1+t} = \sum_{k=0}^{\infty} (-1)^k t^k \qquad -1 < t < 1$$

(6) $$\ln(1+t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{k+1}}{k+1} \qquad -1 < t \le 1$$

(7) $$\arctan t = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{2k+1} \qquad -1 \le t \le 1$$

The following theorem deals with addition, subtraction and multiplication of power series:

Theorem 6.3 *If*

$$F(t) = \sum_{k=0}^{\infty} F_k (t - t_0)^k, \quad -r < t - t_0 < r,$$

$$G(t) = \sum_{k=0}^{\infty} G_k (t - t_0)^k, \quad -r < t - t_0 < r$$

with F_k and G_k the Taylor Coefficients of F and G, then

$$F(t) \pm G(t) = \sum_{k=0}^{\infty} (F_k \pm G_k)(t - t_0)^k$$

and

$$F(t)\,G(t) = \sum_{k=0}^{\infty}\sum_{i=0}^{k} F_i G_{k-i}\,(t-t_0)^k \quad \text{for} \quad |t-t_0| < r.$$

The following theorem is often used.

Theorem 6.4 *If*

$$F(t) = \sum_{k=0}^{\infty} F_k\,(t-t_0)^k \equiv 0, \quad -r < t - t_0 < r,$$

then $F_0 = F_1 = \ldots = 0$ with F_k the Taylor coefficients of F.

Proof. The result follows immediately by recalling that if

$$F(t) \equiv 0, \;\; \text{i.e. } F(t) = 0 \textit{ for all } t,$$

then the derivatives of all orders are also zero. ■

Remark

$F(0) = 0$ does not imply $F'(0) = 0$! For example

$$F(x) = x \text{ implies } F(0) = 0, \text{ but } F'(0) = 1.$$

We now define analytic matrices:

Definition 6.5 *The $n \times m$ matrix $A(t) = [a_{ij}(t)]$ is said to be* **analytic at** $t = t_0$ *if* **every** *entry*

$$a_{ij}(t)\,, \quad i = 1, \ldots, n, j = 1, \ldots, m$$

is analytic at $t = t_0$.

In this definition it may happen that the intervals of convergence of the expansions for the separate entries differ from one another. The interval of convergence of $\mathbf{A}(t)$ is then taken as that interval in which all the entries converge.

The preceding theorems may now be reformulated for analytic $\mathbf{A}(t)$ by simply applying the relevant conditions to each entry of $\mathbf{A}(t)$ and then making the conclusion in a similar way.

The notation

$$\mathbf{A}(t) = \sum_{k=0}^{\infty} \mathbf{A}_k\,(t-t_0)^k\,,$$

with

$$\mathbf{A}_k = \frac{1}{k!}\mathbf{A}^{(k)}(t_0)\,, \quad |t-t_0| < r,$$

is, therefore, just an abbreviated notation for $n \times m$ expansions — one for each entry of $\mathbf{A}(t)$.

Example 6.6

Determine $\mathbf{A}_k$ such that

$$\begin{bmatrix} \cos t & \arctan t \\ e^t & \dfrac{1}{1-t} \end{bmatrix} = \sum_{k=0}^{\infty} \mathbf{A}_k t^k \equiv \mathbf{A}\left(t\right).$$

Solution

Note that $t_0 = 0$. Since $\cos t$ is an even function and $\arctan t$ is an odd function, we rewrite the series expansion

$$\sum_{k=0}^{\infty} \mathbf{A}_k t^k$$

as the sum of its even and odd expansions as follows:

$$\sum_{k=0}^{\infty} \mathbf{A}_k t^k = \sum_{i=0}^{\infty} \mathbf{A}_{2i} t^{2i} + \sum_{i=0}^{\infty} \mathbf{A}_{2i+1} t^{2i+1}.$$

Then

$$\mathbf{A}_{2i} = \begin{bmatrix} \dfrac{(-1)^{2i}}{(2i)!} & 0 \\ \dfrac{1}{(2i)!} & 1 \end{bmatrix}$$

and

$$\mathbf{A}_{2i+1} = \begin{bmatrix} 0 & \dfrac{(-1)^{2i}}{(2i+1)!} \\ \dfrac{1}{(2i+1)!} & 1 \end{bmatrix}$$

for $i = 1, 2, \ldots$ by using the expansions for $\cos t$, e^t and $\arctan t$. The common interval of convergence is $-1 < t < 1$, so that $\mathbf{A}\left(t\right)$ is analytic at $t_0 = 0$.

Exercise 6.7

Find $\mathbf{X}_k$ such that

$$\mathbf{X}(t) = \begin{bmatrix} e^t \\ e^{2t} \end{bmatrix} = \sum_{k=0}^{\infty} \mathbf{X}_k t^k.$$

6.3 SERIES SOLUTIONS FOR $\dot{\mathbf{X}} = \mathbf{A}\left(t\right)\mathbf{X}$

Consider the initial value problem

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}\left(t\right)\mathbf{X}, \\ \mathbf{X}\left(t_0\right) &= \mathbf{X}_0. \end{aligned} \tag{6.4}$$

Assume that $\mathbf{A}\left(t\right)$ is analytic at $t = t_0$.

According to a result which we do not prove, *a unique solution* $\mathbf{X}\left(t\right)$, analytic at t_0, exists for the initial value problem (6.4), i.e. $\mathbf{X}\left(t\right)$ has series expansion

$$\mathbf{X}\left(t\right) = \sum_{k=0}^{\infty} \mathbf{X}_k \left(t - t_0\right)^k. \tag{6.5}$$

If we want to find a solution for (6.4) in the form of the expansion (6.5), we have to determine formulas for $\mathbf{X}_k$. Since $\mathbf{A}(t)$ and $\mathbf{X}(t_0) = \mathbf{X}_0$ are known, we therefore determine the relation between the constant vectors $\mathbf{X}_k$, and $\mathbf{A}(t)$ and $\mathbf{X}_0$.

Since

$$\mathbf{X}_k = \frac{\mathbf{X}^{(k)}(t_0)}{k!}, \quad k = 0, 1, 2, \ldots \tag{6.6}$$

we must differentiate (6.5) $(k-1)$ times, and determine the value of the derivative at $t = t_0$ in each case. In order to do this, we use Leibniz's Rule, according to which

$$\frac{d^k}{dt^k}[\mathbf{A}(t)\mathbf{X}(t)] = \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} \mathbf{A}^{(i)}(t)\mathbf{X}^{(k-i)}(t) \tag{6.7}$$

with

$$\begin{pmatrix} k \\ i \end{pmatrix} = \frac{k!}{i!\,(k-i)!}.$$

From (6.6) we have

$$\begin{aligned}
\mathbf{X}_{k+1} = \frac{\mathbf{X}^{(k+1)}(t_0)}{(k+1)!} &= \frac{1}{(k+1)!}\left[\frac{d^k}{dt^k}\dot{\mathbf{X}}(t)\right]_{t=t_0} \\
&= \frac{1}{(k+1)!}\left[\frac{d^k}{dt^k}(\mathbf{A}(t)\mathbf{X}(t))\right]_{t=t_0} \\
&= \frac{1}{(k+1)!}\sum_{i=0}^{k}\begin{pmatrix} k \\ i \end{pmatrix}\mathbf{A}^{(i)}(t_0)\mathbf{X}^{(k-i)}(t_0)
\end{aligned}$$

from (6.7).

Since

$$\mathbf{A}(t) = \sum_{k=0}^{\infty}\frac{\mathbf{A}^{(k)}(t_0)}{k!}(t-t_0)^k$$

from Taylor's Theorem, we can, by putting

$$\mathbf{A}_k = \frac{\mathbf{A}^{(k)}(t_0)}{k!},$$

and using (6.6), write

$$\mathbf{X}_{k+1} = \frac{k!}{(k+1)!}\sum_{i=0}^{k}\mathbf{A}_i\mathbf{X}_{k-i},$$

i.e.

$$\mathbf{X}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k}\mathbf{A}_i\mathbf{X}_{k-i} \tag{6.8}$$

for $k = 0, 1, \ldots$ and $\mathbf{X}_0 = \mathbf{X}(t_0)$.

Since a unique solution of (6.4), analytic at t_0, exists, (see Chapter 7), we have the following theorem:

Theorem 6.8 *Suppose that* $\mathbf{A}(t)$ *is analytic at* t_0 *with a power series expansion*

$$\mathbf{A}(t) = \sum_{k=0}^{\infty} \mathbf{A}_k (t - t_0)^k, \quad |t - t_0| < r.$$

Let $\mathbf{X}_0$ *be an arbitrary vector. Then*

$$\mathbf{X}(t) = \sum_{k=0}^{\infty} \mathbf{X}_k (t - t_0)^k, \quad |t - t_0| < r$$

is a solution of (6.4)*, analytic at* t_0*, with*

$$\mathbf{X}_{k+1} = \frac{1}{k+1} \sum_{i=0}^{k} \mathbf{A}_i \mathbf{X}_{k-i},$$

i.e.

$$\begin{aligned}
\mathbf{X}_1 &= \mathbf{A}_0 \mathbf{X}_0 \\
\mathbf{X}_2 &= \frac{1}{2} \{\mathbf{A}_0 \mathbf{X}_1 + \mathbf{A}_1 \mathbf{X}_0\} \\
&\vdots \\
\mathbf{X}_{k+1} &= \frac{1}{k+1} \{\mathbf{A}_0 \mathbf{X}_k + \mathbf{A}_1 \mathbf{X}_{k-1} + \ldots + \mathbf{A}_k \mathbf{X}_0\}
\end{aligned} \tag{6.9}$$

Example 6.9 Determine a series solution of

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \mathbf{X} \equiv \mathbf{A}\mathbf{X}.$$

Compare your answer with the solution obtained by the eigenvalue–eigenvector method.

Solution

Since all entries are constants, the matrix $\mathbf{A}$ is analytic at $t = 0$. The sum

$$\sum_{k=0}^{\infty} \mathbf{A}_k (t - t_0)^k$$

can yield the constant matrix

$$\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$$

only if $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots$ are all the zero matrix. Hence, $\mathbf{A}_0 = \mathbf{A}$. From (6.9) we have

$$\begin{aligned}
\mathbf{X}_n &= \frac{1}{n} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \mathbf{X}_{n-1} \\
&= \frac{1}{n(n-1)} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}^2 \mathbf{X}_{n-2} \\
&= \ldots \\
&= \frac{1}{n!} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}^n \mathbf{X}_0.
\end{aligned}$$

Now

$$
\begin{aligned}
\mathbf{A}^2 &= \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 4 \end{bmatrix} \\
\mathbf{A}^3 &= \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 0 & 8 \end{bmatrix} \\
\mathbf{A}^4 &= \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 0 & 8 \\ 0 & 16 \end{bmatrix}
\end{aligned}
$$

so it is easy to see that

$$
\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} 0 & 2^{n-1} \\ 0 & 2^n \end{bmatrix}.
$$

By putting

$$
\mathbf{X}_0 = \begin{bmatrix} a \\ b \end{bmatrix},
$$

we have

$$
\begin{aligned}
\mathbf{X}_n &= \frac{1}{n!} \begin{bmatrix} 0 & 2^{n-1} \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\
&= \frac{1}{n!} \begin{bmatrix} 2^{n-1} b \\ 2^n b \end{bmatrix} \\
&= \frac{2^n}{n!} \begin{bmatrix} \frac{b}{2} \\ b \end{bmatrix}, \quad n \geq 1.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\mathbf{X}(t) &= \begin{bmatrix} a \\ b \end{bmatrix} + \sum_{n=1}^{\infty} \begin{bmatrix} \frac{b}{2} \\ b \end{bmatrix} \frac{2^n t^n}{n!} \\
&= \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} \frac{b}{2} \\ b \end{bmatrix} \sum_{n=1}^{\infty} \frac{(2t)^n}{n!} \\
&= \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} \frac{b}{2} \\ b \end{bmatrix} \left(e^{2t} - 1\right) \\
&= \begin{bmatrix} a - \frac{b}{2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{b}{2} \\ b \end{bmatrix} e^{2t}.
\end{aligned}
$$

(Recall that $e^{2t} - 1 = \sum_{n=1}^{\infty} \frac{(2t)^n}{n!}$.)

By putting $a - \dfrac{b}{2} = c_1$ and $\dfrac{b}{2} = c_2$, we obtain

$$\mathbf{X}(t) = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{2t}.$$

This solution is the same as that obtained by the method of eigenvalues and eigenvectors.

Example 6.10 Find a series solution for the second order initial value problem

$$\ddot{x} + x = 0, \quad x(0) = 1, \quad \dot{x}(0) = 0.$$

Solution

Let $\dot{x} = y$. Then $\dot{y} = \ddot{x} = -x$ so the second order initial value problem becomes a system of two first order equations

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -x \end{aligned}$$

with $x(0) = 1,\ y(0) = 0$.
By putting $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$ the problem can be rewritten as

$$\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{X}, \quad \mathbf{X}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Since all entries are constants, the matrix $\mathbf{A}$ is analytic at $t = 0$. Thus the sum

$$\sum_{k=0}^{\infty} \mathbf{A}_k (t - t_0)^k$$

can yield the constant matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

if and only if $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots$ are all the zero matrix. Hence, $\mathbf{A}_0 = \mathbf{A}$ so that

$$\begin{aligned} \mathbf{X}_n &= \frac{1}{n} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{X}_{n-1} \\ &= \frac{1}{n!} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^n \mathbf{X}_0 = \frac{1}{n!} \mathbf{A}^n \mathbf{X}_0 \,. \qquad (1) \end{aligned}$$

Now from

$$\mathbf{A}\mathbf{X}_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \mathbf{A}^2\mathbf{X}_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\mathbf{A}^3\mathbf{X}_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{A}^4\mathbf{X}_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

we notice the pattern

$$\mathbf{A}^{2k+1}\mathbf{X}_0 = \begin{bmatrix} 0 \\ (-1)^{k+1} \end{bmatrix}, \quad k = 0, 1, \ldots, \qquad \mathbf{A}^{2k}\mathbf{X}_0 = \begin{bmatrix} (-1)^k \\ 0 \end{bmatrix}, \quad k = 0, 1, \ldots$$

so that

$$\mathbf{X}_{2k+1} = \frac{1}{(2k+1)!}\mathbf{A}^{2k+1}\mathbf{X}_0 = \frac{1}{(2k+1)!}\begin{bmatrix} 0 \\ (-1)^{k+1} \end{bmatrix}, \quad k = 0, 1, \ldots,$$

and

$$\mathbf{X}_{2k} = \frac{1}{(2k)!}\mathbf{A}^{2k}\mathbf{X}_0 = \frac{1}{(2k)!}\begin{bmatrix} (-1)^k \\ 0 \end{bmatrix}, \quad k = 0, 1, \ldots$$

Hence we have

$$\begin{aligned}
\mathbf{X}(t) = \sum_{k=0}^{\infty}\mathbf{X}_k t^k &= \sum_{k=0}^{\infty}\mathbf{X}_{2k}t^{2k} + \sum_{k=0}^{\infty}\mathbf{X}_{2k+1}t^{2k+1} \\
&= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\begin{bmatrix} (-1)^k \\ 0 \end{bmatrix} t^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\begin{bmatrix} 0 \\ (-1)^{k+1} \end{bmatrix} t^{2k+1} \\
&= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\begin{bmatrix} 1 \\ 0 \end{bmatrix} t^{2k} + \sum_{k=0}^{\infty}\frac{(-1)^{k+1}}{(2k+1)!}\begin{bmatrix} 0 \\ 1 \end{bmatrix} t^{2k+1} \\
&= \begin{bmatrix} 1 \\ 0 \end{bmatrix}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}t^{2k} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\sum_{k=0}^{\infty}\frac{(-1)^{k+1}}{(2k+1)!}t^{2k+1}
\end{aligned}$$

Remark

If we used the eigenvalue–eigenvector method to solve this problem, we would have found the exact solution

$$x(t) = \cos t, \quad y(t) = -\sin t.$$

This result can easily be obtained from equation (2) by making use of standard series expansions of $\sin t$ and $\cos t$; we find

$$\mathbf{X}(t) = \begin{bmatrix} \cos t \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \sin t \end{bmatrix} = \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix}.$$

Exercise 6.11

1. Show that

$$\mathbf{X}(t) = \sum_{k=0}^{\infty}\frac{(t-t_0)^k}{k!}\mathbf{A}^k\mathbf{X}_0$$

is the solution of

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X}, \quad \mathbf{X}(t_0) = \mathbf{X}_0.$$

2. Use the *power series method* to find a series solution for

(a) $\dot{\mathbf{X}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}\mathbf{X}, \quad \mathbf{X}(0) = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$

Write your final answer in terms of the appropriate exponential function.

(b) $\ddot{x} + 4x = 0$, $x(0) = 1$, $\dot{x}(0) = 0$.

First write the equation above as a system of two first order differential equations and then find the series solution using the power series method. Write your final answer in terms of the appropriate trigonometric functions.

Remark

We have restricted ourselves to series solutions about "ordinary" points, i.e. points where $\mathbf{A}(t)$ is analytic. Series solutions about certain types of singular points are also possible, but for the purpose of this course, we shall not deal with them.

6.4 THE EXPONENTIAL OF A MATRIX

Our object is this section is to generalize the exponential function e^{ta}, where a is a constant. We wish to attach meaning to the symbol $e^{t\mathbf{A}}$ when $\mathbf{A}$ is a constant matrix, say, an $n \times n$ matrix. That this has significance in the process of a study of the system $\dot{\mathbf{X}} = \mathbf{AX}$, and the initial value problem $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$, can be "suspected" from the fact that $e^{ta}y$ is the (unique) solution of $\dot{x} = ax,\ x(0) = y$.
In view of our previous work, it is possible to define in either of the following two ways:

Definition 6.12 (1) *By keeping in mind that*

$$\mathbf{X}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{A}^k\mathbf{X}_0$$

is the unique solution of $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$, *it makes sense to define*

$$e^{t\mathbf{A}}\mathbf{U} = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{A}^k\mathbf{U} \tag{6.10}$$

with $\mathbf{U}$ *a constant* n*–dimensional vector. With this definition*

$$e^{t\mathbf{A}}\mathbf{X}_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{A}^k\mathbf{X}_0 = \mathbf{X}(t)$$

so that the "function" $e^{t\mathbf{A}}$ *corresponds to* e^{ta}, *the solution of* $\dot{x} = ax$.
(2) *By recalling that the solution* $\mathbf{X}(t)$ *of* $\dot{\mathbf{X}} = \mathbf{AX},\ \mathbf{X}(t_0) = \mathbf{X}_0$, *may also be expressed in terms of the normalized fundamental matrix* $\mathbf{\Phi}(t)$ *of* $\mathbf{A}$ *at* $t = 0$, *viz as* $\mathbf{X}(t) = \mathbf{\Phi}(t)\mathbf{X}_0$, *we can also define*

$$e^{t\mathbf{A}} = \mathbf{\Phi}(t). \tag{6.11}$$

Equation (6.11) is interpreted as

$$e^{t\mathbf{A}}\mathbf{U} = \mathbf{\Phi}(t)\mathbf{U}$$

for every vector with n entries.

By using the latter definition, and the properties of a normalized fundamental matrices, it is easy to show that $e^{t\mathbf{A}}$ has all the properties of the exponential function.

Exercise 6.13 Show that

$$e^{(t+s)\mathbf{A}} = e^{t\mathbf{A}}e^{s\mathbf{A}} \ \textit{for } s,t > 0,$$

$$e^{t\mathbf{A}} = \mathbf{I} \ \textit{for } t = 0.$$

Remark

The generalization of the exponential function has been extended to the case where $\mathbf{A}$, which may be dependent on time, is a bounded or unbounded operator with domain and range in a Banach space. A system $\{\mathbf{T}_t : t > 0\} = \{e^{t\mathbf{A}} : t > 0\}$ of operators is then known as a *semigroup* of bounded operators with *infinitesimal generator* $\mathbf{A}$. The concept of semigroups of bounded operators[2] is of cardinal importance in the study of abstract differential equations and has wide application to the field of the qualitative theory of partial differential equations. A good knowledge of ordinary differential equations, as well as a course in Functional Analysis and Distribution Theory will make this exciting field accessible to you.

[2]The concept of semigroup was named in 1904 and after 1930 the theory of semigroups developed rapidly. One of the most authentic books on this subject is *Functional Analysis and Semigroups*, Amer. Math. Soc. Coll. Publ. Vol. 31, Providence R.I., 1957, by E. Hille and R.S. Phillips. A more recent addition to the literature is *A Concise Guide to Semigroups and Evolution Equations* by Aldo Belleni–Morante, World Scientific, Singapore, 1994.

CHAPTER 7

Nonlinear Systems
Existence and Uniqueness Theorem for Linear Systems

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding nonlinear systems:

- autonomous and non–autonomous systems;
- conditions for existence and uniqueness for solutions of differential equations.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- express a non–autonomous system in autonomous form;
- determine under which conditions the initial value problem $\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X}, \quad \mathbf{X}(t_0) = \mathbf{X}_0$ has a unique solution.

7.1 NONLINEAR EQUATIONS AND SYSTEMS

We showed in Chapter 1 that the equation

$$\dot{\mathbf{X}} = \mathbf{A}(t)\mathbf{X} + \mathbf{G}(t) \tag{7.1}$$

with $\mathbf{A}$ an $n \times n$ matrix $[f_{ij}]$,

$$\mathbf{G}(t) = \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is equivalent to the linear first order system

$$\begin{aligned}
\dot{x}_1 &= f_{11}(t)x_1 + f_{12}x_2 + \ldots + f_{1n}x_n + g_1(t) \\
\dot{x}_2 &= f_{21}(t)x_1 + f_{22}x_2 + \ldots + f_{2n}x_n + g_2(t) \\
&\vdots \\
\dot{x}_n &= f_{n1}(t)x_1 + f_{n2}x_n + \ldots + f_{nn}x_n + g_n(t).
\end{aligned}$$

If for some or other i, x_i cannot be expressed as a linear combination of the components of $\mathbf{X}$ (the coefficients of these components may be either functions of t or else constants) we have a nonlinear equation

$$\dot{x}_i = f_i\left(t,\, x_1, \ldots, x_n\right) \equiv f_i\left(t, \mathbf{X}\right).$$

Systems containing nonlinear equations, i.e. nonlinear systems, are written in vector form as

$$\dot{\mathbf{X}} = \mathbf{F}\left(t, \mathbf{X}\right). \tag{7.2}$$

(The matrix notation is obviously not possible.)

From the above it is clear that the system

$$\dot{\mathbf{X}} = \mathbf{F}\left(t, \mathbf{X}\right)$$

is linear iff $\mathbf{F}\left(t, \mathbf{X}\right) = \mathbf{A}\left(t\right)\mathbf{X} + \mathbf{G}\left(t\right),$ for some other matrix $\mathbf{A}\left(t\right)$ and a vector function $\mathbf{G}\left(t\right).$

A special case of nonlinear systems are the *autonomous* systems.

Definition 7.1 *The system* $\dot{\mathbf{X}} = \mathbf{F}\left(t\right)\mathbf{X}$ *is said to be* **autonomous** *if* $\mathbf{F}$ *is independent of* t.

An autonomous system is, therefore, of the form $\dot{\mathbf{X}} = \mathbf{F}\left(\mathbf{X}\right).$ Autonomous systems have certain properties which are not generally valid.

Example 7.2

(i) The system $\dot{\mathbf{X}} = \mathbf{AX}$ with $\mathbf{A}$ a matrix with constant entries, is autonomous.

(ii) The system

$$\begin{aligned} \dot{x}_1 &= x_1 x_2 \\ \dot{x}_2 &= t x_1 + x_2 \end{aligned}$$

is a non–autonomous system.

It is always possible to express a non–autonomous system in autonomous form by introducing a spurious variable.

Example 7.3

Consider the non–autonomous equation

$$\dot{x} = t^2 x - e^t.$$

Put $x_1 = t$ and $x_2 = x$. Then $\dot{x}_1 = 1$ and $\dot{x}_2 = \dot{x} = t^2 x - e^t,$ i.e. $\dot{x}_2 = x_1^2 x_2 - e^{x_1}.$
This yields the system

$$\dot{\mathbf{X}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ x_1^2 x_2 - e^{x_1} \end{bmatrix},$$

the right–hand side of which is independent of t — therefore an autonomous system.

More generally we consider the non–autonomous system

$$\dot{\mathbf{X}} = \mathbf{F}(t, \mathbf{X}) \quad \text{with } \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

Put

$$\mathbf{Y} = \begin{bmatrix} t \\ x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

Then

$$\dot{\mathbf{Y}} = \begin{bmatrix} 1 \\ \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 1 \\ f_1 \\ \vdots \\ f_n \end{bmatrix} = \mathbf{G}(\mathbf{Y})$$

and the system $\dot{\mathbf{Y}} = \mathbf{G}(\mathbf{Y})$ is autonomous.

Example 7.4
Write the second-order equation

$$\ddot{y} - 2t\dot{y} + y = 0$$

as an autonomous system of first-order equations.

Solution Let $x_1 = t$, $x_2 = y$, $x_3 = \dot{y}$. Then $\dot{x}_1 = 1$, $\dot{x}_2 = \dot{y} = x_3$ and $\dot{x}_3 = \ddot{y} = 2t\dot{y} - y = 2x_1x_3 - x_2$. The autonomous system is therefore

$$\begin{aligned} \dot{x}_1 &= 1 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= 2x_1x_3 - x_2 \end{aligned}$$

Exercise 7.5

(1) Which of the following are linear? autonomous?

(a) $\dot{x} = \sin tx$ (b) $\dot{x} = \frac{1}{x}$ (c) $\ddot{x} - x = t$ (d) $\dot{x}_1 = t,\ \dot{x}_2 = x_1 + x_2$

(e) $\dot{x}_1 = x_1,\ \dot{x}_2 = x_2$ (f) $\dot{x} = t^2x + \frac{1}{t}$ (g) $\dot{x}_1 = \frac{x_1}{t},\ \dot{x}_2 = x_1$

(2) Write each nonautonomous system of problem 1 as an autonomous system.

(3) Write

$$\dddot{y} + a\ddot{y} + b\dot{y} + cy = \alpha t$$

as an autonomous system.

It is seldom possible to solve nonlinear equations explicitly[1]. Certain principles, applicable to linear systems are also not even valid in the case of nonlinear equations, as for instance the principle of uniqueness of

[1] One can, however, obtain important information on certain characteristics of the solution of a nonlinear differential equation without actually solving the equation. One method is the method of phase portraits in the phase space. Nonlinear first–order autonomous systems of differential equations also are studied by investigating the points of equilibrium of the so–called linearized system. This subject is dealt with in Chapter 8.

solutions and the principle of superposition of solutions. The following counterexamples illustrate the truth of these statements.

Exercise 7.6

Show that the initial value problem $\dot{y} = \frac{3}{2} y^{\frac{1}{3}}, \;\; y(0) = 0$ has an infinite number of solutions.

Example 7.7

The principle of superposition of solutions is not valid for the equation $\dot{y} = \frac{t}{2} y^{-1}$.

Differentiation confirms that $y(t) = \left(\frac{t^2}{2} + c \right)^{\frac{1}{2}}$ satisfies, for all c, the equation $2yy' = t$, which is equivalent to $y' = \frac{t}{2y}$. However,

$$\left(\frac{t^2}{2} + c_0 \right)^{\frac{1}{2}} + \left(\frac{t^2}{2} + c_1 \right)^{\frac{1}{2}}$$

is in no case a solution, in view of

$$\begin{aligned} & 2\left\{ \sqrt{\frac{t^2}{2} + c_0} + \sqrt{\frac{t^2}{2} + c_1} \right\} \left\{ \frac{t}{2\sqrt{\frac{t^2}{2} + c_0}} + \frac{t}{2\sqrt{\frac{t^2}{2} + c_1}} \right\} \\ & = t + t + t \left\{ \sqrt{\frac{\frac{t^2}{2} + c_0}{\frac{t^2}{2} + c_1}} + \sqrt{\frac{\frac{t^2}{2} + c_1}{\frac{t^2}{2} + c_0}} \right\} \\ & \neq t \;\; \text{for all } c_0, c_1. \end{aligned}$$

It is, indeed, true that the validity of the principle of superposition would imply linearity — thus leading to a contradiction. We prove this:

Suppose that $\mathbf{F}(t, \mathbf{X})$ is a function such that every initial value problem $\dot{\mathbf{X}} = \mathbf{F}(t, \mathbf{X})$ has a solution. Suppose further that the principle of superposition is valid, i.e. if $\mathbf{X}_i(t)$ is a solution of $\dot{\mathbf{X}} = \mathbf{F}(t, \mathbf{X})$ for $i = 1, 2$, then $\mathbf{Z} = a\mathbf{X}_1 + b\mathbf{X}_2$ is a solution of

$$\dot{\mathbf{Z}} = \mathbf{F}(t, \mathbf{Z}), \; \mathbf{Z}(t_0) = a\mathbf{X}_1^0 + b\mathbf{X}_2^0,$$

where we use the notation $\mathbf{X}_i^0 \equiv \mathbf{X}_i(t_0)$.

We show that $\mathbf{F}(t, \mathbf{Z})$ is a linear function of $\mathbf{Z}$, i.e.

$$\mathbf{F}(t, \, a\mathbf{X}_1 + b\mathbf{X}_2) = a\mathbf{F}(t, \mathbf{X}_1) + b\mathbf{F}(t, \mathbf{X}_2).$$

From our assumption we have

$$\begin{aligned} \mathbf{F}(t, \mathbf{X}_1) &= \dot{\mathbf{X}}_1(t), \\ \mathbf{F}(t, \mathbf{X}_2) &= \dot{\mathbf{X}}_2(t), \end{aligned}$$

and

$$\begin{aligned}\mathbf{F}\left(t,\, a\mathbf{X}_1 + b\mathbf{X}_2\right) &= \frac{d}{dt}\left(a\mathbf{X}_1\left(t\right) + b\mathbf{X}_2\left(t\right)\right)\\ &= a\dot{\mathbf{X}}_1\left(t\right) + b\dot{\mathbf{X}}_2\left(t\right)\\ &= a\mathbf{F}\left(t, \mathbf{X}_1\right) + b\mathbf{F}\left(t, \mathbf{X}_2\right).\end{aligned}$$

As far as the solution of nonlinear equations is concerned, we note that the Method of Separation of Variables can, in the case of simple one–dimensional equations, often be applied successfully.

Example 7.8 Determine the solution of

$$\dot{y} = -\frac{\sqrt{1-t^2}}{\sqrt{5+y}} \qquad -1 \le t \le 1, \;\; y > -5. \tag{7.3}$$

Solution

By separation of variables, equation (7.3) is equivalent to

$$\sqrt{1-t^2}\,dt + \sqrt{5+y}\,dy = 0.$$

By integration we obtain that the solution $y\left(t\right)$ is given implicitly by

$$\frac{1}{2}t\sqrt{1-t^2} + \frac{1}{2}\arcsin t + \frac{2}{3}\left(5+y\right)^{\frac{3}{2}} = c, \quad -1 \le t \le 1. \tag{7.4}$$

Note that in order to obtain a unique solution, we consider only principal values of arcsin, i.e. those values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$, otherwise (7.4) would define a multi–valued function.

Exercise 7.9

(a) Show that $\dot{y} = -t + (2y + t^2)^{\frac{1}{2}}$, $y(1) = -\frac{1}{2}$ is solved by $y(t) = -t^2/2$ and $y_1(t) = -t + \frac{1}{2}$.

(b) Separate variables to find solutions of the following equations

(i) $t\dot{y} = y$ (ii) $\dot{y} = e^{t+y}$ (iii) $\dot{y} = f(t)y = 0$.

7.2 NUMERICAL SOLUTIONS OF DIFFERENTIAL EQUATIONS

If a differential equation cannot be solved by means of a standard technique, a *numerical solution* can still be obtained by means of techniques of approximation. By a numerical solution of a differential equation in x, a function of t, is meant a table of values such that for every value of t, a corresponding value of $x\left(t\right)$ is given. Such a table always has a column giving the magnitude of the error involved. The theory of numerical solutions covers a wide field and is outside the scope of this course. Should the reader be interested, an easily readable chapter on numerical solutions is found in Z&W, Chapter 9, as well as in *Ordinary Differential Equations* by M. Tenenbaum and H. Pollard (see references).

It is interesting to note that a computer can yield a "solution" for a differential equation even when theoretically no solution exists. An example is the boundary value problem

$$\ddot{f}\left(x\right) + f\left(x\right) = \sin x, \quad 0 < x < \pi, \quad f\left(0\right) = f\left(\pi\right) = 0.$$

For this reason it is necessary to pay particular attention to those conditions under which a unique solution of a system of differential equations exists. We shall confine ourselves to the case of linear systems of differential equations.

7.3 EXISTENCE AND UNIQUENESS THEOREM FOR LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS

By putting $\mathbf{A}(t)\mathbf{X} = \mathbf{F}(t, \mathbf{X})$, the problem

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}(t)\mathbf{X}, \\ \mathbf{X}(t_0) &= \mathbf{X}_0, \end{aligned}$$

reduces to the problem

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{F}(t, \mathbf{X}), \\ \mathbf{X}(t_0) &= \mathbf{X}_0. \end{aligned}$$

The theorem will, therefore, be proved by applying the Existence and Uniqueness Theorem for the differential equation

$$\dot{x} = f(t, x), \ x(t_0) = x_0.$$

Theorem 7.10 *If f is continuous in an interval $|t - t_0| \leq T$, $||\mathbf{X} - \mathbf{X}_0|| \leq R$ and f satisfies a Lipschitz condition, i.e. a constant $K > 0$ exists such that*

$$||f(t, \mathbf{U}) - f(t, \mathbf{V})|| \leq K ||\mathbf{U} - \mathbf{V}|| \tag{7.5}$$

in $|t - t_0| \leq T$ and $\mathbf{U}, \mathbf{V}$ in $||\mathbf{X} - \mathbf{X}_0|| \leq R$, then the initial value problem

$$\begin{aligned} \dot{\mathbf{X}} &= f(t, \mathbf{X}) \\ \mathbf{X}(t_0) &= \mathbf{X}_0 \end{aligned}$$

has one and only one solution in $|t - t_0| < \delta$.

Remark

(1) In the above theorem the constant δ is defined as $\delta = \min(T, R/M)$, with the constant M the upperbound of $||\mathbf{F}||$. (See note later on.)

(2) If the range of $\mathbf{F}$ is the Banach Space E^n, i.e. the n–dimensional Euclidian space with the Euclidian metric (recall that a Banach space is a normed space which is complete), then $||\mathbf{F}(t, \mathbf{X})||$ is precisely $|\mathbf{F}(t, \mathbf{X})|$.

(3) If $\mathbf{A}(t)$ is an $n \times n$ matrix with real entries, and $\mathbf{X}$ an n–tuple of real numbers, then $\mathbf{A}(t)\mathbf{X}$ is an n–dimensional vector with components real numbers, so that $\mathbf{A}(t)\mathbf{X}$ is an element of the Banach space R^n. As in Chapter 4, Section 5, we define

$$||\mathbf{A}(t)|| = \left(\sum_{i,j=1}^{n} (f_{ij}(t))^2 \right)^{\frac{1}{2}}. \tag{7.6}$$

Since the role of $\mathbf{F}(t, \mathbf{X})$ is played by $\mathbf{A}(t)\mathbf{X}$, we must have $\mathbf{A}(t)\mathbf{X}$ continuous for all t in $|t - t_0| < T$ and every $\mathbf{X}$ in $||\mathbf{X} - \mathbf{X}_0|| \leq R$, with R an arbitrary constant.

It is sufficient to require that $\mathbf{A}(t)$ should be continuous in $|t - t_0| \leq T$, i.e. that $f_{ij}(t)$ should be continuous for $i = 1, \ldots, n,\ j = 1, \ldots, n$. The vector $\mathbf{X}$ is, indeed, continuous as a function of $\mathbf{X}$ for every "interval" $||\mathbf{X} - \mathbf{X}_0|| \leq R$. Since the product of continuous functions is continuous, it follows that $\mathbf{A}(t)\mathbf{X}$ is continuous in $|t - t_0| \leq T$ and $||\mathbf{X} - \mathbf{X}_0|| \leq R$.

For the proof of Theorem 7.10 we need

Lemma 7.11 *If*

$$\mathbf{A}(t) = [f_{ij}(t)], \quad \mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

then we have the inequality

$$||\mathbf{A}(t)\mathbf{X}|| \leq ||\mathbf{A}(t)|| \; ||\mathbf{X}||. \tag{7.7}$$

Proof. By putting

$$\mathbf{A}(t) = \begin{bmatrix} f_{11} & \cdots & f_{1n} \\ f_{21} & \cdots & f_{2n} \\ \vdots & & \\ f_{n1} & \cdots & f_{nn} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_n \end{bmatrix},$$

i.e. $\mathbf{f}_i = (f_{i1}, f_{i2}, \ldots, f_{in}), \quad i = 1, \ldots, n$, we have

$$\mathbf{AX} = \begin{bmatrix} \mathbf{f}_1 \cdot \mathbf{X} \\ \mathbf{f}_2 \cdot \mathbf{X} \\ \vdots \\ \mathbf{f}_n \cdot \mathbf{X} \end{bmatrix}$$

where $\cdot$ indicates the ordinary scalar product. Then, by definition

$$\begin{aligned} ||\mathbf{A}(t)\mathbf{X}||^2 &= (\mathbf{f}_1 \cdot \mathbf{X})^2 + \ldots + (\mathbf{f}_n \cdot \mathbf{X})^2 \\ &\leq ||\mathbf{f}_1||^2 \; ||\mathbf{X}||^2 + \ldots + ||\mathbf{f}_n|| \; ||\mathbf{X}||^2 \quad \text{(Schwarz' Inequality)} \\ &= \left(||\mathbf{f}_1||^2 + ||\mathbf{f}_2||^2 + \ldots + ||\mathbf{f}_n||^2\right) ||\mathbf{X}||^2. \end{aligned}$$

Since

$$||\mathbf{f}_i||^2 = f_{i1}^2 + f_{i2}^2 + \ldots + f_{in}^2 = \sum_{j=1}^{n} (f_{ij})^2,$$

we have

$$\begin{aligned} ||\mathbf{f}_1||^2 + \ldots + ||\mathbf{f}_n||^2 &= \sum_{j=1}^{n} (f_{1j}(t))^2 + \sum_{j=1}^{n} (f_{2j}(t))^2 + \ldots + \sum_{j=1}^{n} (f_{nj}(t))^2 \\ &= \sum_{t,j=1}^{n} (f_{ij}(t))^2 \\ &= ||\mathbf{A}(t)||^2 . \end{aligned}$$

It follows that

$$||\mathbf{A}(t)\mathbf{X}|| \leq (||\mathbf{A}(t)|| \; ||\mathbf{X}||) .$$

■

In proving Theorem 7.10, we finally make use of the fact that if a function f is continuous in $|t - t_0| \leq T$, $||\mathbf{X} - \mathbf{X}_0|| \leq R$, then a constant M, independent of t and $\mathbf{X}$, exists, such that $||f|| \leq M$. Indeed, continuity and boundedness are equivalent concepts in normed spaces. For the proof, the reader may refer to, among others, "Mathematical Analysis" by Apostol, p. 83[2].

Our main result now reads:

Theorem 7.12 *If $\mathbf{A}(t)$ is continuous in $|t - t_0| \leq T$, the initial value problem*

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}(t)\mathbf{X} \\ \mathbf{X}(t_0) &= \mathbf{X}_0 \end{aligned}$$

has a unique solution in $|t - t_0| < \delta$.

Proof. From the assumption that $\mathbf{A}(t)$ is continuous in $|t - t_0| \leq T$ we have, according to our previous remarks, that $\mathbf{A}(t)\mathbf{X}$ is continuous in $|t - t_0| \leq T$ and $||\mathbf{X} - \mathbf{X}_0|| \leq R$ for every R. We also have, from a previous remark, that $||\mathbf{A}(t)| \leq M|$ in $|t - t_0| \leq T$.
Therefore

$$\begin{aligned} ||\mathbf{A}(t)\mathbf{U} - \mathbf{A}(t)\mathbf{V}|| &= ||\mathbf{A}(t)(\mathbf{U} - \mathbf{V})|| \\ &\leq ||\mathbf{A}(t)|| \; ||\mathbf{U} - \mathbf{V}|| \\ &\leq M \, ||\mathbf{U} - \mathbf{V}|| . \end{aligned}$$

Consequently $\mathbf{A}(t)\mathbf{X} = \mathbf{F}(t, \mathbf{X})$ satisfies the Lipschitz condition (7.5) with $K = M$. All the conditions of Theorem 7.10 are now satisfied. We can conclude that the problem

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}(t)\mathbf{X} \\ \mathbf{X}(t_0) &= \mathbf{X}_0 \end{aligned}$$

has a unique solution in $|t - t_0| < \delta$. This completes the proof of the theorem. ■

[2]Mathematical Analysis, T.M. Apostol, Addison–Wesley Publishing Co., London.

CHAPTER 8

Qualitative Theory of Differential Equations
Stability of Solutions of Linear Systems
Linearization of Nonlinear Systems

Objectives for this Chapter

The main objective of this chapter is to gain an understanding of the following concepts regarding the stability of solutions of linear systems and the linearization of nonlinear systems:

- autonomous systems;
- critical point;
- periodic solutions;
- classification of a critical point: stable/unstable node, saddle, center, stable/unstable spiral point, degenerate node;
- stability of critical point;
- linearization and local stability;
- Jacobian matrix;
- phase–plane method.

Outcomes of this Chapter

After studying this chapter the learner should be able to:

- find the critical points of plane autonomous systems;
- solve certain nonlinear systems by changing to polar coordinates;
- apply the stability criteria to determine whether a critical point is locally stable or unstable;
- classify the critical points using the Jacobian matrix.

The final chapter of this module is devoted to the qualitative theory of differential equations. In this branch of theory of differential equations, techniques are developed which will enable us to obtain important information about the solutions of differential equations without actually solving them. We will, for instance, be able to decide whether a solution is stable, or what the long time behaviour of the solution is, without even knowing the form of the solution. This is extremely useful in view of the fact that it is often very difficult, or even impossible, to determine an exact solution of a differential equation.

Study Chapter 10 of Z&W, work through all the examples and do the exercises at the end of each section.

Appendix A

Symmetric Matrices

(Taken from Goldberg & Schwartz - see Preface for full reference.)

You should by now be familiar with the following basic properties of matrices:

(1) $\mathbf{A}$ is symmetric if $\mathbf{A}^T = \mathbf{A}$, where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$;

(2) $(\mathbf{A}^T)^T = \mathbf{A}$;

(3) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

Suppose $\mathbf{A}$ is an $n \times n$ matrix with only real-valued entries. For any given vector $\mathbf{u} = [u_1, u_2, \ldots, u_n]^T$, the vector $\overline{\mathbf{u}}$ is that vector whose entries are the complex conjugates of the entries of $\mathbf{u}$, that is $\overline{\mathbf{u}} = [\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_n]^T$. Obviously $\overline{\mathbf{u}} = \mathbf{u}$, if $\mathbf{u}$ has only real entries.

Lemma A.1 *For any vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathcal{C}^n$ *and any real symmetric matrix* $\mathbf{A}$

$$\mathbf{u}^T\mathbf{A}\overline{\mathbf{v}} = \overline{\mathbf{v}}^T\mathbf{A}\mathbf{u}. \tag{A.1}$$

Proof The product $\mathbf{A}\overline{\mathbf{v}}$ is a column matrix. Since $\mathbf{u}^T$ is a row matrix $\mathbf{u}^T\mathbf{A}\overline{\mathbf{v}}$ is a matrix with one entry. Therefore $(\mathbf{u}^T\mathbf{A}\overline{\mathbf{v}})^T = \mathbf{u}^T\mathbf{A}\overline{\mathbf{v}}$. But $(\mathbf{u}^T\mathbf{A}\overline{\mathbf{v}})^T = \overline{\mathbf{v}}^T\mathbf{A}^T(\mathbf{u}^T)^T = \overline{\mathbf{v}}^T\mathbf{A}\mathbf{u}$, since $\mathbf{A}^T = \mathbf{A}$. Hence the result follows. ■

Lemma A.2 *If* λ *is an eigenvalue of* $\mathbf{A}$ *with corresponding eigenvector* $\mathbf{u}$ *then* $\overline{\mathbf{u}}$ *is an eigenvector of* $\mathbf{A}$ *corresponding to the eigenvalue* $\overline{\lambda}$.

Proof By hypothesis, $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$. Therefore the complex conjugates of both sides are equal, that is $\overline{\mathbf{A}\mathbf{u}} = \mathbf{A}\overline{\mathbf{u}} = \overline{\lambda}\overline{\mathbf{u}}$ (since $\mathbf{A}$ has real entries). Hence $\overline{\mathbf{u}}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue $\overline{\lambda}$. ■

Lemma A.3 *A symmetric matrix has only real eigenvalues.*

Proof Let $\mathbf{v} = \mathbf{u}$ in (A.1). Suppose $\mathbf{A}$ is a symmetric matrix and let $\mathbf{u}$ be an eigenvector of $\mathbf{A}$ which corresponds with the eigenvalue λ. From Lemma A.2 we have that

$$\mathbf{u}^T\mathbf{A}\overline{\mathbf{u}} = \mathbf{u}^T(\overline{\lambda}\overline{\mathbf{u}}) = \overline{\lambda}\mathbf{u}^T\overline{\mathbf{u}}$$

so that from Lemma A.1 it follows that

$$\overline{\mathbf{u}}^T\mathbf{A}\mathbf{u} = \overline{\mathbf{u}}^T(\lambda\mathbf{u}) = \lambda\overline{\mathbf{u}}^T\mathbf{u}.$$

Therefore

$$\lambda\overline{\mathbf{u}}^T\mathbf{u} = \overline{\lambda}\mathbf{u}^T\overline{\mathbf{u}}.$$

Since $\overline{\mathbf{u}}^T\mathbf{u}$ and $\mathbf{u}^T\overline{\mathbf{u}}$ are 1×1 matrices with only one entry $|u_1|^2 + |u_2|^2 + \ldots + |u_n|^2$, and since $\mathbf{u} \neq \mathbf{0}$, this entry is positive. Hence $\lambda\overline{\mathbf{u}}^T\mathbf{u} = \overline{\lambda}\mathbf{u}^T\overline{\mathbf{u}}$ implies $\overline{\lambda} = \lambda$ which means that λ is real. ■

Theorem A.4 *If* $\mathbf{A}$ *is a real* $n \times n$ *symmetric matrix then there exists* n *linearly independent eigenvectors each belonging to* $\mathcal{R}^n$ *(and hence spanning* $\mathcal{R}^n$*).*

Theorem A.5 *If* $\mathbf{A}$ *is symmetric then the initial value problem*

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(t_0) = \mathbf{x_0}$$

has a solution for every $\mathbf{x_0} \in \mathcal{R}^n$.

Appendix B

Refresher notes on Methods of Solution of One-dimensional Differential Equations

I. First order Differential Equations

A few important first order differential equations are Linear equations, Homogeneous equations, Exact equations and the Bernoulli equation.

A. Linear Equations

General form:

$$\frac{dy}{dx} + p(x)y = q(x)$$

where $p(x)$ and $q(x)$ are continuous functions in some or other interval.

Method of Solution:

Multiply be integrating factor (I.F.) $e^{\int p(x)dx}$. This yields

$$e^{\int p(x)dx}(\frac{dy}{dx} + p(x)y) = e^{\int p(x)dx}.q(x),$$

i.e.

$$\frac{d}{dx}(ye^{\int p(x)dx}) = e^{\int p(x)dx}.q(x).$$

By integration we obtain

$$ye^{\int p(x)dx} = \int \{e^{\int p(x)dx}.q(x)\}\,dx + k.$$

Example

Solve

$$\frac{dx}{dt} - \frac{x}{t} = 1. \tag{1}$$

$$I.F. = e^{-\int \frac{dt}{t}} = e^{-\ln t} = \frac{1}{t}$$

$$(1) \times \frac{1}{t}: \quad \frac{1}{t}\frac{dx}{dt} - \frac{x}{t^2} = \frac{1}{t},$$

i.e.

$$\frac{d}{dt}(\frac{x}{t}) = \frac{1}{t}.$$

Therefore

$$\frac{x}{t} = \ln(tc),\ c \quad \text{an arbitrary constant},$$

i.e.

$$x = t\ln(tc).$$

B. The Bernoulli-equation

Definition

The Bernoulli equation is the differential equation

$$\frac{dy}{dx} + p(x)y = q(x)y^n, \quad n \neq 0, 1 \ ,$$

where $p(x)$ and $q(x)$ are continuous functions in some or other interval.

Method of Solution:

By putting $z = y^{1-n}$ the equation is reduced to a linear D.E. in z.

Example

Find the solution of

$$x\frac{dy}{dx} - \frac{y}{2\ln x} = y^2.$$

In this case

$$p(x) = -\frac{1}{2x\ln x}, \qquad q(x) = \frac{1}{x}, \qquad n = 2.$$

By putting

$$z = y^{-1},$$

we obtain the linear D.E.

$$\frac{dz}{dx} + \frac{1}{2x\ln x}\ z = -\frac{1}{x}.$$

Multiplication by

$$e^{\int \frac{dx}{2x\ln x}} = e^{\frac{1}{2}\ln(\ln x)} = (\ln x)^{\frac{1}{2}},$$

yields

$$\frac{d}{dx}(z(\ln x)^{\frac{1}{2}}) = -\frac{(\ln x)^{\frac{1}{2}}}{x}.$$

Therefore

$$z(\ln x)^{\frac{1}{2}} = -\frac{2}{3}(\ln x)^{\frac{3}{2}} + c$$

or

$$z = -\frac{2}{3}\ln x + c(\ln x)^{-\frac{1}{2}},$$

i.e.

$$y = \frac{1}{-\frac{2}{3}\ln x + c(\ln x)^{-\frac{1}{2}}}.$$

C. Homogeneous equations

Definition

A homogeneous differential equation is an equation of the form

$$\frac{dy}{dx} = \frac{g\,(x,y)}{h\,(x,y)} \tag{1}$$

where the functions g and h are homogeneous of the same degree. We recall that a function $g(x,y)$ **homogeneous of degree** n if $g(\alpha x, \alpha y) = \alpha^n g(x,y)$.

Method of Solution:

By putting $y = vx(1)$ is reduced to a D.E. in which the variables are separable.

Example

Find the solution of

$$\frac{dy}{dx} = \frac{x^3+y^3}{xy^2}. \tag{2}$$

Here $g(x,y) = x^3+y^3$ and $h(x,y) = xy^2$ are homogeneous of degree 3. Put $y = vx$. Then (2) becomes

$$v + x\frac{dv}{dx} = \frac{1}{v^2} + v$$

whence

$$v^2 dv = \frac{1}{x}dx.$$

Integrations yields

$$\frac{v^3}{3} = \ln|x| + c,$$

i.e.

$$\frac{1}{3}\frac{y^3}{x^3} = \ln|x| + c.$$

Therefore

$$y = x(d\ln|x| + c)^{\frac{1}{3}}.$$

D. Exact Differential Equations

Definition

A D.E. of the form

$$M(x,y)dx + N(x,y)dy = 0$$

is called **exact** if there exists a function $f(x,y)$ of which the total differential

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

satisfies

$$df = M(x,y)dx + N(x,y)dy.$$

Theorem 1 below provides a method to check whether a D.E. is exact, while Theorem 2 contains a formula for the solution.

Theorem 1

The D.E. $M(x,y)dx + N(x,y)dy = 0$ is exact if and only if $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$.

Theorem 2

The solution of (1) is implicitly given by

$$f(x,y) = \int_{x_0}^{x} M(t,y_0)dt + \int_{y_0}^{y} N(x,s) = C,$$

where (x_0, y_0) is any point in the region in which the functions $M, N, \dfrac{\partial M}{\partial y}$ and $\dfrac{\partial N}{\partial x}$ are continuous. If these functions are continuous in $(0,0)$, it is convenient to choose

$$(x_0, y_0) = (0,0).$$

Example

Solve

$$(3x^2 + 4xy^2)dx + (2y - 3y^2 + 4x^2y)dy = 0.$$

Here

$$M(x,y) = 3x^3 + 4xy^2 \text{ and } N(x,y) = 2y - 3y^2 + 4x^2y.$$

Therefore

$$\frac{\partial M}{\partial y} = 8xy = \frac{\partial N}{\partial x},$$

so that the D.E. is exact. The solution is implicitly given by

$$\int_0^x 3t^2 dt + \int_0^y (2s - 3s^2 + 4x^2s)ds = c,$$

i.e.

$$x^3 + y^2 - y^3 + 2x^2y^2 = c.$$

II. Higher order differential equations with constant coefficients General Form:

$$a_n \frac{d^ny}{dx^n} + a_{n-1}\frac{d^{n-1}}{dx^{n-1}} + \ldots + a_1 \frac{dy}{dx} + a_0 y = F(x) \tag{1}$$

where a_k are constants for $k = 0, \ldots, n$.

By putting

$$\frac{d}{dx} = D, \ldots, \frac{d^n}{dx^n} = D^n,$$

(1) is equivalent to

$$(a_n D^n + a_{n-1}D^{n-1} + \ldots + a_0)y = F(x),$$

i.e.

$$f(D)y = F(x), \tag{2}$$

with $f(D)$ a polynomial in D of degree n. If $F(x) = 0 \; \forall \; x$, we have that the homogeneous equation $f(D)y = 0$. Otherwise (2) is non–homogeneous.

A. Properties of the Differential Operator D, applied to some Standard Functions.

(1) $f(D)[e^{ax}] = f(a)e^{ax}$.

(2) $f(D)[e^{ax}V(x)] = e^{ax}f(D+a)[V(x)]$ (Shift Property).

(3) $$f(D^2)\begin{bmatrix}\sin ax\\ \cos ax\end{bmatrix} = f(-a^2)\begin{bmatrix}\sin ax\\ \cos ax\end{bmatrix}.$$

Proof

We prove (3) only for $f(D^2) = \sum\limits_{k=0}^{2} c_k D^{2k}$ where $c_k,\ k = 0, 1, 2$, are constants.

$$\begin{aligned}\left\{\sum_0^2 c_k D^{2k}\right\}[\sin ax] &= \sum c_k D^{2k}[\sin ax]\\ &= \sum c_k(-a^2)^k \sin ax \text{ (since } D^2(\sin ax) = D(a\cos ax) = -a^2\sin ax)\\ &= \left\{\sum c_k(-a^2)^k\right\}[\sin ax]\\ &= f(-a^2)\sin ax \text{ (by the definition of } f\left(D^2\right)).\end{aligned}$$

B. Properties of the Inverse Operator $\dfrac{1}{f(D)}$ applied to Standard Functions

(1) $\dfrac{e^{ax}}{f(D)} = \dfrac{e^{ax}}{f(a)}$ provided $f(a) \neq 0$

(2) $\dfrac{e^{ax}V(x)}{f(D)} = e^{ax}\dfrac{V(x)}{f(D+a)}$.

This property is known as the Shift Property and is particularly useful whenever $f(a) = 0$.

(3)
$$\frac{\cos ax}{f(D^2, D)} = \frac{\cos ax}{f(-a^2, D)} \quad\text{and}\quad \frac{\sin ax}{f(D^2, D)} = \frac{\sin ax}{f(-a^2, D)}$$

provided $f(-a^2, D) \neq 0$.

Example

$$\frac{1}{D^2 + a^2}[\sin bx] = \frac{\sin bx}{(a^2 - b^2)} \text{ provided } b \neq a.$$

(4) $$\frac{1}{(D^2 + a^2)}[\sin ax] = \frac{-x}{2a}\cos ax$$

$$\frac{1}{(D^2 + a^2)}[\cos ax] = \frac{+x}{2a}\sin ax.$$

(5) $$\frac{\sin ax}{\lambda D + \mu} = \frac{-1}{(\lambda^2 a^2 + \mu^2)}(\lambda a\cos ax - \mu\sin ax).$$

$$\frac{\cos ax}{\lambda D + \mu} = \frac{1}{(\lambda^2 a^2 + \mu^2)}(\lambda a\sin ax + \mu\cos ax),\ \lambda, \mu \text{ real}.$$

Proof

We prove only (1) and (5).

(1) We have to show that

$$f(D)\left\{\frac{e^{ax}}{f(a)}\right\} = e^{ax}.$$

From $f(D)[e^{ax}] = f(a)e^{ax}$, it follows that

$$f(D)\left\{\frac{e^{ax}}{f(a)}\right\} = \frac{1}{f(a)}f(D)e^{ax} = e^{ax}.$$

(5) By applying $\dfrac{\lambda D-\mu}{\lambda D-\mu}$, it follows that

$$\begin{aligned}
\frac{\sin ax}{\lambda D+\mu} &= \frac{(\lambda D-\mu)}{(\lambda^2D^2-\mu^2)}[\sin ax]. \\
&= \frac{(\lambda D-\mu)}{(-\lambda^2a^2-\mu^2)}[\sin ax] \quad \text{(by (3))} \\
&= \frac{-(\lambda a\cos ax-\mu\sin ax)}{(\lambda^2a^2+\mu^2)} \quad \text{(by differentiation).}
\end{aligned}$$

Solution of $f(D)y = F(x)$

The solution is given by $y(x) = y_{C.F.} + y_{P.I.}$, where $y_{C.F.}$ denotes the complementary function, which solves $f(D)y = 0$, and $y_{P.I.}$ a particular solution, the so–called Particular Integral.

On finding the Complementary Function

The auxiliary equation $f(m) = 0$ is formed from $f(D)$ — solve this for m. Suppose $f(m) = 0$ has

(i) n different real roots $m_1, \ldots, m_n$. Then

$$y_{C.F.}(x) = A_1e^{m_1x} + A_2e^{m_2x} + \ldots + A_ne^{m_nx}$$

with $A_i (i = 1, \ldots, n)$ arbitrary constants.

(ii) n coincident real roots m. Then

$$y_{C.F.}(x) = (A_1 + A_2x + \ldots + A_nx^{n-1})e^{mx}.$$

(iii) We first assume that $f(D)$ is of degree 2 (i.e. $f(m) = 0$ has only 2 roots). Suppose the roots are complex conjugates $\xi \pm \eta i$, $(\xi, \eta$ real). Then

$$y_{C.F.}(x) = e^{\xi x}(A_1\cos\eta x + A_2\sin\eta x).$$

Suppose for $f(D)$ of degree $2n$ the roots of $f(m) = 0$ occur in complex conjugate pairs $\alpha_r \pm i\beta_r$, $r = 1, \ldots, n$. Then

$$y_{C.F.}(x) = \sum_{r=1}^{n} e^{\alpha_rx}(A_r\cos\beta_rx + B_r\sin\beta_rx),$$

with A_r, B_r for $r = 1, \ldots, n$ arbitrary constants. If $(m) = 0$ has k–fold roots, i.e.

$$(m - (\alpha + i\beta))^k \; (m - (\alpha - i\beta))^k = 0,$$

then

$$\begin{aligned} y_{C.F.}(x) \; &= e^{\alpha x}(C_1 + C_2 x + \ldots + C_k x^{k-1})[A_1 \cos \beta x + \\ &\quad B_1 \sin \beta x + x(A_2 \cos \beta x + B_2 \sin \beta x) + \\ &\quad + \ldots + x^{k-1}(A_k \cos \beta x + B_k \sin \beta x)] \end{aligned}$$

Note that $y_{C.F.}(x)$ is the solution of $f(D)y = 0$.

Example 1

(i) Solve

$$(D^2 - 7D + 12)y = 0$$

$f(m) = m^2 - 7m + 12$. Put $f(m) = 0$. Then $(m^2 - 7m + 12) = 0$, i.e.

$$(m - 3)(m - 4) = 0.$$

Therefore

$$y = y_{C.F.} = Ae^{3x} + Be^{4x}.$$

(ii) Solve

$$(8D^3 - 12D^2 + 16D - 1)y = 0.$$

Put

$$f(m) = 0.$$

Now

$$\begin{aligned} & 8m^3 - 12m^2 + 6m - 1 = 0 \\ \Longleftrightarrow \quad & 8m^3 - 1 - 6m(2m - 1) = 0 \\ \Longleftrightarrow \quad & (2m - 1)(4m^2 + 2m + 1) - 6m(2m - 1) = 0 \\ \Longleftrightarrow \quad & (2m - 1)(4m^2 - 4m + 1) = 0 \\ \Longleftrightarrow \quad & (2m - 1)^3 = 0 \end{aligned}$$

Therefore

$$y = y_{C.F.} = (A + Bx + Cx^2)e^{\frac{x}{2}}.$$

(iii) Solve

$$\begin{aligned} (D^2 - 2D + 5)y \; &= 0 \\ m^2 - 2m + 5 \; &= 0 \Longleftrightarrow (m - 1)^2 - 4i^2 = 0 \\ \Longleftrightarrow m - 1 \; &= \pm 2i, \; \text{ i.e. } \; m = 1 \pm 2i. \end{aligned}$$

Therefore

$$y \; = \; y_{C.F} = e^x(A \cos 2x + B \sin 2x),$$

with A, B arbitrary constants.

(iv) Solve

$$(D-4)^3(D^2+9)^2y = 0.$$

The roots of the auxiliary equation are

$$m = 4, 4, 4, \qquad m = \pm 3i, \pm 3i.$$

Therefore

$$\begin{aligned} y = y_{C.F.} &= (A_1 + A_2x + A_3x^2)e^{4x} + \\ &[(C_4 \cos 3x + C_5 \sin 3x) + \\ &x(C_6 \cos 3x + C_7 \sin 3x)], \\ A_i(i = 1, 2, 3), &\quad C_j(j = 4, 5, 6, 7) \text{ arbitrary constants.} \end{aligned}$$

On finding the Particular Integral

For the D.E. $f(D)y = F(x)$, $y_{P.I}(x)$ is given by $y_{P.I}(x) = \dfrac{F(x)}{f(D)}$. Now apply B.

Example 2

(i) Determine $Y_{P.I.}$ for the D.E. $(D^2 + 6D + 9)y = 50e^{-3x}$. In this case we have

$$f(D) = D^2 + 6D + 9, \text{ so that } f(a) = 0.$$

It now follows from $B(2)$, by writing

$$\frac{50e^{-3x}}{(D+3)^2} \quad \text{as} \quad \frac{e^{0x}50e^{-3x}}{(D+3)^2}$$

(i.e. $V(x) = e^{0x} = 1$), that

$$\begin{aligned} y_{P.I.} &= 50e^{-3x} . \frac{1}{((D-3)+3)^2} \\ &= 50e^{-3x}\frac{x^2}{2}. \end{aligned}$$

Note that $\dfrac{1}{D^k}$ is interpreted as integration k times.

(ii) Find $y_{P.I.}$ for

$$(D^2 + 9)y = 40 \sin 4x.$$

By $B(3)$, we have

$$y_{P.I.} = \frac{40 \sin 4x}{-16+9} = -\frac{40 \sin 4x}{7}.$$

(iii) Find $y_{P.I.}$ for

$$(D^2 + 9)y = \sin 3x$$

By $B(4)$ we have

$$y_{P.I.} = -\frac{x}{6} \cos 3x.$$

Find $y_{P.I.}$ for

$$(D^3 - 2D^2 + 2D + 7)y = \sin 2x.$$

By $B(3)$ we have

$$\begin{aligned} y_{P.I.} &= \frac{\sin 2x}{D(-4) - 2(-4) + 2D + 7} = \frac{\sin 2x}{-2D + 15} \\ &= \frac{(-2D - 15)[\sin 2x]}{4D^2 - 225} \\ &= -(2D + 15)\frac{[\sin 2x]}{-16 - 225} \qquad (B(3)) \\ &= \frac{-4\cos 2x - 15\sin 2x}{-241} \ \text{(by differentiation)} \end{aligned}$$

Therefore

$$y_{P.I.} = \frac{4\cos 2x + 15\sin 2x}{241}$$

If $f(x)$ is not one of the standard functions in B, (or even if it is), then $y_{P.I.}$ may be obtained with the aid of

The Method of Undetermined Coefficients

Example 1

Find

$$y_{P.I.} \text{ for } (D^2 + 1)y = 2 + x^{2.} \tag{3}$$

Suppose

$$y_{P.I.} = a + bx + cx^2$$

Then

$$Dy = y' = b + 2cx, Dy^2 = 2c.$$

Substitute in (3). We then have $2c + a + bx + cx^2 = 2 + x^2$. (4)

Compare coefficients of powers of x in (4).

$$\begin{aligned} x^2 &: c = 1 \\ x &: b = 0 \\ x^0 &: 2c + a = 2 \Rightarrow a = 0. \end{aligned}$$

Therefore

$$y_{P.I.}(x) = x^{2.}$$

Example 2

Solve

$$(3D^2 + 3D)x = e^t \cos t \tag{5}$$

$$3m^2 + 3m = 0 \Rightarrow m = 0, m = -1.$$

Therefore

$$x_{C.F.}(t) = Ae^{-t} + B.$$

Suppose

$$x_{P.I.} = Ae^t \cos t + Be^t \sin t.$$

Then

$$\begin{aligned} Dx_{P.I.} &= Ae^t(\cos t - \sin t) + Be^t(\sin t + \cos t), \\ D^2x_{P.I.} &= -2Ae^t \sin t + 2Be^t \cos t. \end{aligned}$$

Substitute in (5):

$$(3D^2 + 3D)x = (-9A + 3B)e^t \sin t + (3A + 9B)e^t \cos t.$$

Compare coefficients of $\cos t$ and $\sin t$:

$$\begin{aligned} \cos t\text{:} \quad 3A + 9B &= 1 \\ \sin t\text{:} \ -9A + 3B &= 0 \end{aligned}$$

This yields

$$A = \frac{1}{30}, \qquad B = \frac{1}{10}.$$

Therefore

$$x_{P.I}(t) = Ae^{-t} + B + \frac{1}{30}e^t \cos t + \frac{1}{10}e^t \sin t.$$