Introduction
to ordinary differential equations

Roberto Monti

Padova, 28th April 2010

Contents

Introduction

Let $\Omega \subset \mathbb{R}^{n+2}$, $n \in \mathbb{N}$, be an open set and let $F : \Omega \to \mathbb{R}$ be a continuous function. An equation of the form

$$F(x, y, y', ..., y^{(n)}) = 0 \tag{0.0.1}$$

is called ordinary differential equation of order n. Here, x is a real variable, y is a real valued unknown function, and y',...,$y^{(n)}$ are its derivatives.

A function $\varphi \in C^n(I)$ is a solution of the differential equation if:

i) $I \subset \mathbb{R}$ is an open interval;
ii) $(x, y(x), ..., y^{(n)}(x)) \in \Omega$ for all $x \in I$;
iii) $F(x, y(x), ..., y^{(n)}(x)) = 0$ for all $x \in I$.

The main problems concerning ordinary differential equations are:

1) Existence of solutions;
2) Uniqueness of solutions (with suitable initial conditions or boundary value data);
3) Regularity and stability of solutions (e.g. dependence on the initial conditions, large time stability, higher regularity depending on F or on parameters);
4) Computation of solutions.

The existence of solutions can be proved by fixed point theorems, by approximation and compactness, by variational methods (minimization and critical point theorems), by the implicit function theorem in Banach spaces, by Functional Analysis techniques. The problem of uniqueness is typically more difficult. Only in very special cases, it is possible to compute the solutions in some explicit form.

CHAPTER 1

Some methods of resolution

1. First order linear equations

1.1. First order linear equations. Let $I \subset \mathbb{R}$ be an open interval and let $a, b \in C(I)$ be two continuous functions. A first order differential equation of the form

$$y' + a(x)y = b(x), \quad x \in I, \tag{1.1.2}$$

is called linear. In the case $b = 0$, the equation is said to be homogeneous

$$y' + a(x)y = 0, \quad x \in I. \tag{1.1.3}$$

We solve the homogeneous equation, first. Assuming $y \neq 0$, e.g. $y > 0$, the differential equation (1.1.3) has the form $y'/y = -a(x)$. A primitive of y'/y is $\log y$. Then, denoting by A a primitive of a, i.e. $A'(x) = a(x)$ for all $x \in I$, we have

$$-A = \log y + d,$$

for some constant $d \in \mathbb{R}$. It follows that $y = \exp(-d - A)$ and letting $c = e^{-d}$ we find the solution

$$y(x) = ce^{-A(x)}, \quad x \in I. \tag{1.1.4}$$

This function is a solution to the homogeneous equation for any $c \in \mathbb{R}$ (i.e. the restriction $y > 0$ can be dropped).

Now we look for a solution of the form (1.1.4) for the non homogeneous equation (1.1.2), where now $c \in C^1(I)$ is a function that has to be determined (this method is called "Variation of constants"). Plugging $y' = c'e^{-A} - ace^{-A}$ into (1.1.2) we get

$$c'e^{-A} = b, \quad \text{that is} \quad c' = be^{A}.$$

Integrating this equation on some interval $(x_0, x) \subset I$ we get

$$c(x) = c(x_0) + \int_{x_0}^{x} b(t)e^{A(t)}dt,$$

and we find

$$y(x) = \Big(c(x_0) + \int_{x_0}^{x} b(t)e^{A(t)}dt\Big)e^{-A(x)}, \quad x \in I, \tag{1.1.5}$$

where $c(x_0) \in \mathbb{R}$ is a real number and $x_0 \in I$.

Proposition 1.1.1. *Let $x_0 \in I$ and A be a primitive of a. Then the function in (1.1.5) is a solution to (1.1.2). Moreover, any solution of (1.1.2) is of the form (1.1.5) for some $c(x_0) \in \mathbb{R}$.*

PROOF. The first statement is a computation. Let $z \in C^1(I)$ be a solution to (1.1.2) and let

$$w(x) = e^{A(x)}z(x) - \int_{x_0}^{x} b(t)e^{A(t)}dt.$$

For we have

$$w' = (az + z')e^A - be^A = 0,$$

the function w is constant on I, and the second claim is proved.

□

2. Separation of variables

Let $I, J \subset \mathbb{R}$ be two open intervals and let $f \in C(I)$ and $g \in C(J)$ be two continuous functions. We look for solutions to the first order differential equation

$$y' = f(x)g(y). \tag{1.2.6}$$

Let $x_0 \in I$ and $y_0 \in J$. If $g(y_0) = 0$ for some $y_0 \in J$, then the constant function $y(x) = y_0$, $x \in I$, is a solution to the differential equation (1.2.6). Assume that $g(y_0) \neq 0$. Then it is $g \neq 0$ in a neighborhood of y_0 and we can divide the equation be $g(y)$ (separation of variables). We find

$$\frac{y'(x)}{g(y(x))} = f(x). \tag{1.2.7}$$

Let $G \in C^1(J_1)$ be the primitive of $1/g(y)$ (in the variable y), defined in some interval J_1 containing y_0. The function G is strictly monotonic, because $G'(y) \neq 0$, and thus invertible. Moreover, let $F \in C^1(I)$ be a primitive of f. Upon integrating (1.2.7), we get

$$G(y(x)) = F(x) + C, \quad x \in I_1, \tag{1.2.8}$$

for some interval $I_1 \subset I$. Here $C \in \mathbb{R}$ is a real constant. The general solution of the differential equation is then

$$y(x) = G^{-1}(F(x) + C), \quad x \in I_1, \tag{1.2.9}$$

where $G^{-1} : G(J_1) \to J_1$ is the inverse function of G. The constant C is uniquely determined by the initial condition $y(x_0) = y_0$, i.e. $C = G(y_0) - F(x_0)$.

This argument identifies two kinds of solutions to the equation (1.2.6): constant solutions and solutions such that $g(y) \neq 0$. There could be other solutions (see Section 5). If e.g. $g \in C^1(J)$, however, there are no other solutions (see Chapter 2).

EXAMPLE 1.2.1. We look for the solution to the Cauchy Problem

$$\begin{cases} y' = \dfrac{1 + 2x}{\cos y} \\ y(0) = \pi. \end{cases} \tag{1.2.10}$$

The differential equation is of the form $y' = f(x)g(y)$ with $f(x) = 1 + 2x$ and $g(y) = 1/\cos y$. ular, g is defined for $\cos y \neq 0$, i.e. for $y \neq \pi/2 + k\pi$ with $k \in \mathbb{Z}$.

Separating the variables we get $y' \cos y = 1 + 2x$, and integrating we find the general solution in implicit form

$$\sin y = x + x^2 + C,$$

where $C \in \mathbb{R}$ is a constant, which is determined by the initial condition $y(0) = \pi$, i.e. $C = \sin y(0) = 0$. The function

$$z(x) = \arcsin(x + x^2)$$

is not, however, the solution to (1.2.10) because $z(0) = \arcsin(0) = 0$. In order to determine the correct solution, notice the arcsin is the inverse function of sin when restricted to $[-\pi/2, \pi/2]$, whereas y takes values in a neighborhood of π. Letting $w(x) = y(x) - \pi$, we have $w(0) = y(0) - \pi = 0$ and $\sin w = \sin(y - \pi) = -\sin y = -(x + x^2)$. Now we can invert the sine function, obtaining $w = -\arcsin(x + x^2)$ and thus

$$y(x) = \pi - \arcsin(x + x^2).$$

The solution y is defined in a suitable neighborhood of the origin.

3. Equations of homogeneous type

A differential equation of the form

$$y' = f\Big(\frac{y}{x}\Big)$$

is called of homogeneous type. Here $f : I \to \mathbb{R}$ is a (continuous) function in some interval $I \subset \mathbb{R}$. With the change of variable $y = xz$, where z is the new unknown function, we get $y' = z + xz'$ and the differential equation transforms into

$$xz' + z = f(z).$$

This equation can be solved on separating the variables.

For instance, the following differential equation is of homogeneous type (see Exercise **??**)

$$y' = \frac{x^2 + y^2}{xy} = f\Big(\frac{y}{x}\Big), \quad \text{with} \quad f(t) = \frac{1}{t} + t.$$

4. Bernoulli's equations

A differential equation of the form

$$y' + a(x)y = b(x)y^\alpha, \quad x \in I, \tag{1.4.11}$$

where α is a real parameter such that $\alpha \neq 0, 1$ is said to be of Bernoulli type. Letting

$$y = z^{\frac{1}{1-\alpha}}, \quad y' = \frac{1}{1-\alpha} z^{\frac{\alpha}{1-\alpha}} z',$$

the equations transforms into

$$z' + (1-\alpha)a(x)z = (1-\alpha)b(x).$$

This is a linear equation.

5. Exact equations

Let $\Omega \subset \mathbb{R}^2$ be an open set and $f, g \in C(\Omega)$ be continuous function. We look for a solution $y \in C^1(I)$, $I \subset \mathbb{R}$ open interval, of the differential equation

$$f(x,y) + g(x,y)y' = 0, \tag{1.5.12}$$

satisfying the condition $y(x_0) = y_0$ for some $(x_0, y_0) \in \Omega$ with $x_0 \in I$.

To this aim, consider the differential form ω in Ω

$$\omega = f(x,y)dx + g(x,y)dy, \tag{1.5.13}$$

where dx, dy is the dual basis of the basis $e_1 = (1,0)$ and $e_2 = (0,1)$.

Definition 1.5.1 (Exact forms). The differential form ω is exact if there exists a function $F \in C^1(\Omega)$ such that

$$\omega = dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy \quad \text{in } \Omega.$$

The function F is called a potential of ω. In this case, the differential equation (1.5.12) is called exact.

Theorem 1.5.2. *Assume that ω is an exact form with potential F such that*

$$\frac{\partial F(x_0,y_0)}{\partial y} \neq 0. \tag{1.5.14}$$

Then the equation $F(x,y) = F(x_0,y_0)$ implicitly defines a function $y \in C^1(I)$ for some open interval I containing x_0 solving the differential equation (1.5.12) *along with the condition $y(x_0) = y_0$. This solution is unique on the interval I.*

Proof. Assume w.l.g. that $F(x_0,y_0) = 0$. By the implicit function theorem, there exist $\delta, \eta > 0$ and $y \in C^1(x_0 - \delta, x_0 + \delta)$ such that

$$\big\{(x,y) \in \Omega : |x - x_0| < \delta, |y - y_0| < \eta, F(x,y) = 0\big\} = \big\{(x, y(x)) \in \Omega : |x - x_0| < \delta\big\}. \tag{1.5.15}$$

Differentiating the identity $F(x,y(x)) = 0$ we get

$$\begin{aligned} 0 = \frac{d}{dx}F(x,y(x)) &= \frac{\partial F(x,y(x))}{\partial x} + \frac{\partial F(x,y(x))}{\partial y}y'(x) \\ &= f(x,y(x)) + g(x,y(x))y'(x), \end{aligned} \tag{1.5.16}$$

i.e. y is a solution of the differential equation, and moreover $y(x_0) = y_0$.

On the other hand, if $z \in C^1(I)$ is a solution to the equation (1.5.12) such that $z(x_0) = y_0$, then the same argument as in (1.5.16) shows that

$$\frac{d}{dx}F(x,z(x)) = 0,$$

and therefore $F(x,z(x)) = F(x_0, z(x_0)) = F(x_0,y_0) = 0$. By (1.5.15) it must by $z = y$. □

DEFINITION 1.5.3 (Closed forms). Assume that $f, g \in C^1(\Omega)$. The differential form $\omega = f dx + g dy$ is closed in Ω if

$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial g(x,y)}{\partial x} \quad \text{for all } (x,y) \in \Omega.$$

An exact differential form in an open set Ω with a potential $F \in C^2(\Omega)$ is closed in Ω because mixed derivatives are equal by Schwarz theorem

$$\frac{\partial^2 F}{\partial x \partial y}(x,y) = \frac{\partial^2 F}{\partial y \partial x}(x,y), \quad (x,y) \in \Omega.$$

The converse is also true if Ω is simply connected.

THEOREM 1.5.4. *If $\Omega \subset \mathbb{R}^2$ is a simply connected open set, then any closed differential form in Ω is exact.*

Convex and starshaped open sets are simply connected. In particular, closed forms always have a potential locally.

If $\varphi \in C(\Omega)$ is a function such that $\varphi \neq 0$ in Ω, then the differential equation (1.5.12) and the differential equation

$$\varphi(x,y)\big\{f(x,y) + g(x,y)y'\big\} = 0, \tag{1.5.17}$$

have the same solutions. For a suitable choice of φ, the differential equation (1.5.17) may happen to be exact, even though (1.5.12) is not exact. The function φ is then called integrating factor (or multiplier). If $f, g \in C^1(\Omega)$, a necessary condition for a function $\varphi \in C^1(\Omega)$ to be a multiplier is

$$\frac{\partial}{\partial y}\varphi f = \frac{\partial}{\partial x}\varphi g \quad \text{in } \Omega. \tag{1.5.18}$$

EXAMPLE 1.5.5. The differential equation

$$xy^2 + y - xy' = 0 \tag{1.5.19}$$

is not exact. In fact, with $f = xy^2 + y$ and $g = -x$, we have

$$\frac{\partial f(x,y)}{\partial y} = 2xy + 1 \quad \text{and} \quad \frac{\partial g(x,y)}{\partial x} = -1.$$

We look for a function φ such that (1.5.18) holds. We try with the ansatz $\varphi = \varphi(y)$, i.e. φ depends only on y. We get the necessary condition

$$\varphi'(y)(xy^2 + y) + \varphi(y)(2xy + 1) = -\varphi(y),$$

that is implied by $y\varphi' + 2\varphi = 0$ (this equation does not depend on x). A solution for this linear equation is $\varphi(y) = 1/y^2$. Now the differential equation

$$x + \frac{1}{y} - \frac{xy'}{y^2} = 0$$

is exact, where $y \neq 0$. A potential F for this exact differential equation can be found on solving

[1]

An integration yields

$$F(x,y) = \frac{x^2}{2} + \frac{x}{y} + C, \quad C \in \mathbb{R}.$$

The equation $F(x,y) = 0$ implicitly defines solutions to the differential equation (1.5.19) (for $x \neq 0$ and $y \neq 0$).

6. Second order linear equations with constant coefficients

Let $f : I \to \mathbb{R}$ be a continuous function on the interval $I \subset \mathbb{R}$ and let $a, b, c \in$ be real numbers such that $a \neq 0$. The differential equation

$$ay'' + by' + cy = f(x), \quad x \in I, \tag{1.6.20}$$

is a second order linear differential equation with constant coefficients. When $f = 0$ the equation is called homogeneous. The general solution of the homogeneous equation

$$ay'' + by' + cy = 0, \quad x \in \mathbb{R}, \tag{1.6.21}$$

is of the form $y_{GH} = C_1 y_1 + C_2 y_2$ where $C_1, C_2 \in \mathbb{R}$ are real numbers and y_1, y_2 are two solutions of (1.6.21) which are linearly independent, i.e. such that for real numbers $\alpha, \beta \in \mathbb{R}$

$$\alpha y_1 + \beta y_2 = 0 \text{ in } \mathbb{R} \quad \Rightarrow \quad \alpha = \beta = 0.$$

The general solution of the inhomogeneous equation (1.6.20) is a function $y \in C^2(I)$ of the form $y = y_{GH} + y_P$, where $y_P \in C^2(I)$ is a particular solution of the inhomogeneous equations. We describe some practical methods to compute y_{GH} and y_P. The general theory is dealt with in Chapter 3.

6.1. Homogeneous equation. The solutions to the homogeneous equation

$$ay'' + by' + cy = 0 \tag{1.6.22}$$

are a real vector space, i.e. any linear combination of solutions is still a solution. We shall prove in Chapter 3 that this vector space has dimension 2. It is therefore sufficient to find two linearly independent solutions to the equation. We look for solutions of the form $y(x) = e^{\lambda x}$ for some complex number $\lambda \in \mathbb{C}$. Inserting y, y', y'' into (1.6.22) we get $e^{\lambda x}(a\lambda^2 + b\lambda + c) = 0$. Since $e^{\lambda x} \neq 0$, the complex number λ must solve the characteristic equation

$$a\lambda^2 + b\lambda + c = 0. \tag{1.6.23}$$

According to the sign of $\Delta = b^2 - 4ac$ we distinguish three cases.

Case 1: $\Delta > 0$. In this case the characteristic equation has two real (simple) solutions

$$\lambda_1 = \frac{-b + \sqrt{\Delta}}{2a}, \quad \lambda_2 = \frac{-b - \sqrt{\Delta}}{2a}.$$

[1]MANCA TESTO

The general solution of the homogeneous equation (1.6.22) is

$$y_{GH}(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x},$$

where $C_1, C_2 \in \mathbb{R}$ are real constant.

Case 2: $\Delta = 0$. In this case the characteristic equation $a\lambda^2 + b\lambda + c = 0$ has one real double solution $\lambda = -b/2a$. The ansatz yields only the solution $y_1(x) = e^{\lambda x}$. A direct computation shows that the function $y_2(x) = xe^{\lambda x}$ is also a solution which is linearly independent from the first one. The general solution of the homogeneous equation (1.6.22) is then

$$y_{GH}(x) = e^{\lambda x}(C_1 + C_2 x),$$

where $C_1, C_2 \in \mathbb{R}$ are real constants.

Case 3: $\Delta < 0$. In this case the characteristic equation has two complex conjugate solutions

$$\lambda_1 = \alpha + i\beta \quad \text{and} \quad \lambda_2 = \alpha - i\beta, \quad \text{where} \quad \alpha = -\frac{b}{2a}, \quad \beta = \frac{\sqrt{|\Delta|}}{2a}.$$

We get the complex valued solutions

$$z_1(x) = e^{(\alpha+i\beta)x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)),$$
$$z_2(x) = e^{(\alpha-i\beta)x} = e^{\alpha x}(\cos(\beta x) - i\sin(\beta x)),$$

and the real valued solutions

$$y_1(x) = \frac{1}{2}(z_1(x) + z_2(x)) = e^{\alpha x}\cos(\beta x),$$
$$y_2(x) = \frac{1}{2i}(z_1(x) - z_2(x)) = e^{\alpha x}\sin(\beta x).$$

The general solution of the homogeneous equation is

$$y_{GH}(x) = e^{\alpha x}\big(C_1\cos(\beta x) + C_2\sin(\beta x)\big),$$

where $C_1, C_2 \in \mathbb{R}$ are real constants.

6.2. Inhomogeneous equation. Similar solutions. Consider the inhomogeneous equation (1.6.20), where the right hand side is a function $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = e^{\alpha x}\big(P_0(x)\cos(\beta x) + Q_0(x)\sin(\beta x)\big),$$

for some $\alpha, \beta \in \mathbb{R}$ and real polynoms P_0 and Q_0.

We describe a practical method to find a particular solution y_P to the differential equation

$$ay'' + by' + cy = f(x). \tag{1.6.24}$$

A more systematic method is given in the next section.

Consider the real or complex number $\lambda = \alpha + i\beta$ and denote by $m \in \{0, 1, 2\}$ the multiplicity of λ as a solution of the characteristic equation (1.6.23). The case $m = 0$ means that λ is not a solution, the case $m = 1$ means that λ is a simple solution, the case $m = 2$ means that λ is a double solution.

It is always possible to find a particular solution y_P for the inhomogeneous equation (1.6.24) of the form

$$y_P(x) = x^m e^{\alpha x}\big(P(x)\cos(\beta x) + Q(x)\sin(\beta x)\big),$$

where P and Q are unknown polynoms with degree equal to the maximum of the degrees of P_0 and Q_0. These polynoms can be determined inserting y_P, y_P', y_P'' into (1.6.24) and comparing the coefficients of the trigonometric functions in the left and right hand side.

6.3. Inhomogeneous equation. Variation of constants. We look for a particular solution of the inhomogeneous equation (1.6.20) when $f : I \to \mathbb{R}$ is any continuous function on some interval $I \subset \mathbb{R}$.

The general solution of the homogeneous equation (1.6.21) is of the form

$$y = C_1 y_1 + C_2 y_2, \tag{1.6.25}$$

where C_1, C_2 are two real constants and y_1 and y_2 are two linearly independent solutions of the homogeneous equation. These solutions are known.

The method of the variation of constants consists in letting C_1, C_2 be functions of the variable x. We look for a particular solution y of (1.6.20) of the form (1.6.25), where now C_1 and C_2 are functions. We have to determine C_1 and C_2. On differentiating y, we get

$$y' = C_1' y_1 + C_1 y_1' + C_2' y_2 + C_2 y_2'. \tag{1.6.26}$$

We impose on C_1' and C_2' the condition

$$C_1' y_1 + C_2' y_2 = 0. \tag{1.6.27}$$

On differentiating y' we find

$$y'' = C_1' y_1' + C_1 y_1'' + C_2' y_2' + C_2 y_2''. \tag{1.6.28}$$

Plugging (1.6.25), (1.6.26), (1.6.28) into the inhomogeneous equation (1.6.20) we find

$$a(C_1' y_1' + C_2' y_2') = f(x). \tag{1.6.29}$$

We also used (1.6.27) and the fact that y_1, y_2 are solution to the homogeneous equation.

As y_1 and y_2 are linearly independent, it is

$$\det\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \neq 0. \tag{1.6.30}$$

This fact follows from Exercise **??**. Then it is possible to solve the system of equations

$$\begin{cases} C_1' y_1 + C_2' y_2 = 0 \\ C_1' y_1' + C_2' y_2' = f(x)/a. \end{cases} \tag{1.6.31}$$

Finally, the functions C_1 and C_2 can be computed on integrating C_1' and C_2'.

7. Euler's second order equations

Let $f : I \to \mathbb{R}$ be a continuous function on the interval $I \subset \mathbb{R}^+$ and let $a, b, c \in$ be real numbers such that $a \neq 0$. The differential equation

$$ax^2y'' + bxy' + cy = f(x), \quad x \in I, \tag{1.7.32}$$

is a second order differential equation of Euler's type. We consider only the case $f = 0$, i.e. the homogeneous equation

$$ax^2y'' + bxy' + cy = 0, \quad x \in \mathbb{R}^+. \tag{1.7.33}$$

The differential equation is singular at $x = 0$ because the coefficient of y'' vanishes. We look for solutions on the half line $\mathbb{R}^+ = (0, +\infty)$. For the differential equation is linear, solutions are a two dimensional vector space. We look for two linearly independent solutions of the form

$$y(x) = x^\lambda = e^{\lambda \log(x)} = e^{(\alpha+i\beta)\log x} = x^\alpha(\cos(\beta \log x) + i \sin(\beta \log x)),$$

where $\lambda = \alpha + i\beta$ is a complex parameter. Plugging y, $y' = \lambda x^{\lambda-1}$, and $y'' = \lambda(\lambda-1)x^{\lambda-2}$ into (1.7.33) we get $x^\lambda(a\lambda(\lambda-1) + b\lambda + c) = 0$. Because $x^\lambda \neq 0$, λ must solve the characteristic equation

$$a\lambda^2 + (b-a)\lambda + c = 0. \tag{1.7.34}$$

According to the sign of $\Delta = (b-a)^2 - 4ac$ we distinguish three cases.

Case 1: $\Delta > 0$. In this case the characteristic equation has two real (simple) solutions $\lambda_1, \lambda_2 \in \mathbb{R}$ and the general solution of the homogeneous equation (1.7.33) is

$$y(x) = C_1x^{\lambda_1} + C_2x^{\lambda_2},$$

where $C_1, C_2 \in \mathbb{R}$ are real constant.

Case 2: $\Delta = 0$. In this case the characteristic equation has one real double solution $\lambda \in \mathbb{R}$ and we get the solution $y_1(x) = x^\lambda$. A direct computation shows that the function $y_2(x) = x^\lambda \log x$ is also a solution which is linearly independent from the first one. The general solution of the homogeneous equation (1.7.33) is then

$$y(x) = x^\lambda(C_1 + C_2 \log x),$$

where $C_1, C_2 \in \mathbb{R}$ are real constants.

Case 3: $\Delta < 0$. In this case the characteristic equation has two complex conjugate solutions

$$\lambda_1 = \alpha + i\beta \quad \text{and} \quad \lambda_2 = \alpha - i\beta.$$

We get the complex valued solutions

$$z_1(x) = x^{\alpha+i\beta} = x^\alpha\big(\cos(\beta \log x) + i \sin(\beta \log x)\big),$$
$$z_2(x) = x^{\alpha-i\beta} = x^\alpha\big(\cos(\beta \log x) - i \sin(\beta \log x)\big),$$

and the real valued solutions

$$y_1(x) = \frac{1}{2}(z_1(x) + z_2(x)) = x^\alpha \cos(\beta \log x),$$
$$y_2(x) = \frac{1}{2i}(z_1(x) - z_2(x)) = x^\alpha \sin(\beta \log x).$$

The general solution of the homogeneous equation is

$$y(x) = x^\alpha\big(C_1 \cos(\beta \log x) + C_2 \sin(\beta \log x)\big),$$

where $C_1, C_2 \in \mathbb{R}$ are real constants.

CHAPTER 2

Existence and uniqueness in the Lipschitz case

1. Banach fixed point theorem

DEFINITION 2.1.1 (Contraction). Let (X, d) be a metric space. A mapping $T : X \to X$ is a *contraction* if there exists $0 < \lambda < 1$ such that $d(T(x), T(y)) \le \lambda d(x, y)$ for all $x, y \in X$.

THEOREM 2.1.2. *Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction. Then there exists a unique $x \in X$ such that $x = T(x)$.*

PROOF. Let $x_0 \in X$ be a given point and define $x_n = T^n(x_0) = T \circ ... \circ T(x_0)$, n-times. The sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy. In fact, by the triangle inequality we have for all $n, k \in \mathbb{N}$

$$\begin{aligned} d(x_{n+k}, x_n) &\le \sum_{h=1}^{k} d(x_{n+h}, x_{n+h-1}) = \sum_{h=1}^{k} d(T^{n+h}(x_0), T^{n+h-1}(x_0)) \\ &\le d(T(x_0), x_0) \sum_{h=1}^{k} \lambda^{n+h-1} \le \lambda^n d(T(x_0), x_0) \sum_{h=1}^{\infty} \lambda^{h-1}. \end{aligned}$$

The series converges and $\lambda^n \to 0$ as $n \to \infty$, because $\lambda < 1$. For X is complete, there exists $x \in X$ such that $x = \lim_{n\to\infty} T^n(x_0)$.

We show that $x = T(x)$. The mapping T is continuous and so we have

$$x = \lim_{n\to\infty} T^n(x_0) = \lim_{n\to\infty} T(T^{n-1}(x_0)) = T(\lim_{n\to\infty} T^{n-1}(x_0)) = T(x).$$

Finally, we prove that the fixed point is unique. Let $\bar{x} \in X$ be such that $\bar{x} = T(\bar{x})$. Then we have

$$d(x, \bar{x}) = d(T(x), T(\bar{x})) \le \lambda d(x, \bar{x}) \quad \Rightarrow \quad d(x, \bar{x}) = 0,$$

because $\lambda < 1$, and thus $x = \bar{x}$.

□

THEOREM 2.1.3. *Let (X, d) be a complete metric space and let $T : X \to X$ be a mapping such that T^n is a contraction, for some $n \in \mathbb{N}$. Then there exists a unique $x \in X$ such that $x = T(x)$.*

PROOF. There exists a unique $x \in X$ such that $T^n(x) = x$. Then we have for some $0 \le \lambda < 1$

$$d(x, T(x)) = d(T^n(x), T(T^n(x))) = d(T^n(x), T^n(T(x))) \le \lambda d(x, T(x)),$$

and thus $d(x, T(x)) = 0$, which is equivalent to $T(x) = x$.

Now assume that for some $y \in X$ it is $y = T(y)$. Then we also have $y = T^n(y)$ and thus $x = y$, because the fixed point of T^n is unique. □

2. Excursus. Other fixed point theorems

Some fixed point theorems are based on the notion of convexity.

THEOREM 2.2.1 (Brouwer). *Let $K \subset \mathbb{R}^n$ be a closed ball and let $T : K \to K$ be continuous. Then there exists $x \in K$ such that $T(x) = x$.*

For an analytical proof, see Evans, *Partial Differential Equations*, p.441. Brouwer's theorem extends to the infinite dimensional case.

THEOREM 2.2.2 (Schauder I). *Let $(X, \|\cdot\|)$ be a Banach space and let $K \subset X$ be a nonempty, convex, and compact set. Any continuous mapping $T : K \to K$ has at least one fixed point in K, i.e., there exists $x \in K$ tale che $T(x) = x$.*

See Evans, *Partial Differential Equations*, p.502. The assumption on K to be compact can be be transferred to the mapping T.

THEOREM 2.2.3 (Schauder II). *Let $(X, \|\cdot\|)$ be a Banach space and let $K \subset X$ be a nonempty, convex, closed and bounded set. Let $T : K \to K$ be a mapping such that:*

(i) *T is continuous;*
(ii) *T is compact, i.e., $T(K) \subset K$ is precompact.*

Then there exists $x \in K$ such that $T(x) = x$.

Tarki's Fixed Point theorem relies upon the notion of partial order.

THEOREM 2.2.4 (Tarski). *Let $(X, \leq)$ be a partially ordered set such that any subset $Y \subset X$ has a supremum. Let $T : X \to X$ be an order preserving mapping, i.e. a mapping such that*

$$x \leq y \;\Rightarrow T(x) \leq T(y).$$

Then there exists $x \in X$ such that $x = T(x)$.

The proof of Tarki's Lemma is an exercise.

3. Cauchy Problem. Introduction

In $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$, $n \geq 1$, we introduce the coordinates $x \in \mathbb{R}$ and $y \in \mathbb{R}^n$. Let $\Omega \subset \mathbb{R}^{n+1}$ be an open set and let $f \in C(\Omega; \mathbb{R}^n)$ be a continuous function. Given a point $(x_0, y_0) \in \Omega$ we consider the Cauchy Problem

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0. \end{cases} \tag{2.3.35}$$

A function $y \in C^1(I; \mathbb{R}^n)$ is a solution to the problem if:

i) $I \subset \mathbb{R}$ is an interval such that $x_0 \in I$;

ii) $(x, y(x)) \in \Omega$ for all $x \in I$;
iii) $y'(x) = f(x, y(x))$ for all $x \in I$ (the differential equations is solved);
iv) $y(x_0) = y_0$ (the initial datum is attained).

We are interested in the following questions:

a) Existence of solutions;
b) Uniqueness of solutions;
c) Dependence of solutions from x_0, y_0, and f (regularity and stability).

Integrating the differential equation $y' = f(x, y)$ on the interval with end-points x_0 and x we get the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt = Ty(x), \tag{2.3.36}$$

where $y \mapsto Ty$ is a mapping defined on a suitable functional space. A solution to the Cauchy Problem is then a fixed point of the mapping T. On the other hand, if a continuous function y solves the fixed point equation (2.3.36) then y is of class C^1 and solves the Cauchy Problem (2.3.35).

We fix the functional space. For a $\delta > 0$ consider the real vector space

$$V = C([x_0 - \delta, x_0 + \delta]; \mathbb{R}^n). \tag{2.3.37}$$

Endowed with the norm

$$\|y\| = \max_{x \in [x_0 - \delta, x_0 + \delta]} |y(x)|, \quad y \in V, \tag{2.3.38}$$

the vector space V is a Banach space, because any Cauchy sequence in V converges to a function in V. For any $\varepsilon > 0$, the subset X of V

$$X = \big\{ y \in V : y(x_0) = y_0,\ \|y - y_0\| \le \varepsilon \big\} \tag{2.3.39}$$

is closed because both conditions $y(x_0) = y_0$ and $\|y - y_0\| \le \varepsilon$ are preserved by the uniform (pointwise) convergence. Then the metric space (X, d) is complete w.r.t. the metric $d(y, z) = \|y - z\|$.

We shall see that for a suitable choice of δ and ε the mapping $T : X \to X$

$$Ty(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt \tag{2.3.40}$$

is well defined, i.e. it is $Ty \in X$ for all $y \in X$.

4. Local existence and uniqueness under the Lipschitz condition

Definition 2.4.1. Let $\Omega \subset \mathbb{R}^{n+1}$ be an open set. We say that a function $f \in C(\Omega; \mathbb{R}^n)$ has the local Lipschitz property in y if for any compact set $K \subset \Omega$ there exists a constant $L > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \le L |y_1 - y_2| \tag{2.4.41}$$

for all $(x, y_1), (x, y_2) \in K$.

THEOREM 2.4.2. *Let $\Omega \subset \mathbb{R}^{n+1}$ be an open set, $(x_0, y_0) \in \Omega$, and let $f \in C(\Omega; \mathbb{R}^n)$ be a function with the local Lipschitz property in y. Then there exists $\delta > 0$ such that the Cauchy Problem (2.3.35) has a unique solution $y \in C^1(I; \mathbb{R}^n)$ in the interval $I = [x_0 - \delta, x_0 + \delta]$.*

PROOF. Let $\delta > 0$ and $\varepsilon > 0$ be such that $K = [x_0 - \delta, x_0 + \delta] \times \{y \in \mathbb{R}^n : |y - y_0| \leq \varepsilon\} \subset \Omega$. Let $H \subset \Omega$ be any compact set such that $K \subset \mathrm{int}(H)$. For f is continuous on H, the number

$$M = \sup_{(x,y) \in H} |f(x, y)| < +\infty$$

is finite. Let X be the set introduced in (2.3.39) and let T be the mapping (2.3.40). For any $y \in X$ we have for $x \in I$

$$|Ty(x) - y_0| \leq \Big| \int_{x_0}^{x} |f(t, y(t))| dt \Big| \leq M|x - x_0| \leq \delta M.$$

In fact it is $(t, y(t)) \in K$ for all $t \in I$. Possibly choosing a smaller $\delta > 0$ (this does not affect M), we can assume that $\delta M \leq \varepsilon$. With such a choice, it is $Ty \in X$ for all $y \in X$. The choice of $\delta > 0$ is independent from x_0 and y_0 as long as $K \subset \mathrm{int}(H)$.

We prove that the mapping $T : X \to X$ has a unique fixed point. It is enough to show that, for some $k \in \mathbb{N}$, the iterated mapping T^k is a contraction. Let $y, \bar{y} \in X$ and $x \in I$. We have (with e.g. $x \geq x_0$)

$$\begin{aligned} |Ty(x) - T\bar{y}(x)| &= \Big| \int_{x_0}^{x} \big(f(t, y(t)) - f(t, \bar{y}(t)\big) dt \Big| \\ &\leq \int_{x_0}^{x} |f(t, y(t)) - f(t, \bar{y}(t)| dt \\ &\leq L \int_{x_0}^{x} |y(t) - \bar{y}(t)| dt \leq L|x - x_0| \cdot \|y - \bar{y}\|. \end{aligned}$$

Here, L is the Lipschitz constant for f relative to the compact set H. Analogously, it is (e.g. $x \geq x_0$)

$$\begin{aligned} |T^2 y(x) - T^2 \bar{y}(x)| &= \Big| \int_{x_0}^{x} (f(t, Ty(t)) - f(t, T\bar{y}(t)) dt \Big| \\ &\leq L \int_{x_0}^{x} |Ty(t) - T\bar{y}(t)| dt \\ &\leq L^2 \|y - \bar{y}\| \int_{x_0}^{x} (t - x_0) dt \leq L^2 \frac{(x - x_0)^2}{2} \|y - \bar{y}\|. \end{aligned}$$

By induction, we get for any $k \in \mathbb{N}$ and $x \in I$

$$|T^k y(x) - T^k \bar{y}(x)| \leq \frac{L^k |x - x_0|^k}{k!} \|y - \bar{y}\|,$$

which implies

$$\|T^k y - T^k \bar{y}\| \leq \frac{(L\delta)^k}{k!} \|y - \bar{y}\|.$$

Because

$$\lim_{k\to+\infty} \frac{(L\delta)^k}{k!} = 0,$$

there exists $k \in \mathbb{N}$ such that $\frac{(L\delta)^k}{k!} < 1$. For such a k, the mapping T^k is a contraction. Then T has a unique fixed point $y \in X$. Then it is $y \in C^1([x_0 - \delta, x_0 + \delta]; \mathbb{R}^n)$ and y solves the Cauchy Problem (2.3.35). □

5. Peano's example

Consider the Cauchy problem

$$\begin{cases} y'(x) = 2\sqrt{|y(x)|}, & x \in \mathbb{R}, \\ y(0) = 0 \end{cases} \tag{2.5.42}$$

The function $f(x, y) = 2\sqrt{|y|}$ is not locally Lipschitz in the variable y. The Lipschitz property (2.4.41) fails in a neighborhood of $y = 0$ and the assumptions of Theorem 2.4.2 are not fulfilled. The Cauchy Problem could have more than one solution.

In fact, a solution is the constant function $y = 0$. A second solution can be found separating the variables: $2 = y'/\sqrt{|y|}$. Integrating this equation on the interval between 0 and $x \in \mathbb{R}$ we get

$$2x = \int_0^x \frac{y'(t)}{\sqrt{|y(t)|}} dt = \int_0^{y(x)} \frac{1}{\sqrt{|z|}} dz = \begin{cases} 2\sqrt{y(x)}, & \text{if } y(x) > 0 \\ -2\sqrt{-y(x)}, & \text{if } y(x) < 0. \end{cases}$$

In the change of variable $z = y(t)$ we used the initial datum $y(0) = 0$. Then we find the solution $y \in C^1(\mathbb{R})$

$$y(x) = \begin{cases} x^2 & \text{if } x \geq 0, \\ -x^2 & \text{if } x < 0. \end{cases}$$

On the other hand, for all real numbers $\alpha \leq 0 \leq \beta$, the function

$$y_{\alpha\beta}(x) = \begin{cases} (x-\beta)^2 & \text{if } x \geq \beta, \\ 0 & \text{if } \alpha < x < \beta, \\ -(x-\alpha)^2 & \text{if } x \leq \alpha \end{cases}$$

is of class $C^1(\mathbb{R})$ and solves the Cauchy Problem (2.5.42). So there is a continuum of solutions.

6. Maximal solutions

Let $f \in C(\Omega; \mathbb{R}^n)$ be a function satisfying the local Lipschitz condition (2.4.41) and let $(x_0, y_0) \in \Omega$.

PROPOSITION 2.6.1. *Under the hypotheses of Theorem 2.4.2, let I_1 and I_2 be two open intervals containing x_0 and assume that $y_1 \in C^1(I_1; \mathbb{R}^n)$ and $y_2 \in C^1(I_2; \mathbb{R}^n)$ are solutions to the Cauchy Problem* (2.3.35)*. Then it is $y_1 = y_2$ on $I_1 \cap I_2$.*

Proof. The set $A = \{x \in I_1 \cap I_2 : y_1(x) = y_2(x)\}$ is relatively closed in $I_1 \cap I_2$ because y_1 and y_2 are continuous. We show that A is also open in $I_1 \cap I_2$. Since $I_1 \cap I_2$ is connected it then follows that $A = I_1 \cap I_2$.

Let $\bar{x}_0 \in A$ and $\bar{y}_0 = y_1(\bar{x}_0) = y_2(\bar{x}_0)$. By Theorem 2.4.2 there exists $\delta > 0$ such that the Cauchy Problem

$$\begin{cases} y' = f(x, y) \\ y(\bar{x}_0) = \bar{y}_0 \end{cases} \tag{2.6.43}$$

has a unique solution $y \in C^1(I; \mathbb{R}^n)$ with $I = [\bar{x}_0 - \delta, \bar{x}_0 + \delta]$. For a small $\delta > 0$ it is $I \subset I_1 \cap I_2$. It then follows that $y = y_1 = y_2$ in I, and thus $I \subset A$. □

Consider the family $\mathcal{A}$ of all pairs (J, y_J) where $J \subset \mathbb{R}$ is an open interval containing x_0 and $y_J \in C^1(J; \mathbb{R}^n)$ is a solution to the Cauchy Problem (2.3.35). By Theorem 2.4.2, it is $\mathcal{A} \neq \emptyset$.

Let $I \subset \mathbb{R}$ be the interval $I = \bigcup J$, where the union is over all intervals J such that $(J, y_J) \in \mathcal{A}$. Let $y \in C^1(I; \mathbb{R}^n)$ be the function defined by

$$y(x) = y_J(x) \quad \text{if} \quad x \in J. \tag{2.6.44}$$

The function y is well defined because by Proposition 2.6.1 it is $y_J = y_{J'}$ on $J \cap J'$. Moreover, y is a solution to the Cauchy Problem (2.3.35).

Definition 2.6.2 (Maximal solution). The function y defined in (2.6.44) is called maximal solution to the Cauchy Problem (2.3.35).

Theorem 2.6.3 (Continuation criterion). *Let $I = (a_0, b_0) \subset \mathbb{R}$ be an open interval with $-\infty < a_0 < b_0 < +\infty$, $\Omega = I \times \mathbb{R}^n$, and $f \in C(\Omega; \mathbb{R}^n)$ be a function satisfying the local Lipschitz property in y. If $y \in C^1((a, b); \mathbb{R}^n)$ is the maximal solution to the Cauchy Problem* (2.3.35)*, for some interval $(a, b) \subset (a_0, b_0)$, then we have either*

i) $b = b_0$; *or,*
ii) $\lim_{x \uparrow b} |y(x)| = +\infty$.

There is an analogous statement for a.

Proof. Assume by contradiction that $b < b_0$ and there exists a sequence $x_k \in (a, b)$, $k \in \mathbb{N}$, such that

$$\lim_{k \to +\infty} x_k = b \quad \text{and} \quad \sup_{k \in \mathbb{N}} |y(x_k)| \leq M_0,$$

for some constant $M_0 < +\infty$. Letting $\bar{y}_k = y(x_k) \in \mathbb{R}^n$, possibly taking a subsequence we can assume that

$$\lim_{k \to +\infty} \bar{y}_k = \bar{y}_0$$

for some $\bar{y}_0 \in \mathbb{R}^n$.

We study the Cauchy Problem

$$\begin{cases} z'(x) = f(x, z(x)) \\ z(x_k) = \bar{y}_k. \end{cases} \tag{2.6.45}$$

Fix a compact set $H \subset \Omega$ such that $(b, \bar{y}_0) \in \text{int}(H)$ and let

$$M = \max_{(x,y)\in H} |f(x,y)| < +\infty.$$

For some $\varepsilon > 0$ and for some large enough $k \in \mathbb{N}$, the compact set

$$K = [x_k, 2b - x_k] \times \{y \in \mathbb{R}^n : |y - \bar{y}_k| \leq \varepsilon\}$$

is contained in H. Let us introduce the functional space

$$X = \big\{z \in C([x_k, 2b - x_k]; \mathbb{R}^n) : z(x_k) = \bar{y}_k,\ \|z - \bar{y}_k\| \leq \varepsilon\big\}.$$

For large enough $k \in \mathbb{N}$ we also have $2(b - x_k)M \leq \varepsilon$. Then, the integral operator

$$Tz(x) = \bar{y}_k + \int_{x_k}^{x} f(t, z(t))dt$$

maps X into itself, i.e. $T : X \to X$.

As in the proof of Theorem 2.4.2, some iterated of T is a contraction on X and therefore by Theorem 2.1.3 there exists a unique solution $z \in C^1([x_k, 2b - x_k]; \mathbb{R}^n)$ to the Cauchy Problem (2.6.45).

On the other hand, the function y solves the same Cauchy Problem on the interval $[x_k, b)$ and by uniqueness it is $y = z$ on $[x_k, b)$. This shows that y can be continued as a solution to the Cauchy Problem (2.3.35) beyond b. This contradicts the maximality of y. □

7. Gronwall's Lemma

LEMMA 2.7.1. *Let $I \subset \mathbb{R}$ be an interval, $x_0 \in I$, and $\varphi \in C(I)$ be a non negative $\varphi \geq 0$ continuous function. If there exist $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta \geq 0$, such that*

$$\varphi(x) \leq \alpha + \beta \int_{x_0}^{x} \varphi(t)\, dt, \quad \text{for all } x \in I \text{ with } x \geq x_0, \tag{2.7.46}$$

then

$$\varphi(x) \leq \alpha e^{\beta(x - x_0)} \quad \text{for all } x \in I \text{ with } x \geq x_0. \tag{2.7.47}$$

PROOF. Let $\Phi : I \to \mathbb{R}$ be the function

$$\Phi(x) = \alpha + \beta \int_{x_0}^{x} \varphi(t)\, dt.$$

It is $\Phi \in C^1(I)$ and moreover, $\Phi'(x) = \beta\varphi(x)$ for all $x \in I$, by the Fundamental Theorem of Calculus. From (2.7.46) it follows that $\Phi'(x) \leq \beta\Phi(x)$ for $x \in I$, because $\beta \geq 0$. The function $\Psi(x) = e^{-\beta(x - x_0)}\Phi(x)$ satisfies

$$\Psi'(x) = -\beta e^{-\beta(x-x_0)}\Phi(x) + e^{-\beta(x-x_0)}\Phi'(x) = e^{-\beta(x-x_0)}\big(- \beta\Phi(x) + \Phi'(x)\big) \leq 0$$

and $\Psi(x_0) = \Phi(x_0) = \alpha$. It follows that $\Psi(x) \leq \alpha$ for $x \geq x_0$, i.e.

$$\Phi(x) \leq \alpha e^{\beta(x - x_0)}$$

for all $x \in I$ with $x \geq x_0$. This implies (2.7.47), because $\varphi(x) \leq \Phi(x)$, by (2.7.46). □

8. Existence in the large

THEOREM 2.8.1 (Global solutions). *Let* $I = (a_0, b_0)$ *with* $-\infty \le a_0 < b_0 \le +\infty$, $\Omega = I \times \mathbb{R}^n$, *and* $f \in C(\Omega; \mathbb{R}^n)$ *be a continuous function satisfying the local Lipschitz assumption* (2.4.41). *Assume that for any compact set* $K \subset I$ *there exists a constant* $C \ge 0$ *such that*

$$|f(x, y)| \le C(1 + |y|), \quad \text{for all } x \in K \text{ and } y \in \mathbb{R}^n. \tag{2.8.48}$$

Then the Cauchy Problem (2.3.35), *with* $x_0 \in I$ *and* $y_0 \in \mathbb{R}^n$, *has a (unique) global solution defined on* I.

PROOF. Let $y \in C^1(J; \mathbb{R}^n)$ be the maximal solution to the Cauchy Problem (2.3.35), with $J = (a, b) \subset I$. Assume by contradiction that $b < b_0$. By Theorem 2.6.3 it is

$$\lim_{x \uparrow b} |y(x)| = +\infty. \tag{2.8.49}$$

Let $K = [x_0, b]$ and $C > 0$ such that (2.8.48) holds. From

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt, \quad x \in J,$$

we get for $x \in J$ with $x \ge x_0$

$$|y(x)| \le |y_0| + C \int_{x_0}^{x} (1 + |y(t)|)dt \le |y_0| + C(b - x_0) + C \int_{x_0}^{x} |y(t)|\, dt.$$

By the Gronwall's Lemma it follows that

$$|y(x)| \le \big\{|y_0| + C(b - x_0)\big\} e^{C(x - x_0)}, \quad x \in (x_0, b),$$

and therefore (2.8.49) cannot hold. □

CHAPTER 3

Linear systems

1. Introduction

Denote by $M_n(\mathbb{R})$, $n \geq 1$, the vector space of $n \times n$ matrices with real entries. If $A = (a_{ij})_{i,j=1,\dots,n} \in M_n(\mathbb{R})$ and $y \in \mathbb{R}^n$, we denote by Ay the standard matrix-vector product where y is thought of as a column vector, and precisely

$$Ay = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} y_j \\ \vdots \\ \sum_{j=1}^n a_{nj} y_j \end{pmatrix}.$$

The matrix norm of A is

$$\|A\| = \max_{|y|=1} |Ay|.$$

The matrix norm has the following properties:

i) $|Ay| \leq \|A\| |y|$ for all $y \in \mathbb{R}^n$;
ii) $\|A + B\| \leq \|A\| + \|B\|$ for all $A, B \in M_n(\mathbb{R})$;
iii) $\|AB\| \leq \|A\| \cdot \|B\|$ for all $A, B \in M_n(\mathbb{R})$.

Let $I = (a, b) \subset \mathbb{R}$ be an interval. A function $A : I \to M_n(\mathbb{R})$ is continuous if $A(x) = (a_{ij}(x))_{i,j=1,\dots,n}$ for $x \in I$ and $a_{ij} \in C(I)$ for all $i, j = 1, ..., n$.

Let $A : I \to M_n(\mathbb{R})$ be continuous and let $b : I \to \mathbb{R}^n$ be a continuous mapping. A system of differential equations of the form

$$y' = A(x)y + b(x) \tag{3.1.50}$$

is called linear. The function $f : I \times \mathbb{R}^n \to \mathbb{R}^n$

$$f(x, y) = A(x)y + b(x).$$

has the following properties:

1) $f \in C(I \times \mathbb{R}^n; \mathbb{R}^n)$;
2) f has the local Lipschitz property in y;
3) for any compact set $K \subset I$ there is a constant $C > 0$ such that

$$|f(x, y)| \leq C(1 + |y|), \quad \text{for all } x \in K,\ y \in \mathbb{R}^n;$$

In fact, for any compact set $K \subset I$ it is $L = \max_{x \in K} \|A(x)\| < +\infty$ and thus

$$\begin{aligned} |f(x, y_1) - f(x, y_2)| &= |A(x)y_1 - A(x)y_2| = |A(x)(y_1 - y_2)| \\ &\leq \|A(x)\| |y_1 - y_2| \leq L |y_1 - y_2| \end{aligned}$$

for all $x \in K$ and $y_1, y_2 \in \mathbb{R}^n$. This shows 2). Moreover, let $M = \max_{x\in K} |b(x)|$ and $C = \max\{L, M\}$. Then we have

$$|f(x,y)| \leq |A(x)y| + |b(x)| \leq C(|y|+1), \quad x \in K,\, y \in \mathbb{R}^n.$$

By Theorem 2.4.2, the Cauchy problem

$$\begin{cases} y' = A(x)y + b(x) \\ y(x_0) = y_0 \end{cases} \tag{3.1.51}$$

has a unique local solution, for any $x_0 \in I$ and $y_0 \in \mathbb{R}^n$. On the other hand, by Theorem 2.8.1 the maximal solution of the Cauchy Problem (3.1.51) is defined on the whole interval I. In the following, by solution of the differential equation (3.1.50) we mean a maximal solution.

2. Homogeneous equations

A differential equation of the form (3.1.50) with $b = 0$ is called homogeneous.

THEOREM 3.2.1. *Let $A : I \to M_n(\mathbb{R})$ be continuous. The set of solutions of the differential equation*

$$y' = A(x)y, \quad x \in I, \tag{3.2.52}$$

is a real vector space of dimension $n \in \mathbb{N}$.

PROOF. Let $S = \{y \in C^1(I;\mathbb{R}^n) : y \text{ is a solution of } (3.2.52)\}$ be the set of solutions. If $y, z \in S$, then $\alpha y + \beta z \in C^1(I;\mathbb{R}^n)$ is also a solution, for any $\alpha, \beta \in \mathbb{R}$:

$$(\alpha y + \beta z)' = \alpha y' + \beta z' = \alpha A(x)y + \beta A(x)z = A(x)(\alpha y + \beta z), \quad x \in I.$$

Then S is a linear subspace of $C^1(I;\mathbb{R}^n)$.

We show that the dimension of S is n. For some fixed $x_0 \in I$, define the mapping $T : S \to \mathbb{R}^n$

$$T(y) = y(x_0). \tag{3.2.53}$$

T is linear: $T(\alpha y+\beta z) = \alpha y(x_0)+\beta z(x_0) = \alpha T(y)+\beta T(z)$. T is injective, i.e. $T(y) = 0$ implies $y = 0$. In fact, y solves equation (3.2.52) with initial condition $y(x_0) = 0$. The solution to this problem is unique and 0 is a solution. Then it is $y = 0$. Finally, T is surjective because for any $y_0 \in \mathbb{R}^n$ the differential equation (3.2.52) with initial datum $y(x_0) = y_0$ has a solution $y \in C^1(I;\mathbb{R}^n)$. □

PROPOSITION 3.2.2. *Let $S \subset C^1(I;\mathbb{R}^n)$ be the space of solutions to* (3.2.52) *and let $y_1, ..., y_n \in S$. The following are equivalent:*

i) *$y_1, ..., y_n$ are a basis of S;*
ii) *$\det[y_1(x_0), ..., y_n(x_0)] \neq 0$ for all $x_0 \in I$;*
iii) *$\det[y_1(x_0), ..., y_n(x_0)] \neq 0$ for some $x_0 \in I$.*

By $[y_1, ..., y_n]$ we mean the $n \times n$ matrix with columns $y_1, ..., y_n \in \mathbb{R}^n$.

Definition 3.2.3 (Fundamental matrix). If one of the three equivalent conditions of Proposition 3.2.2 holds, then the functions $y_1, ..., y_n$ are called a fundamental system of solutions of the differential equation $y' = Ay$. The matrix $Y = [y_1, ..., y_n]$ is then called a fundamental matrix for the equation.

Proof of Proposition 3.2.2. i)⇒ii) Let $x_0 \in I$, and let $T : S \to \mathbb{R}^n$ be the isomorphism defined in (3.2.53). Then $y_1(x_0) = Ty_1, ..., y_n(x_0) = Ty_n$ form a basis for $\mathbb{R}^n$. This is equivalent with ii).

iii)⇒i) Let $x_0 \in I$ be such that iii) holds and let $T : S \to \mathbb{R}^n$ be the isomorphism (3.2.53) relative to x_0. Then $T^{-1} : \mathbb{R}^n \to S$ is also an isomorphisms. It follows that $y_1 = T^{-1}(y_1(x_0)), ..., y_n = T^{-1}(y_n)(x_0)$ is a basis of S. □

Definition 3.2.4 (Wronski determinant). Let $y_1, ..., y_n \in S$ be solutions to the differential equations (3.2.52). The function $w \in C^1(I; \mathbb{R}^n)$

$$w(x) = \det[y_1(x), ..., y_n(x)], \quad x \in I, \tag{3.2.54}$$

is called Wronski determinant of $y_1, ..., y_n$.

Theorem 3.2.5. *The Wronski determinant w of $y_1, ..., y_n \in S$ solves the differential equation*

$$w' = \mathrm{tr}A(x)\, w, \quad x \in I, \tag{3.2.55}$$

where $\mathrm{tr}A(x) = \sum_{i=1}^{n} a_{ii}(x)$ *is the trace of the matrix* $A(x) = (a_{ij}(x))_{i,j=1,...,n}$.

Proof. If $y_1, ..., y_n$ are linearly dependent then $w(x) = 0$ for all $x \in I$ and equation (3.2.55) trivially holds. Assume that $y_1, ..., y_n$ are linearly independent, i.e. $w(x) \neq 0$ for all $x \in I$. Denote by $Y : I \to M_n(\mathbb{R})$ the fundamental matrix having as columns the solutions $y_1, ..., y_n$. Letting $y_j = (y_{1j}, ..., y_{nj})^T$, $j = 1, ..., n$, we have

$$Y(x) = (y_{ij}(x))_{i,j=1,...,n}, \quad x \in I.$$

We check equation (3.2.55) at the point $x_0 \in I$, i.e. we show that $w'(x_0) = \mathrm{tr}A(x_0)w(x_0)$. To this aim, let $z_j \in C^1(I; \mathbb{R}^n)$ be the solution to the Cauchy problem

$$\begin{cases} z' = A(x)z \\ z(x_0) = e_j, \end{cases} \tag{3.2.56}$$

where $e_j = (0, ..., 0, 1, 0, ..., 0)$ with 1 at the jth position. The functions $z_1, ..., z_n$ are a basis for the space if solutions to the differential equation $z' = Az$. Letting, as above,

$$Z(x) = (z_{ij}(x))_{i,j=1,...,n}, \quad x \in I,$$

there exists an invertible matrix $C \in GL_n(\mathbb{R})$ such that

$$Y(x) = CZ(x), \quad x \in I.$$

We show that the function $v(x) = \det Z(x)$ solves $v'(x_0) = \mathrm{tr}A(x_0)$. In fact, we have

$$v'(x) = \frac{d}{dx} \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}\sigma} \prod_{i=1}^{n} z_{i\sigma(i)}(x) = \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}\sigma} \sum_{j=1}^{n} z'_{j\sigma(j)}(x) \prod_{i \neq j} z_{i\sigma(i)}(x),$$

where

$$\prod_{i\neq j} z_{i\sigma(i)}(x_0) = 0 \quad \text{unless } \sigma = \mathrm{Id},$$

and

$$z'_{jj}(x_0) = (A(x_0)z_j(x_0))_j = \sum_{k=1}^n a_{jk}(x_0)z_{kj}(x_0) = \sum_{k=1}^n a_{jk}(x_0)\delta_{kj}(x_0) = a_{jj}(x_0).$$

Then it is $v'(x_0) = \mathrm{tr}A(x_0)$. Now the general result follows on differentiating the identity

$$w = \det Y = \det(CZ) = \det C \det Z = \det C v.$$

In fact,

$$w'(x_0) = \det C v'(x_0) = \det C\, \mathrm{tr}A(x_0) = \mathrm{tr}A(x_0)w(x_0),$$

because $v(x_0) = 1$. □

3. Inhomogeneous equations

Consider an inhomogeneous linear differential equation of the form

$$y' = A(x)y + b(x), \tag{3.3.57}$$

with $A \in C(I; M_n(\mathbb{R}))$ and $b \in C(I;\mathbb{R}^n)$ for some open interval $I \subset \mathbb{R}$.

Let Y be a fundamental matrix for the homogeneous equation $y' = A(x)y$, i.e. $Y' = AY$ and $\det Y \neq 0$ on I. Then, any solution y to this equation is of the form

$$y(x) = Y(x)c, \quad x \in I, \tag{3.3.58}$$

for some (column) vector $c \in \mathbb{R}^n$. We look for a solution to (3.3.57) of the form (3.3.58) with $c \in C^1(I;\mathbb{R}^n)$. This method is called "variation of constants". In this case,

$$y' = Y'c + Yc' = AYc + Yc' = Ay + Yc'.$$

Plugging this identity into (3.3.57), we get $Yc' = b$. Being Y invertible, by an integration over an interval $[x_0, x]$ we find

$$c(x) = c_0 + \int_{x_0}^x Y(t)^{-1}b(t)dt,$$

for some $c_0 \in \mathbb{R}^n$. Thus we find the solution

$$y(x) = Y(x)\Big(c_0 + \int_{x_0}^x Y(t)^{-1}b(t)dt\Big). \tag{3.3.59}$$

THEOREM 3.3.1. *Let Y be a fundamental matrix for the homogeneous equation $y' = Ay$. For any $c_0 \in \mathbb{R}^n$ the function y in* (3.3.59) *is a solution to* (3.3.57). *Moreover, any solution to* (3.3.57) *is of the form* (3.3.59) *for some $c_0 \in \mathbb{R}^n$.*

PROOF. The first statement is an easy computation. Let y be the function (3.3.59) and let $z \in C^1(I;\mathbb{R}^n)$ be a solution to (3.3.57). Then

$$(z - y)' = z' - y' = Az + b - (Ay + b) = A(z - y).$$

It follows that $z - y = Yc_1$ for some $c_1 \in \mathbb{R}^n$ and the claim follows. □

4. Exponential of a matrix

For a matrix $A \in M_n(\mathbb{C})$ define the exponential matrix $e^A \in M_n(\mathbb{C})$ on letting

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

In order to prove that the series converges, we show that the sequence of matrices $(B_k)_{k\in\mathbb{N}} \subset M_n(\mathbb{C})$

$$B_k = \sum_{h=0}^{k} \frac{A^h}{h!}, \quad k \in \mathbb{N},$$

is a Cauchy sequence in the norm $\|\cdot\|$. In fact, for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $k \geq N$ and for all $p \in \mathbb{N}$ we have

$$\|B_{k+p} - B_k\| = \Big\| \sum_{h=k+1}^{k+p} \frac{A^h}{h!} \Big\| \leq \sum_{h=k+1}^{k+p} \frac{\|A\|^h}{h!} \leq \varepsilon.$$

Notice that the normed space $(M_n(\mathbb{C}), \|\cdot\|)$ is complete.

We list some properties of the exponential matrix.

4.1. Exponential of the sum. If $A, B \in M_n(\mathbb{C})$ and $AB = BA$, then

$$e^{A+B} = e^A e^B. \tag{3.4.60}$$

The proof of this fact is left as an exercise.

4.2. Diagonal matrix. Let $\lambda_1, ..., \lambda_n \in \mathbb{C}$. The exponential matrix of a diagonal matrix $A \in M_n(\mathbb{C})$ of the form

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{is} \quad e^A = \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}.$$

This follows directly from the formula for the exponential.

4.3. Block matrix. Let $A_j \in M_{k_j}(\mathbb{C})$ for $j = 1, ..., p$, with $k_1 + ... + k_p = n$. The exponential matrix of a block matrix $A \in M_n(\mathbb{C})$ of the form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix} \quad \text{is} \quad e^A = \begin{pmatrix} e^{A_1} & & 0 \\ & \ddots & \\ 0 & & e^{A_p} \end{pmatrix}.$$

This also follows directly from the formula for the exponential.

4.4. Fundamental Jordan block. Consider a matrix $A \in M_n(\mathbb{C})$ of the form

$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}, \quad A = \lambda I_n + J, \quad \text{with} \quad J = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix},$$

and $\lambda \in \mathbb{C}$. The matrix A is called fundamental Jordan block of order n relative to $\lambda \in \mathbb{C}$. Later, we shall use the notation $A = J_n(\lambda)$.

We show that for any $t \in \mathbb{R}$ we have

$$e^{tA} = e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2! & \dots & t^{n-1}/(n-1)! \\ & 1 & t & & \vdots \\ & & 1 & \ddots & t^2/2! \\ & & & \ddots & t \\ 0 & & & & 1 \end{pmatrix}. \tag{3.4.61}$$

The matrix $J^k = J \dots J$ k-times, $k = 0, 1, ..., n-1$, has 1 on the $(k+1)$-th left-right downwards diagonal and 0 otherwise. Moreover, it is $J^k = 0$ for $k \geq n$. Then we have

$$\begin{aligned} \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} &= \sum_{k=0}^{+\infty} \frac{t^k}{k!} \sum_{h=0}^{k} \binom{k}{h} \lambda^h J^{k-h} \\ &= \sum_{h=0}^{+\infty} \lambda^h \sum_{k=h}^{h+n-1} \frac{t^k}{k!} \binom{k}{h} J^{k-h} \\ &= \sum_{h=0}^{+\infty} \frac{\lambda^h t^h}{h!} \sum_{p=0}^{n-1} \frac{t^p}{p!} J^p. \end{aligned}$$

4.5. Conjugation and exponentiation. Let $A, B \in M_n(\mathbb{C})$ and $C \in GL_n(\mathbb{C})$ be matrices such that $A = CBC^{-1}$. Then we have

$$e^A = Ce^B C^{-1}.$$

In fact

$$e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!} = \sum_{k=0}^{+\infty} \frac{(CBC^{-1})^k}{k!} = \sum_{k=0}^{+\infty} \frac{CB^k C^{-1}}{k!} = Ce^B C^{-1}.$$

5. Linear systems with constant coefficients

Let $A \in M_n(\mathbb{R})$ be an $n \times n$ matrix and consider the differential equation

$$y' = Ay, \quad x \in \mathbb{R}. \tag{3.5.62}$$

This is a linear, homogeneous system of differential equations with constant coefficients. The solutions are defined on $\mathbb{R}$ and the set of solutions is a real vector space of dimension n. For some $x_0 \in \mathbb{R}$, fix the initial data $y(x_0) = y_0 \in \mathbb{R}^n$. The solution to the differential equation with this initial data is a fixed point of the mapping $T : X \to X$

$$Ty(x) = y_0 + \int_{x_0}^{x} Ay(t)\,dt = y_0 + A\int_{x_0}^{x} y(t)\,dt, \quad x \in \mathbb{R}, \tag{3.5.63}$$

where $X = \{y \in C(\mathbb{R};\mathbb{R}^n) : y(x_0) = y_0\}$. We can interchange integral and A, because A has constant coefficients.

The fixed point is unique and can be obtained as the limit of $T^k y$ for $k \to +\infty$, for any $y \in X$. In particular, we can choose the constant function $y = y_0$. In this case we have

$$Ty(x) = y_0 + (x - x_0)Ay_0,$$

and, in general, we find for any $k \in \mathbb{N}$

$$T^k y(x) = \sum_{h=0}^{k} \frac{(x - x_0)^h}{h!} A^h y_0.$$

This formula can be checked by induction. It holds for $k = 0, 1$, with the convention $A^0 = I_n$, the identity matrix. Assume it holds for k. Then we have

$$\begin{aligned} T^{k+1}y(x) &= T(T^k y)(x) = y_0 + A\int_{x_0}^{x} T^k y(t)\,dt \\ &= y_0 + A\sum_{h=0}^{k} A^h y_0 \int_{x_0}^{x} \frac{(t - x_0)^h}{h!}dt \\ &= y_0 + \sum_{h=0}^{k} \frac{(x - x_0)^{h+1}}{(h+1)!} A^{h+1} y_0 = \sum_{h=0}^{k+1} \frac{(x - x_0)^h}{h!} A^h y_0. \end{aligned}$$

For any compact set $K \subset \mathbb{R}$, the sequence of matrices

$$B_k(x) = \sum_{h=0}^{k} \frac{(x - x_0)^h}{h!} A^h, \quad k \in \mathbb{N},$$

converges uniformly for $x \in K$. From the theory of power series, it follows that the function $\varphi : \mathbb{R} \to M_n(\mathbb{R})$

$$\varphi(x) = e^{(x-x_0)A} = \sum_{h=0}^{+\infty} \frac{(x - x_0)^h}{h!} A^h$$

is of class C^∞, and in fact it is analytic.

Proposition 3.5.1. *Let $A \in M_n(\mathbb{R})$. For any $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^n$, the function $y \in C^\infty(\mathbb{R};\mathbb{R}^n)$*

$$y(x) = e^{(x-x_0)A} y_0$$

is the unique solution to the Cauchy Problem $y' = Ay$ and $y(x_0) = y_0$.

PROOF. The function y is the unique fixed point of the mapping T in (3.5.63). Alternatively, the function y can be differentiated term by term, because the series of the derivatives converges uniformly on compact sets. Then we find

$$y'(x) = \sum_{h=1}^{+\infty} \frac{(x-x_0)^{h-1}}{(h-1)!} A^h y_0 = A \sum_{h=1}^{+\infty} \frac{(x-x_0)^{h-1}}{(h-1)!} A^{h-1} y_0 = Ay(x), \quad x \in \mathbb{R}.$$

Moreover, $y(x_0) = y_0$. □

DEFINITION 3.5.2 (Jordan block). A matrix $A \in M_n(\mathbb{C})$ of the form

$$A = \begin{pmatrix} J_{k_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{k_p}(\lambda) \end{pmatrix},$$

where $\lambda \in \mathbb{C}$, $k_1 + ... + k_p = n$, and $J_{k_1}(\lambda), ..., J_{k_p}(\lambda)$ are fundamental Jordan blocks, is called Jordan block of orders $k_1, ..., k_p$ relative to $\lambda \in \mathbb{C}$. We denote $A = J_{k_1...k_p}(\lambda)$.

The exponential of a Jordan block can be computed using the rules of Section 4.

By known results from Linear Algebra, for any matrix $A \in M_n(\mathbb{R})$ with complex eigenvalues $\lambda_1, ..., \lambda_m \in \mathbb{C}$ there exists a matrix $C \in GL_n(\mathbb{C})$ such that $A = CBC^{-1}$, where B is the Jordan normal form of A, i.e.

$$B = \begin{pmatrix} J_{k_1^1...k_{p_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_1^m...k_{p_m}^m}(\lambda_m) \end{pmatrix}, \tag{3.5.64}$$

with $k_1^1 + ... + k_{p_1}^1 + ... + k_1^m + ... + k_{p_m}^m = n$ and $J_{k_1^1...k_{p_1}^1}(\lambda_1)$, ..., $J_{k_1^m...k_{p_m}^m}(\lambda_m)$ are Jordan blocks relative to the eigenvalues $\lambda_1, ..., \lambda_m$.

PROPOSITION 3.5.3. *Let $A \in M_n(\mathbb{R})$, $A = CBC^{-1}$ where $C \in GL_n(\mathbb{C})$ and B is as in* (3.5.64). *A fundamental system of solutions of the homogeneous linear equation $y' = Ay$ is given by the columns of the (real) matrix*

$$e^{xA} = C \begin{pmatrix} e^{xJ_{k_1^1...k_{p_1}^1}(\lambda_1)} & & 0 \\ & \ddots & \\ 0 & & e^{xJ_{k_1^m...k_{p_m}^m}(\lambda_m)} \end{pmatrix} C^{-1}, \quad x \in \mathbb{R}.$$

PROOF. This follows from Proposition 3.5.1 and by the computation rules of Section 4. □

6. Higher order linear equations

Let $f, a_k \in C(I)$, $k = 0, 1, ..., n-1$, be continuous functions in some interval $I \subset \mathbb{R}$. We transform the linear n-th order differential equation

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + ... + a_1(x)y' + a_0(x)y = f(x), \quad x \in I, \tag{3.6.65}$$

into a linear system. Consider the mapping $T : C^n(I) \to C^1(I;\mathbb{R}^n)$

$$Ty = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}.$$

The vector of functions $z = Ty$ satisfies the system of equations

$$\begin{cases} z_i' = z_{i+1}, \quad i = 1, ..., n-1, \\ z_n' = f(x) - \displaystyle\sum_{k=0}^{n-1} a_k(x) z_{k+1}, \end{cases}$$

which can be written in the following way

$$z' = Az + F, \quad \text{with} \quad A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ -a_0 & -a_1 & & \dots & -a_{n-1} \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f \end{pmatrix}. \tag{3.6.66}$$

Proposition 3.6.1. *Let $y \in C^n(I)$ and $z = Ty \in C^1(I;\mathbb{R}^n)$. Then y solves equation* (3.6.65) *if and only if z solves system* (3.6.66)*. Moreover, the set of solutions $y \in C^n(I)$ of equation* (3.6.65) *with $f = 0$ is a real vector space of dimension n.*

The proof of this proposition is straightforward.

7. Higher order linear equations with constant coefficients

We solve the differential equation (3.6.65) in the homogeneous case $f = 0$ and with constant coefficients $a_0, a_1, ..., a_k \in \mathbb{R}$. Equivalently, we solve the linear system

$$z' = Az, \quad \text{with} \quad A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ -a_0 & -a_1 & & \dots & -a_{n-1} \end{pmatrix} \in M_n(\mathbb{R}). \tag{3.7.67}$$

We establish some algebraic properties of the matrix A. The characteristic polynomial in the variable $\lambda \in \mathbb{C}$ of the matrix A is

$$p(\lambda) = \det(A - \lambda I) = (-1)^n \sum_{k=0}^{n} a_k \lambda^k,$$

with $a_n = 1$. In fact, we can develop the determinant in the last row:

$$
\begin{aligned}
p(\lambda) &= \det \begin{pmatrix} -\lambda & 1 & & & \\ & -\lambda & 1 & & \\ & & -\lambda & \ddots & \\ & & & \ddots & 1 \\ -a_0 & -a_1 & & \dots & -a_{n-1}-\lambda \end{pmatrix} \\
&= -(-1)^{n+1}a_0 - (-1)^{n+2}a_1(-\lambda)^1 - ... - (-1)^{2n}(a_{n-1}+\lambda)(-\lambda)^{n-1} \\
&= (-1)^n\big(a_0 + a_1\lambda + ... + (a_{n-1}+\lambda)\lambda^{n-1}\big).
\end{aligned}
$$

The geometric multiplicity (i.e. the dimension of the eigenspace) of any eigenvalue $\lambda \in \mathbb{C}$ of A is 1 and a corresponding eigenvector is

$$
v_\lambda = \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}. \tag{3.7.68}
$$

Indeed, let $v = (v^1, ..., v^n)^t$ be an eigenvector of A, $Av = \lambda v$. If $v^1 = 0$ then it follows that $v = 0$. We can then assume $v^1 = 1$ and from $v^{i+1} = \lambda v^i$ we deduce that $v = v_\lambda$ as in (3.7.68).

A Jordan chain of vectors $v_0, v_1, ..., v_{r-1}$ relative to the eigenvector v with eigenvalue λ of algebraic multiplicity $r \geq 1$ is defined through the recursive relations $v_0 = v$ and $(A - \lambda)v_{i+1} = v_i$, $i = 0, 1, ..., r-2$. Jordan chains will be used to transform A into its Jordan normal form.

In our case, a Jordan chain relative to the eigenvalue $\lambda \in \mathbb{C}$ of algebraic multiplicity r_λ is given by the vectors

$$
v_{\lambda,i} = \frac{1}{i!} D^i_\lambda v_\lambda, \quad i = 0, 1, ..., r_\lambda - 1,
$$

where D^i_λ is the i-th derivative operator w.r.t. λ. Explicitly, we have

$$
v_{\lambda,0} = \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{n-1} \end{pmatrix}, \quad v_{\lambda,1} = \begin{pmatrix} 0 \\ 1 \\ 2\lambda \\ \vdots \\ (n-1)\lambda^{n-2} \end{pmatrix}, \quad v_{\lambda,2} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ \frac{(n-1)(n-2)}{2}\lambda^{n-3} \end{pmatrix}, \quad \text{etc.} \tag{3.7.69}
$$

We check that $(A - \lambda)v_{\lambda,i+1} = v_{\lambda,i}$ for all $i = 0, 1, ..., r_\lambda - 1$. Let us introduce the following notation:

$$
v_{\lambda,i} = (v^1_i, ..., v^n_i)^t, \quad v^j_i = \frac{1}{i!} D^i \lambda^{j-1}.
$$

Then we have to check that

$$
v^{j+1}_{i+1} - \lambda v^j_{i+1} = v^j_i, \quad j = 1, ..., n-1, \quad -\sum_{k=0}^{n-1} a_k v^{k+1}_{i+1} - \lambda v^n_{i+1} = v^n_i, \quad i + 1 \leq r_\lambda - 1.
$$

The last equation is equivalent with

$$\begin{aligned} 0 &= \sum_{k=0}^{n-1} a_k D^{i+1}\lambda^k + \lambda D^{i+1}\lambda^{n-1} + (i+1)D^i\lambda^{n-1} \\ &= \sum_{k=0}^{n-1} a_k D^{i+1}\lambda^k + D^{i+1}\lambda^n \\ &= \sum_{k=0}^{n} a_k D^{i+1}\lambda^k = (-1)^n D^{i+1}p(\lambda). \end{aligned}$$

The equation $D^{i+1}p(\lambda) = 0$ is satisfied as soon as $i+1 \le r_\lambda - 1$, because it is $p(\lambda) = 0$ with (algebraic) multiplicity r_λ.

Now we determine the Jordan normal form of the matrix A. Let $\lambda_1, ..., \lambda_p \in \mathbb{C}$ be the eigenvalues of A and $v_{\lambda_1}, ..., v_{\lambda_p}$ the corresponding eigenvectors. Denote by r_{λ_j} the algebraic multiplicity of λ_j, for $j = 1, ..., p$. Finally, let $v_{\lambda_j}^k$ with $k = 0, 1, ..., r_{\lambda_j} - 1$ be a Jordan chain relative to $v_{\lambda_j} = v_{\lambda_j}^0$.

Let $C \in GL_n(\mathbb{C})$ be the matrix

$$C = [v_{\lambda_1}^0 \dots v_{\lambda_1}^{r_{\lambda_1}-1} \dots v_{\lambda_p}^0 \dots v_{\lambda_p}^{r_{\lambda_p}-1}].$$

Then A has the Jordan normal form

$$A = C \begin{pmatrix} J_{r_{\lambda_1}}(\lambda_1) & & \\ & \ddots & \\ & & J_{r_{\lambda_p}}(\lambda_p) \end{pmatrix} C^{-1},$$

where $J_{r_{\lambda_1}}(\lambda_1)$,..., $J_{r_{\lambda_p}}(\lambda_p)$ are fundamental Jordan blocks. The exponential of A is then

$$e^{xA} = C \begin{pmatrix} e^{xJ_{r_{\lambda_1}}(\lambda_1)} & & \\ & \ddots & \\ & & e^{xJ_{r_{\lambda_p}}(\lambda_p)} \end{pmatrix} C^{-1},$$

where the exponential of a fundamental Jordan block is computed in (3.4.61).

The column of the matrix $e^{xA}C$ are a fundamental system of complex valued solutions for the system of equations $z' = Az$. The n functions appearing in the first row of the matrix $e^{xA}C$ are thus n linearly independent complex valued solutions of equation (3.6.65) with $f = 0$. Then the following functions are a system of n linearly independent complex valued solution to the equation

$$e^{\lambda_1 x}, xe^{\lambda_1 x}, ..., x^{r_{\lambda_1}-1}e^{\lambda_1 x}, \dots, e^{\lambda_p x}, xe^{\lambda_p x}, ..., x^{r_{\lambda_p}-1}e^{\lambda_p x}. \tag{3.7.70}$$

In order to get real valued solutions notice that $\lambda \in \mathbb{C}$ is an eigenvalue for A if and only if $\bar{\lambda}$ is an eigenvalue, because A has real coefficients. Complex valued solutions are thus coupled, and by linear combinations we obtain real valued solutions.

THEOREM 3.7.1. *Let* $a_k \in \mathbb{R}$, $k = 0, 1, ..., n-1$, *and* $a_n = 1$. *Let* $\mu_1, ..., \mu_q \in \mathbb{R}$ *and* $\lambda_1 = \alpha_1 + i\beta_1, ..., \lambda_p = \alpha_p + i\beta_p, \bar{\lambda}_1, ..., \bar{\lambda}_p \in \mathbb{C} \setminus \mathbb{R}$ *be the real respectively complex* *solutions of the equation*

$$\sum_{k=0}^{n} a_k \lambda^k = 0.$$

Let $r_{\mu_i} \geq 1$ *be the algebraic multiplicity of* μ_i*, and let* $r_{\lambda_j} \geq 1$ *be the algebraic multiplicity of* λ_j *(and so also of* $\bar{\lambda}_j$*). A basis of solutions to the differential equation*

$$y^{(n)} + a_{n-1} y^{(n-1)} + ... + a_1 y' + a_0 y = 0, \quad x \in \mathbb{R}, \tag{3.7.71}$$

is given by the functions

$$\begin{gathered} e^{\mu_1 x}, x e^{\mu_1 x}, \ldots, x^{r_{\mu_1}-1} e^{\mu_1 x} \\ \vdots \\ e^{\mu_q x}, x e^{\mu_q x}, \ldots, x^{r_{\mu_q}-1} e^{\mu_q x} \end{gathered}$$

along with

$$\begin{gathered} e^{\alpha_1 x} \sin(\beta_1 x), x e^{\alpha_1 x} \sin(\beta_1 x), \ldots, x^{r_{\lambda_1}-1} e^{\alpha_1 x} \sin(\beta_1 x) \\ e^{\alpha_1 x} \cos(\beta_1 x), x e^{\alpha_1 x} \cos(\beta_1 x), \ldots, x^{r_{\lambda_1}-1} e^{\alpha_1 x} \cos(\beta_1 x) \\ \vdots \\ e^{\alpha_p x} \sin(\beta_p x), x e^{\alpha_p x} \sin(\beta_p x), \ldots, x^{r_{\lambda_p}-1} e^{\alpha_p x} \sin(\beta_p x) \\ e^{\alpha_p x} \cos(\beta_p x), x e^{\alpha_p x} \cos(\beta_p x), \ldots, x^{r_{\lambda_p}-1} e^{\alpha_p x} \cos(\beta_p x). \end{gathered}$$