Algerian Democratic and Popular Republic
وزارة التعليم العالي والبحث العلمي
Ministry of Higher Education and Scientific Research

University of 20 Août 1955-Skikda
Department of Mathematics

DYNAMICAL SYSTEMS 1

First semester

For first-year Master's degree students

Realized by: Amel Boulfoul

2024/2025

Contents

1 Preliminaries of Ordinary Differential Equations **1**

1.1 General Definition of an Ordinary Differential System 1

1.2 Classification of differential systems 2

1.2.1 Examples 3

1.3 Linear differential system 5

1.3.1 Homogeneous linear systems 5

1.3.2 Nonhomogeneous linear systems 28

1.4 Exercises 39

2 General Theory of ODEs **41**

2.1 ODEs, IVPs, solutions 41

2.1.1 Ordinary differential equation, initial value problem 41

2.1.2 Solutions of an ODE 43

2.2 Existence and uniqueness theorems 46

2.2.1 Successive approximations method 46

2.2.2 Local existence and uniqueness - Proof by fixed point theorem 47

2.2.3 Local existence and uniqueness - Proof by successive approximations method 55

2.2.4 Local existence (non Lipschitz case) 59

2.2.5 Some examples of existence and uniqueness 65

2.3 Continuation of solutions 70

2.3.1 Maximal interval of existence 72

2.3.2 Maximal and global solutions 74

2.4 Exercises 76

3 Stability in Linear and Nonlinear Systems **78**
3.1 Linear stability analysis . . . 78
3.1.1 Equilibria . . . 79
3.1.2 Stability of Equilibria . . . 79
3.1.3 Classification of fixed points in $2D$ . . . 81
3.2 Nonlinear stability analysis . . . 83
3.2.1 Linearization Around Equilibria . . . 85
3.2.2 Stability and Lyapunov's method . . . 87
3.3 First integral . . . 93
3.4 Conservative and dissipative systems . . . 94
3.4.1 Conservative Systems . . . 95
3.4.2 Dissipative Systems . . . 99
3.5 Exercises . . . 100

4 Periodic Solutions and Their Stability **102**
4.1 Limit sets and invariant sets . . . 102
4.2 Periodic solutions and limit cycles . . . 109
4.3 Poincaré maps . . . 113
4.4 Stability of periodic orbits . . . 118
4.5 Bendixon's and Dulac's criteria . . . 118
4.6 The Poincaré-Bendixon theorem . . . 119
4.7 Exercises . . . 123

5 Introduction to Local Bifurcations **126**
5.1 Basic concepts of bifurcation . . . 126
5.2 Bifurcation in one-dimensional systems . . . 127
5.2.1 Saddle-node bifurcation . . . 128
5.2.2 Pitchfork bifurcation . . . 131
5.2.3 Transcritical bifurcation . . . 132
5.3 Bifurcation in two-dimensional systems . . . 134
5.3.1 Saddle-node bifurcation . . . 135
5.3.2 Transcritical and pitchfork bifurcations . . . 136
5.3.3 Hopf bifurcations . . . 140
5.4 Exercises . . . 144

Preface

This course book serves as an introduction to dynamical systems, focusing on ordinary differential equations (ODEs), their stability, periodic solutions, and bifurcations. It is aimed at students in mathematics, particularly those in the first year of their Master's degree.

The book is structured into five main chapters, each addressing important aspects of dynamical systems. The first chapter, **Preliminaries of Ordinary Differential Equations**, begins with a review of differential systems, followed by their classification, and linear differential systems, including homogeneous and nonhomogeneous cases.

The second chapter, **General Theory of ODEs**, introduces the fundamental aspects of ordinary differential equations, initial value problems, and solutions. It includes key existence and uniqueness theorems, different proof methods, and examples. Additionally, the continuation of solutions and maximal intervals of existence are discussed in detail.

The third chapter, **Stability in Linear and Nonlinear Systems**, explores the concept of stability, starting with linear systems and extending to nonlinear systems. It covers important tools such as Lyapunov's method and the analysis of conservative and dissipative systems, which are fundamental for understanding system behavior over time.

In the fourth chapter, **Periodic Solutions and Their Stability**, we delve into the nature of periodic solutions, limit cycles, and their stability. Concepts such as Poincaré maps, Bendixson's and Dulac's criteria, and the Poincaré-Bendixson theorem are explored, offering a deep understanding of the behavior of dynamical systems in the long term.

Finally, the fifth chapter, **Introduction to Local Bifurcations**, introduces the

concept of bifurcation, focusing on one-dimensional and two-dimensional systems. Key bifurcations, including saddle-node, pitchfork, transcritical, and Hopf bifurcations, are explored, laying the foundation for further study and applications in complex systems.

In this book, my objective was to collect the most important definitions, information, and tools from the most significant references such as [[12], [14], [15], [10], [5]], and the references therein, to help students focus on the essential notions of dynamical systems. Throughout the book, a formal and clear approach is taken, with numerous examples and illustrations. The topics covered are foundational for the study of dynamical systems and provide the necessary tools to tackle more advanced subjects in mathematical modeling, control theory, and applied mathematics.

It is my hope that this book serves as a learning resource for students and a reference for those seeking to deepen their understanding of dynamical systems.

Chapter 1

Preliminaries of Ordinary Differential Equations

Differential systems play a central role in various branches of mathematics and their applications, providing a framework to model and analyze dynamic processes in science and engineering. These systems describe the relationships between functions and their derivatives, capturing how quantities evolve over time or interact with one another. The study of differential systems encompasses both linear and nonlinear dynamics, offering insights into stability, control, and long-term behavior. In this chapter, we present a concise review of key concepts and results in the theory of differential systems.

1.1 General Definition of an Ordinary Differential System

An ordinary differential system (ODS) is a set of equations involving one or more functions and their derivatives with respect to a single independent variable, typically time (t). Formally, such a system can be expressed as

$$\frac{dY}{dt} = F(t, Y), \tag{1.1.1}$$

where $Y = [y_1(t), y_2(t), \ldots, y_n(t)]^T$ is a vector of unknown functions, and $F(t, Y) = [f_1(t, Y), f_2(t, Y), \ldots, f_n(t, Y)]^T$ is a vector-valued function that determines the sys-

tem's dynamics.

1.2 Classification of differential systems

Differential systems can be classified based on various criteria, including their order, linearity, and whether they are autonomous. In this subsection, we provide the following primary classifications:

1. Order

- First-order systems: These involve only the first derivatives of the unknown functions.
- Higher-order systems: These include derivatives of order greater than one. Such systems can often be transformed into equivalent first-order systems by introducing additional variables.

2. Linearity

- Linear systems: These are systems where the unknown functions and their derivatives appear linearly. A general linear system can be written as:

$$\frac{dY}{dt} = A(t)Y + B(t), \tag{1.2.2}$$

  where $A(t)$ is a matrix and $B(t)$ is a vector.

- Nonlinear systems: These involve nonlinear terms in the unknown functions or their derivatives. Examples include:

$$\begin{cases} \dfrac{dx}{dt} = x^2 + y, \\ \dfrac{dy}{dt} = \sin(x)y. \end{cases}$$

3. Autonomy

- Autonomous systems: The system does not explicitly depend on the independent variable (e.g., time t):

$$\frac{dY}{dt} = F(Y).$$

- Non-autonomous systems: The system explicitly depends on the independent variable:

$$\frac{dY}{dt} = F(t, Y).$$

4. Homogeneity

- Homogeneous systems: All terms involve the unknown functions or their derivatives, with no independent terms (e.g., $B(t) = 0$ in the linear case (1.2.2)).

- nonhomogeneous systems: The system includes independent terms or external inputs (e.g., $B(t) \neq 0$).

5. Dimensionality

- Single-variable systems: These involve a single unknown function.

- Multi-variable systems: These involve multiple interdependent unknown functions.

6. Special Types

- Coupled systems: Systems where the equations are interconnected and cannot be solved independently, as in many physical or biological models.

- Decoupled systems: Systems where the equations can be solved independently of each other.

1.2.1 Examples

Example 1.2.1. *(First-order Linear System)*

$$\begin{cases} \dfrac{dx}{dt} = -2x + y, \\ \dfrac{dy}{dt} = 3x - 4y. \end{cases}$$

This system describes the linear interaction of two variables x and y.

Example 1.2.2. *(Nonlinear System)*

$$\begin{cases} \dfrac{dx}{dt} = x(1-x) - xy, \\ \dfrac{dy}{dt} = -y + xy. \end{cases}$$

This system could represent a predator-prey model with logistic growth for the prey x.

Example 1.2.3. *(Autonomous System)*

$$\begin{cases} \dfrac{dx}{dt} = x^2, \\ \dfrac{dy}{dt} = y - y^3. \end{cases}$$

The system does not explicitly depend on time.

Example 1.2.4. *(Non-autonomous System)*

$$\begin{cases} \dfrac{dx}{dt} = t - x, \\ \dfrac{dy}{dt} = \cos(t) - y. \end{cases}$$

Here, time t explicitly appears in the equations.

Example 1.2.5. *(Homogeneous system)*

$$\begin{cases} \dfrac{dx}{dt} = -x + y, \\ \dfrac{dy}{dt} = -2x - 3y. \end{cases}$$

This system is homogeneous because all terms involve the variables x and y without any independent forcing terms or constants.

Example 1.2.6. *(nonhomogeneous system)*

$$\begin{cases} \dfrac{dx}{dt} = 3x + 4y + \sin(t), \\ \dfrac{dy}{dt} = -x + 2y + e^t. \end{cases}$$

This system includes independent forcing terms $(sin(t), e^t)$.

Example 1.2.7. *(Coupled System)*

$$\begin{cases} \dfrac{dx}{dt} = x + y, \\ \dfrac{dy}{dt} = -x + y. \end{cases}$$

This system describes two interdependent variables often found in physics or engineering contexts, such as coupled oscillators.

Example 1.2.8. *(Decoupled System)*

$$\begin{cases} \dfrac{dx}{dt} = 2x, \\ \dfrac{dy}{dt} = -y. \end{cases}$$

Each equation can be solved independently of the other.

1.3 Linear differential system

Definition 1.3.1. *A linear system is a system of differential equations of the form*

$$\begin{cases} \dot{x}_1 = a_{11}x_1 + \cdots + a_{1n}x_n + f_1, \\ \dot{x}_2 = a_{21}x_1 + \cdots + a_{2n}x_n + f_2, \\ \vdots\vdots\cdots\vdots\vdots \\ \dot{x}_m = a_{m1}x_1 + \cdots + a_{mn}x_n + f_m, \end{cases} \tag{1.3.3}$$

where $\dot{} = d/dt$. The functions $a_{ij}(t)$ and $f_j(t)$ are defined on some interval $a < t < b$ with $1 \leq i \leq m$ and $1 \leq j \leq n$. The unknowns are the functions $x_1(t), \ldots, x_n(t)$.

1.3.1 Homogeneous linear systems

The system (1.3.3) is called homogeneous if all $f_j = 0$, otherwise, it is called non-homogeneous. Here we consider the system (1.3.3) where $f_j = 0$.

Definition 1.3.2. *(Matrix notation for homogeneous linear systems). A homogeneous system of linear differential equations (1.3.3) with $f_j = 0$ is written as the*

equivalent vector-matrix system $\dot{\mathbf{x}} = A(t)\mathbf{x}$, *where*

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

The fundamental theorem of Picard and Lindelöf applied to the matrix system $\dot{\mathbf{x}} = A(t)\mathbf{x}$ says that a unique solution $\mathbf{x}(t)$ exists for each initial value problem and the solution exists on the common interval of continuity of the entries in $A(t)$.
Two particular results are highlighted here to illustrate the application of Picard theory to linear systems.

Theorem 1.3.1. *(Unique Zero Solution). Let $A(t)$ be an $m \times n$ matrix with continuous components on $a < t < b$. Then the initial value problem*

$$\dot{\mathbf{x}} = A(t)\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{0}$$

has unique solution $\mathbf{x}(t) = \mathbf{0}$ on $a < t < b$.

Theorem 1.3.2. *(Existence-Uniqueness for Constant Linear Systems). Let $A(t) = A$ be an $m \times n$ matrix with constant components and let $\mathbf{x}_0$ be any m-vector. Then the initial value problem*

$$\dot{\mathbf{x}} = A\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0$$

has a unique solution $\mathbf{x}(t)$ defined for all values of t.

Theorem 1.3.3. *(Linear Structure). Let $\dot{\mathbf{x}} = A(t)\mathbf{x}$ have two solutions $\mathbf{x}_1(t), \mathbf{x}_2(t)$. If k_1, k_2 are constants, then $\mathbf{x}(t) = k_1\mathbf{x}_1(t) + k_2\mathbf{x}_2(t)$ is also a solution of $\dot{\mathbf{x}} = A(t)\mathbf{x}$.*

The standard basis $\{\mathbf{w}_k\}_{k=1}^n$:
The Picard-Lindelöf theorem applied to initial conditions $\mathbf{x}(t_0) = \mathbf{x}_0$, with $\mathbf{x}_0$ successively set equal to the columns of the $n \times n$ identity matrix, produces n solutions $\mathbf{w}_1, \ldots, \mathbf{w}_n$ to the equation $\mathbf{x}' = A(t)\mathbf{x}$, all of which exist on the same interval $a < t < b$.
The linear structure theorem implies that for any choice of the constants $c_1, \ldots, c_n$,

the vector linear combination

$$\mathbf{x}(t) = c_1\mathbf{w}_1(t) + c_2\mathbf{w}_2(t) + \cdots + c_n\mathbf{w}_n(t) \tag{1.3.4}$$

is a solution of $\dot{\mathbf{x}} = A(t)\mathbf{x}$.

Conversely, if $c_1, \ldots, c_n$ are taken to be the components of a given vector x_0, then the above linear combination must be the unique solution of the initial value problem with $\mathbf{x}(t_0) = \mathbf{x}_0$. Therefore, all solutions of the equation $\dot{\mathbf{x}} = A(t)\mathbf{x}$ are given by the expression above, where $c_1, \ldots, c_n$ are taken to be arbitrary constants. In summary:

Theorem 1.3.4. *(Basis). The solution set of $\dot{\mathbf{x}} = A(t)\mathbf{x}$ is an n-dimensional subspace of the vector space of all vector-valued functions $\mathbf{x}(t)$. Every solution has a unique basis expansion (1.3.4).*

Definition 1.3.3. *(Augmented matrix). Let $u_1(t), u_2(t), \ldots, u_n(t)$ be vector functions, where $u_i(t) \in \mathbb{R}^n$. The augmented matrix formed by these vectors is defined as*

$$\mathrm{aug}(u_1(t), u_2(t), \ldots, u_n(t)) = \begin{bmatrix} u_{1,1}(t) & u_{2,1}(t) & \cdots & u_{n,1}(t) \\ u_{1,2}(t) & u_{2,2}(t) & \cdots & u_{n,2}(t) \\ \vdots & \vdots & \ddots & \vdots \\ u_{1,n}(t) & u_{2,n}(t) & \cdots & u_{n,n}(t) \end{bmatrix},$$

where each $u_i(t) = \begin{bmatrix} u_{i,1}(t) \\ u_{i,2}(t) \\ \vdots \\ u_{i,n}(t) \end{bmatrix}$ *is a column vector.*

Theorem 1.3.5. *(Abel's formula and the Wronskian). Let $\mathbf{x}_h(t) = c_1\mathbf{u}_1(t) + \cdots + c_n\mathbf{u}_n(t)$ be a candidate general solution to the equation $\dot{\mathbf{x}} = A(t)\mathbf{x}$. In particular, the vector functions $\mathbf{u}_1(t), \ldots, \mathbf{u}_n(t)$ are solutions of $\dot{\mathbf{x}} = A(t)\mathbf{x}$. Define the Wronskian by $w(t) = \det(\mathbf{aug}(\mathbf{u}_1(t), \ldots, \mathbf{u}_n(t)))$. Then Abel's formula holds*

$$w(t) = e^{\int_{t_0}^{t} \mathrm{trace}(A(s))ds} w(t_0).$$

In particular, $w(t)$ is either everywhere nonzero or everywhere zero, accordingly as $w(t_0) \neq 0$ or $w(t_0) = 0$.

Theorem 1.3.6. *(Abel's Wronskian test for independence). The vector solutions* $\mathbf{u}_1, \ldots, \mathbf{u}_n$ *of* $\dot{\mathbf{x}} = A(t)\mathbf{x}$ *are independent if and only if the Wronskian* $w(t)$ *is nonzero for some* $t = t_0$.

Clever use of the point t_0 in Abel's Wronskian test can lead to succinct independence tests. For instance, let

$$\mathbf{u}_1 = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} e^{2t} \\ 2e^{2t} \\ 4e^{2t} \end{pmatrix}.$$

Then $w(t)$ might appear to be complicated, but $w(0)$ is obviously zero because it has two duplicate columns. Therefore, Abel's Wronskian test detects dependence of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.

To illustrate Abel's Wronskian test when it detects independence, consider the column vectors

$$\mathbf{u}_1 = \begin{pmatrix} e^t \\ e^t \\ -e^{-t} + 4e^{2t} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} e^{-t} \\ e^t \\ 2e^{-t} \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ e^{2t} \\ -2e^{2t} \end{pmatrix}.$$

At $t = t_0 = 0$, they become the column vectors

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}.$$

Then $w(0) = \det\left(\mathbf{aug}\left(\mathbf{u}_1(0), \mathbf{u}_2(0), \mathbf{u}_3(0)\right)\right) = 1$ is nonzero, testing independence of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.

Definition 1.3.4. *(Equilibria). An equilibrium point* $\mathbf{x}_0$ *of a linear system* $\mathbf{x}' = A(t)\mathbf{x}$ *is a constant solution,* $\mathbf{x}(t) = \mathbf{x}_0$ *for all* t. *Mostly, this makes sense when* $A(t)$ *is constant, although the definition applies to continuous systems. For a solution* $\mathbf{x}$ *to be constant means* $\mathbf{x}' = \mathbf{0}$, *hence all equilibria are determined from the equation*

$$A(t)\mathbf{x}_0 = \mathbf{0} \quad \text{for all } t.$$

This is a homogeneous system of linear algebraic equations to be solved for $\mathbf{x}_0$. *The solution* $\mathbf{x}_0$ *must not depend on* t*; otherwise, it does not represent an equilibrium. The theory for a constant matrix* $A(t) \equiv A$ *states that either* $\mathbf{x}_0 = \mathbf{0}$ *is the unique solution, or there exist infinitely many solutions for* $\mathbf{x}_0$*, corresponding to the case where the nullity of* A *is positive.*

Matrix Exponential

The problem

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{x}_0$$

has a unique solution, according to the Picard-Lindelöf theorem. Solve this problem n times, where $\mathbf{x}_0$ is chosen as each column of the identity matrix. Denote the resulting n solutions by $\mathbf{w}_1(t), \ldots, \mathbf{w}_n(t)$. Define the matrix exponential by packaging these n solutions into a matrix:

$$e^{At} \equiv \text{aug}\left(\mathbf{w}_1(t), \ldots, \mathbf{w}_n(t)\right).$$

By construction, any solution of $\dot{\mathbf{x}} = A\mathbf{x}$ can be uniquely expressed in terms of the matrix exponential e^{At} using the formula

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0).$$

Properties of Matrix Exponential

Here are various formulas and identities for the matrix exponential e^{At}.

1. **Derivative property**

$$\left(e^{At}\right)' = Ae^{At}, \quad \text{Columns of } e^{At} \text{ satisfy } \dot{\mathbf{x}} = A\mathbf{x}.$$

In fact. Let $\mathbf{x}_0$ denote a column of the identity matrix. Define $\mathbf{x}(t) = e^{At}\mathbf{x}_0$. Then

$$\left(e^{At}\right)'\mathbf{x}_0 = \mathbf{x}'(t) = A\mathbf{x}(t) = Ae^{At}\mathbf{x}_0.$$

Because this identity holds for all columns of the identity matrix, then $\left(e^{At}\right)'$ and Ae^{At} have identical columns, hence we have proved the identity $\left(e^{At}\right)' = Ae^{At}$.

2. **Commutativity with another matrix**

If $AB = BA$, then

$$Be^{At} = e^{At}B.$$

Define $\mathbf{w}_1(t) = e^{At}B\mathbf{w}_0$ and $\mathbf{w}_2(t) = Be^{At}\mathbf{w}_0$. Calculate $\mathbf{w}_1'(t) = A\mathbf{w}_1(t)$ and $\mathbf{w}_2'(t) = BAe^{At}\mathbf{w}_0 = ABe^{At}\mathbf{w}_0 = A\mathbf{w}_2(t)$, due to $BA = AB$. Because $\mathbf{w}_1(0) = \mathbf{w}_2(0) = \mathbf{w}_0$, then the uniqueness theorem implies that $\mathbf{w}_1(t) = \mathbf{w}_2(t)$. Because $\mathbf{w}_0$ is any vector, then $e^{At}B = Be^{At}$.

3. **Addition of exponents**

If $AB = BA$, then

$$e^{At}e^{Bt} = e^{(A+B)t}.$$

Let $\mathbf{x}_0$ be a column of the identity matrix. Define $\mathbf{x}(t) = e^{At}e^{Bt}\mathbf{x}_0$ and $\mathbf{y}(t) = e^{(A+B)t}\mathbf{x}_0$. We must show that $\mathbf{x}(t) = \mathbf{y}(t)$ for all t. Define $\mathbf{u}(t) = e^{Bt}\mathbf{x}_0$. We will apply the result $e^{At}B = Be^{At}$, valid for $BA = AB$. We have

$$\begin{aligned}
\mathbf{x}'(t) &= \left(e^{At}\mathbf{u}(t)\right)' && (1.3.5)\\
&= Ae^{At}\mathbf{u}(t) + e^{At}\mathbf{u}'(t) && (1.3.6)\\
&= A\mathbf{x}(t) + e^{At}B\mathbf{u}(t) && (1.3.7)\\
&= A\mathbf{x}(t) + Be^{At}\mathbf{u}(t) && (1.3.8)\\
&= (A+B)\mathbf{x}(t) && (1.3.9)
\end{aligned}$$

We also know that $\mathbf{y}'(t) = (A+B)\mathbf{y}(t)$ and since $\mathbf{x}(0) = \mathbf{y}(0) = \mathbf{x}_0$, then the uniqueness theorem implies that $\mathbf{x}(t) = \mathbf{y}(t)$ for all t. This completes the proof.

4. **Time addition property**

$$e^{At}e^{As} = e^{A(t+s)}, \quad \text{since } A(t) \text{ and } A(s) \text{ commute.}$$

Let t be a variable and consider s fixed. Define $\mathbf{x}(t) = e^{At}e^{As}\mathbf{x}_0$ and $\mathbf{y}(t) = e^{A(t+s)}\mathbf{x}_0$. Then $\mathbf{x}(0) = \mathbf{y}(0)$ and both satisfy the differential equation $\mathbf{u}'(t) = A\mathbf{u}(t)$. By the uniquenes theorem, $\mathbf{x}(t) = \mathbf{y}(t)$, which implies $e^{At}e^{As} = e^{A(t+s)}$.

Theorem 1.3.7. *(Picard Series Expansion) The matrix exponential can be expressed*

as the infinite series

$$e^{At} = \sum_{n=0}^{\infty} A^n \frac{t^n}{n!}.$$

Special Formulas for e^{At}

(a) $e^{\mathbf{diag}(\lambda_1,\ldots,\lambda_n)t} = \mathbf{diag}\left(e^{\lambda_1 t},\ldots,e^{\lambda_n t}\right)$, where $\lambda_1,\ldots,\lambda_n$ are real or complex constants.

(b) $e^{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} t} = e^{at}\begin{pmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{pmatrix}$, where a, b are real constants.

(c) If $A = \mathrm{diag}\,(B_1,\ldots,B_k)$ and each of $B_1,\ldots,B_k$ is a square matrix, then $e^{At} = \mathbf{diag}\left(e^{B_1 t},\ldots,e^{B_k t}\right)$.

Putzer's spectral formula

The spectral formula of Putzer applies to a system $\mathbf{x}' = A\mathbf{x}$ to find the general solution, using matrices $P_1,\ldots,P_n$ constructed from A and the eigenvalues $\lambda_1,\ldots,\lambda_n$ of A, matrix multiplication, and the solution $\mathbf{r}(t)$ of the first order $n\times n$ initial value problem

$$\mathbf{r}'(t) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_2 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda_3 & \cdots & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \cdots & 1 & \lambda_n \end{pmatrix} \mathbf{r}(t), \quad \mathbf{r}(0) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The system is solved by first-order scalar methods and back-substitution. We will derive the formula separately for the 2×2 case (the one used most often) and the $n\times n$ case.

Putzer's 2×2 spectral formula

The general solution of $\dot{\mathbf{x}} = A\mathbf{x}$ is given by the formula

$$\mathbf{x}(t) = \left(r_1(t)P_1 + r_2(t)P_2\right)\mathbf{x}(0),$$

where r_1, r_2, P_1, P_2 are defined as follows.
The eigenvalues $r = \lambda_1, \lambda_2$ are the two roots of the quadratic equation

$$\det(A - rI) = 0$$

Define 2×2 matrices P_1, P_2 by the formulae $P_1 = I, \quad P_2 = A - \lambda_1 I$.
The functions $r_1(t), r_2(t)$ are defined by the differential system

$$\begin{aligned} \dot{r}_1 &= \lambda_1 r_1, & r_1(0) &= 1, \\ \dot{r}_2 &= \lambda_2 r_2 + r_1, & r_2(0) &= 0. \end{aligned}$$

Proof: The Cayley-Hamilton formula $(A - \lambda_1 I)(A - \lambda_2 I) = \mathbf{0}$ is valid for any 2×2 matrix A and the two roots $r = \lambda_1, \lambda_2$ of the determinant equality $\det(A - rI) = 0$. The Cayley-Hamilton formula is the same as $(A - \lambda_2) P_2 = \mathbf{0}$, which implies the identity $AP_2 = \lambda_2 P_2$. The following computation

$$\begin{aligned} \dot{\mathbf{x}}(t) &= (\dot{r}_1(t) P_1 + \dot{r}_2(t) P_2)\, \mathbf{x}(0) \\ &= (\lambda_1 r_1(t) P_1 + r_1(t) P_2 + \lambda_2 r_2(t) P_2)\, \mathbf{x}(0) \\ &= (r_1(t) A + \lambda_2 r_2(t) P_2)\, \mathbf{x}(0) \\ &= (r_1(t) A + r_2(t) A P_2)\, \mathbf{x}(0) \\ &= A\,(r_1(t) I + r_2(t) P_2)\, \mathbf{x}(0) \\ &= A\mathbf{x}(t), \end{aligned}$$

proves that $\mathbf{x}(t)$ is a solution. Because $\Phi(t) \equiv r_1(t) P_1 + r_2(t) P_2$ satisfies $\Phi(0) = I$, then any possible solution of $\dot{\mathbf{x}} = A\mathbf{x}$ can be represented by the given formula. The proof is complete.

Based on the eigenvalues λ_1 and λ_2 we distinguish the following cases:

Real distinct eigenvalues. Suppose A is 2×2 having real distinct eigenvalues λ_1, λ_2 and $\mathbf{x}(0)$ is real. Then

$$r_1 = e^{\lambda_1 t}, \quad r_2 = \frac{e^{\lambda_1 t} - e^{\lambda_2 T}}{\lambda_1 - \lambda_2}$$

and

$$\mathbf{x}(t) = \left(e^{\lambda_1 t} I + \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2} (A - \lambda_1 I) \right) \mathbf{x}(0).$$

The matrix exponential formula for real distinct eigenvalues:

$$e^{At} = e^{\lambda_1 t} I + \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2} (A - \lambda_1 I).$$

Real equal eigenvalues. Suppose A is 2×2 having real equal eigenvalues $\lambda_1 = \lambda_2$ and $\mathbf{x}(0)$ is real. Then $r_1 = e^{\lambda_1 t}, r_2 = te^{\lambda_1 t}$ and

$$\mathbf{x}(t) = \left(e^{\lambda_1 t} I + te^{\lambda_1 t} (A - \lambda_1 I) \right) \mathbf{x}(0).$$

The matrix exponential formula for real equal eigenvalues:

$$e^{At} = e^{\lambda_1 t} I + te^{\lambda_1 t} (A - \lambda_1 I).$$

Complex Eigenvalues. Suppose A is 2×2 having complex eigenvalues $\lambda_1 = a + bi$ with $b > 0$ and $\lambda_2 = a - bi$. If $\mathbf{x}(0)$ is real, then a real solution is obtained by taking the real part of the spectral formula. This formula is formally identical to the case of real distinct eigenvalues. Then

$$\begin{aligned} \mathrm{Re}(\mathbf{x}(t)) &= \left(\mathrm{Re}\left(r_1(t) \right) I + \mathrm{Re}\left(r_2(t) \left(A - \lambda_1 I \right) \right) \right) \mathbf{x}(0) \\ &= \left(\mathrm{Re}\left(e^{(a+ib)t} \right) I + \mathrm{Re}\left(e^{at} \frac{\sin bt}{b} (A - (a + ib) I) \right) \right) \mathbf{x}(0) \\ &= \left(e^{at} \cos bt I + e^{at} \frac{\sin bt}{b} (A - aI) \right) \right) \mathbf{x}(0). \end{aligned}$$

The matrix exponential formula for complex conjugate eigenvalues:

$$e^{At} = e^{at} \left(\cos bt I + \frac{\sin bt}{b} (A - aI) \right) \right).$$

How to remember Putzer's 2×2 formula. The expressions

$$\begin{aligned} e^{At} &= r_1(t) I + r_2(t) \left(A - \lambda_1 I \right), \\ r_1(t) &= e^{\lambda_1 t}, \quad r_2(t) = \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2}. \end{aligned} \tag{*}$$

are enough to generate all the previous three formulae. The fraction $r_2(t)$ is a

difference quotient with limit $te^{\lambda_1 t}$ as $\lambda_2 \to \lambda_1$, therefore the formula includes the case $\lambda_1 = \lambda_2$ by limiting. If $\lambda_1 = \bar{\lambda}_2 = a + ib$ with $b > 0$, then the fraction r_2 is already real, because it has for $z = e^{\lambda_1 t}$ and $w = \lambda_1$ the form

$$r_2(t) = \frac{z - \bar{z}}{w - \bar{w}} = \frac{\sin bt}{b}.$$

Taking real parts of expression (*) then gives the complex case formula for e^{At}.

Putzer's $n \times n$ spectral formula

The general solution of $\dot{\mathbf{x}} = A\mathbf{x}$ is given by the formula

$$\mathbf{x}(t) = (r_1(t)P_1 + r_2(t)P_2 + \cdots + r_n(t)P_n)\,\mathbf{x}(0).$$

where $r_1, r_2, \ldots, r_n, P_1, P_2, \ldots, P_n$ are defined as follows.
The eigenvalues $r = \lambda_1, \ldots, \lambda_n$ are the roots of the polynomial equation

$$\det(A - rI) = 0$$

Define $n \times n$ matrices $P_1, \ldots, P_n$ by the formulae

$$P_1 = I, \quad P_k = P_{k-1}\,(A - \lambda_{k-1}I)\,, \quad k = 2, \ldots, n.$$

More succinctly, $P_k = \prod_{j=1}^{k-1} (A - \lambda_j I)$. The functions $r_1(t), \ldots, r_n(t)$ are defined by the differential system

$$\begin{aligned}
\dot{r}_1 &= \lambda_1 r_1, & r_1(0) &= 1,\\
\dot{r}_2 &= \lambda_2 r_2 + r_1, & r_2(0) &= 0,\\
&\vdots\\
\dot{r}_n &= \lambda_n r_n + r_{n-1}, & r_n(0) &= 0.
\end{aligned}$$

Proof: The Cayley-Hamilton formula $(A - \lambda_1 I)\cdots(A - \lambda_n I) = \mathbf{0}$ is valid for any $n \times n$ matrix A and the n roots $r = \lambda_1, \ldots, \lambda_n$ of the determinant equality $\det(A - rI) = 0$. Two facts will be used: (1) The Cayley-Hamilton formula implies $AP_n = \lambda_n P_n$; (2) The definition of P_k implies $\lambda_k P_k + P_{k+1} = AP_k$ for $1 \leq k \leq n - 1$.

Compute as follows.

$$\begin{aligned}
\mathbf{x}'(t) &= (r_1'(t)P_1 + \cdots + r_n'(t)P_n)\,\mathbf{x}(0) \\
&= \left(\sum_{k=1}^{n} \lambda_k r_k(t)P_k + \sum_{k=2}^{n} r_{k-1}P_k\right)\mathbf{x}(0), \\
&= \left(\sum_{k=1}^{n-1} \lambda_k r_k(t)P_k + r_n(t)\lambda_n P_n + \sum_{k=1}^{n-1} r_k P_{k+1}\right)\mathbf{x}(0) \\
&= \left(\sum_{k=1}^{n-1} r_k(t)(\lambda_k P_k + P_{k+1}) + r_n(t)\lambda_n P_n\right)\mathbf{x}(0) \\
&= \left(\sum_{k=1}^{n-1} r_k(t)AP_k + r_n(t)AP_n\right)\mathbf{x}(0) \\
&= A\left(\sum_{k=1}^{n} r_k(t)P_k\right)\mathbf{x}(0) \\
&= A\mathbf{x}(t).
\end{aligned}$$

This proves that $\mathbf{x}(t)$ is a solution. Because $\Phi(t) \equiv \sum_{k=1}^{n} r_k(t)P_k$ satisfies $\Phi(0) = I$, then any possible solution of $\dot{\mathbf{x}} = A\mathbf{x}$ can be represented. The proof is complete.

The eigenanalysis method

The general solution $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$ of the linear system $\mathbf{x}' = A\mathbf{x}$,
where $' = dt/dx$, can be fully determined through the eigenanalysis of the matrix A).
A particularly practical scenario arises when the $n \times n$ matrix A possesses n independent eigenvectors associated with its eigenpairs: $(\lambda_1, \mathbf{v}_1)\,, \quad (\lambda_2, \mathbf{v}_2)\,, \quad \ldots, \quad (\lambda_n, \mathbf{v}_n)\,.$
The eigenvalues $\lambda_1, \ldots, \lambda_n$ need not be distinct and may be either real or complex.

The eigenanalysis method when A is a 2×2 matrix

Suppose that A is 2×2 real and has eigenpairs $(\lambda_1, \mathbf{v}_1)$, $(\lambda_2, \mathbf{v}_2)\,$, with $\mathbf{v}_1, \mathbf{v}_2$ independent. The eigenvalues λ_1, λ_2 can be both real. Also, they can be a complex conjugate pair $\lambda_1 = \bar{\lambda}_2 = a + ib$ with $b > 0$.
It will be shown that the general solution of $\mathbf{x}' = A\mathbf{x}$ can be written as

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2.$$

Because we have

$$\begin{aligned}
\mathbf{x}' &= c_1\left(e^{\lambda_1 t}\right)'\mathbf{v}_1 + c_2\left(e^{\lambda_2 t}\right)'\mathbf{v}_2 && \text{Differentiate the formula for } \mathbf{x}.\\
&= c_1 e^{\lambda_1 t}\lambda_1\mathbf{v}_1 + c_2 e^{\lambda_2 t}\lambda_2\mathbf{v}_2 &&\\
&= c_1 e^{\lambda_1 t}A\mathbf{v}_1 + c_2 e^{\lambda_2 t}A\mathbf{v}_2 && \text{Use } \lambda_1\mathbf{v}_1 = A\mathbf{v}_1, \lambda_2\mathbf{v}_2 = A\mathbf{v}_2.\\
&= A\left(c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2\right) && \text{Factor } A \text{ left.}\\
&= A\mathbf{x} && \text{Definition of } \mathbf{x}.
\end{aligned}$$

Let's rewrite the solution $\mathbf{x}$ in the vector-matrix form

$$\mathbf{x}(t) = \operatorname{aug}(\mathbf{v}_1, \mathbf{v}_2)\begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t}\end{pmatrix}\begin{pmatrix} c_1 \\ c_2\end{pmatrix}.$$

Since the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are assumed to be linearly independent, the matrix $\mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2)$ is invertible. Setting $t = 0$ in the previous equation yields

$$\begin{pmatrix} c_1 \\ c_2\end{pmatrix} = \mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2)^{-1}\mathbf{x}(0).$$

Since the constants c_1 and c_2 can be chosen to produce any initial condition $\mathbf{x}(0)$, it follows that $\mathbf{x}(t)$ represents the general solution of the system $\mathbf{x}' = A\mathbf{x}$.
The general solution, expressed as $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$, leads to the following exponential matrix relation

$$e^{At} = \mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2)\begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t}\end{pmatrix}\mathbf{aug}(\mathbf{v}_1, \mathbf{v}_2)^{-1}.$$

This formula becomes particularly useful when the eigenpairs are real.

λ_1 and λ_2 Complex Conjugate Eigenvalues: First, the eigenpair $(\lambda_2, \mathbf{v}_2)$ is never computed or used, since $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ implies $A\overline{\mathbf{v}}_1 = \bar{\lambda}_1\overline{\mathbf{v}}_1$, which leads to the result that $\lambda_2 = \bar{\lambda}_1$ has eigenvector $\mathbf{v}_2 = \overline{\mathbf{v}}_1$.

If A is real, then e^{At} is also real. Taking the real parts of the formula for e^{At} will yield a real formula. Due to the complexity of the complex algebra, we will report the real answer and justify it with minimal use of complex numbers.

Define the symbols a, b, P for the eigenpair $(\lambda_1, \mathbf{v}_1)$ as follows

$$\lambda_1 = a + ib, \quad b > 0, \quad P = \text{aug}\left(\text{Re}(\mathbf{v}_1), \text{Im}(\mathbf{v}_1)\right).$$

Then the solution for e^{At} is

$$e^{At} = e^{at} P \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} P^{-1}. \tag{1.3.10}$$

Justification of formula (1.3.10). The formula is established by showing that the matrix $\Phi(t)$ on the right satisfies $\Phi(0) = I$ and $\Phi'(t) = A\Phi(t)$. Therefore, by definition, $e^{At} = \Phi(t)$. For exposition, let

$$R(t) = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}, \quad \Phi(t) = PR(t)P^{-1}.$$

We verify $\Phi(0) = I$ as follows

$$\Phi(0) = PR(0)P^{-1} = Pe^0 IP^{-1} = I.$$

Next, we expand the eigenpair relation

$$A\left(\text{Re}(\mathbf{v}_1) + i\ \text{Im}(\mathbf{v}_1)\right) = (a + ib)\left(\text{Re}(\mathbf{v}_1) + i\ \text{Im}(\mathbf{v}_1)\right),$$

which, when expressed in real and imaginary parts, shows that

$$AP = P \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Now, we compute $\Phi'(t)$

$$\Phi'(t)\Phi^{-1}(t) = PR'(t)P^{-1}PR^{-1}(t)P^{-1} = PR'(t)R^{-1}(t)P^{-1} = P\left(aI + \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}\right) P^{-1}.$$

Simplifying, we get $\Phi'(t) = A\Phi(t)$. Thus, the proof of $\Phi'(t) = A\Phi(t)$ is complete.

The formula for e^{At} implies that the general solution in this special case is

$$\mathbf{x}(t) = e^{at}\ \mathrm{aug}\left(\mathrm{Re}(\mathbf{v}_1), \mathrm{Im}(\mathbf{v}_1)\right) \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The values c_1, c_2 are related to the initial condition $\mathbf{x}(0)$ by the matrix identity

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathrm{aug}\left(\mathrm{Re}(\mathbf{v}_1), \mathrm{Im}(\mathbf{v}_1)\right)^{-1} \mathbf{x}(0).$$

The eigenanalysis method when A is a 3×3 matrix

Suppose that A is 3×3 real and has eigenpairs $(\lambda_1, \mathbf{v}_1)$, $(\lambda_2, \mathbf{v}_2)$, $(\lambda_3, \mathbf{v}_3)$, with $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ independent. The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ can be all real. Also, there can be one real eigenvalue λ_3 and a complex conjugate pair of eigenvalues $\lambda_1 = \bar{\lambda}_2 = a + ib$ with $b > 0$.
The general solution of $\mathbf{x}' = A\mathbf{x}$ can be written as

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 + c_3 e^{\lambda_3 t}\mathbf{v}_3.$$

The verification of the previous solution is similar to the 2×2 case.
The solution $\mathbf{x}$ is written in vector-matrix form

$$\mathbf{x}(t) = \mathbf{aug}\left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right) \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Because the three eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are assumed independent, then $\mathrm{aug}\left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right)$ is invertible. Setting $t = 0$ in the previous equation gives

$$\begin{pmatrix} c_1 \\ c_2 \\ c_2 \end{pmatrix} = \mathbf{aug}\left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right)^{-1} \mathbf{x}(0).$$

The constants c_1, c_2, c_3 can be chosen to produce any initial condition $\mathbf{x}(0)$, therefore $\mathbf{x}(t)$ is the general solution of the 3×3 system $\mathbf{x}' = A\mathbf{x}$. There is a corresponding

exponential matrix relation

$$e^{At} = \mathbf{aug}\left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right) \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{pmatrix} \mathbf{aug}\left(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right)^{-1}.$$

This formula is normally used when the eigenpairs are real. When there is a complex conjugate pair of eigenvalues $\lambda_1 = \bar{\lambda}_2 = a + ib, b > 0$, then as was shown in the 2×2 case it is possible to extract a real solution $\mathbf{x}$ from the complex formula and report a real form for the exponential matrix

$$e^{At} = P \begin{pmatrix} e^{at}\cos bt & e^{at}\sin bt & 0 \\ -e^{at}\sin bt & e^{at}\cos bt & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{pmatrix} P^{-1},$$

where $P = \operatorname{aug}\left(\operatorname{Re}\left(\mathbf{v}_1\right), \operatorname{Im}\left(\mathbf{v}_1\right), \mathbf{v}_3\right).$

The eigenanalysis method for an $n \times n$ Matrix

The general solution formula and the formula for e^{At} generalize easily from the 2×2 and 3×3 cases to the general case of an $n \times n$ matrix.

Theorem 1.3.8. *(The Eigenanalysis Method). Let the $n \times n$ real matrix A have eigenpairs*

$$\left(\lambda_1, \mathbf{v}_1\right), \left(\lambda_2, \mathbf{v}_2\right), \ldots, \quad \left(\lambda_n, \mathbf{v}_n\right),$$

with n independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then the general solution of the linear system $\mathrm{x}' = A\mathrm{x}$ is given by

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} + \cdots + c_n\mathbf{v}_n e^{\lambda_n t}.$$

The vector-matrix form of the general solution is

$$\mathbf{x}(t) = \operatorname{aug}\left(\mathbf{v}_1, \ldots, \mathbf{v}_n\right) \operatorname{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

This form is real provided all eigenvalues are real. A real form can be made from a complex form by following the example of a 3×3 matrix A.
The plan is to list all complex eigenvalues first, in pairs, $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_p, \bar{\lambda}_p$. Then the real eigenvalues $r_1, \ldots, r_q$ are listed, $2p + q = n$. Define
$P = \text{aug}\left(\text{Re}\left(\mathbf{v}_1\right), \text{Im}\left(\mathbf{v}_1\right), \ldots, \text{Re}\left(\mathbf{v}_{2p-1}\right), \text{Im}\left(\mathbf{v}_{2p-1}\right), \mathbf{v}_{2p+1}, \ldots, \mathbf{v}_n\right),$

$$R_\lambda(t) = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}, \quad \text{where } \lambda + a + ib, \quad b > 0.$$

Then the real vector-matrix form of the general solution is

$$\mathbf{x}(t) = P \text{ diag}\left(R_{\lambda_1}(t), \ldots, R_{\lambda_p}(t), e^{r_1 t}, \ldots, e^{r_q t}\right) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \text{ and}$$

$$e^{At} = P \, \mathbf{diag}\left(R_{\lambda_1}(t), \ldots, R_{\lambda_p}(t), e^{r_1 t}, \ldots, e^{r_q t}\right) P^{-1}.$$

Jordan form and eigenanalysis

Generalized eigenanalysis

The fundamental result in generalized eigenanalysis is **Jordan's theorem**, which states that for any real or complex square matrix A, there exists an invertible matrix P such that:

$$A = PJP^{-1},$$

where J is an upper triangular matrix known as the **Jordan form** of A.
In what follows, we describe the procedure to compute the invertible matrix P, whose columns consist of generalized eigenvectors of A, and the corresponding Jordan form J:

$$J = \begin{pmatrix} \lambda_1 & J_{12} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & J_{n-1\,n} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

The entries $J_{i\,i+1}$ of J along its super-diagonal are either 0 or 1, while the diagonal entries λ_i correspond to the eigenvalues of A. Consequently, the Jordan form J is a

band matrix with nonzero entries restricted to its diagonal and super-diagonal.
The *geometric multiplicity* of an eigenvalue λ, denoted as $g(\lambda)$, is given by the nullity of $A - \lambda I$ i.e. $\dim \ker(A - \lambda I)$. This represents the number of basis vectors in the solution space of $(A - \lambda I)\mathbf{x} = \mathbf{0}$, or equivalently, the number of free variables in the system.
The *algebraic multiplicity* of λ, denoted as $a(\lambda)$, is the largest integer k such that $(r - \lambda)^k$ divides the characteristic polynomial $\det(A - rI)$.

Theorem 1.3.9. *(Algebraic and geometric multiplicity). Let A be a square real or complex matrix. then*

$$1 \leq g(\lambda) \leq a(\lambda). \tag{1.3.11}$$

In addition, there are the following relationships between the Jordan form J and algebraic and geometric multiplicities.

1. *$g(\lambda)$ equals the number of Jordan blocks $B(\lambda, m)$ that appear in J,*

2. *$a(\lambda)$ equals the number of times λ is repeated along the diagonal of J.*

The relation $A = PJP^{-1}$, which is equivalent to $AP = PJ$, can be interpreted in terms of the columns of the matrix P. When J consists of a single Jordan block $B(\lambda, m)$, the columns $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of P satisfy the following system of equations

$$\begin{cases} A\mathbf{v}_1 &= \lambda\mathbf{v}_1, \\ A\mathbf{v}_2 &= \lambda\mathbf{v}_2 + \mathbf{v}_1, \\ &\vdots \\ A\mathbf{v}_m &= \lambda\mathbf{v}_m + \mathbf{v}_{m-1}. \end{cases}$$

These equations demonstrate how each vector $\mathbf{v}_k$ in the chain relates to the eigenvalue λ and its predecessors.

Chains of Generalized Eigenvectors For a given eigenvalue λ of the matrix A, generalized eigenanalysis identifies a Jordan block $B(\lambda, m)$ in J by constructing an m-chain of generalized eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$, which form the columns of P in the relation $A = PJP^{-1}$. The first vector $\mathbf{v}_1$ in the chain is an eigenvector, satisfying

the equation

$$(A - \lambda I)\mathbf{v}_1 = \mathbf{0}.$$

The subsequent vectors $\mathbf{v}_2, \ldots, \mathbf{v}_m$ are generalized eigenvectors, not eigenvectors themselves. They satisfy the following relations

$$(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1, \quad \ldots, \quad (A - \lambda I)\mathbf{v}_m = \mathbf{v}_{m-1}.$$

The term *m-chain* implies that the equation $(A - \lambda I)\mathbf{x} = \mathbf{v}_m$ has no solution. The size of the chain, m, is constrained by the inequality $1 \leq m \leq a(\lambda)$.
The Jordan form J may contain several Jordan blocks for one eigenvalue λ. To illustrate, if J has only one eigenvalue λ and $a(\lambda) = 3$, then J might be constructed as follows:

$$J = \mathbf{diag}(B(\lambda,1), B(\lambda,1), B(\lambda,1)) \quad \text{or} \quad J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

$$J = \mathbf{diag}(B(\lambda,1), B(\lambda,2)) \quad \text{or} \quad J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

$$J = B(\lambda,3) \quad \text{or} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

The three generalized eigenvectors for this example correspond to

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \leftrightarrow \quad \text{Three 1-chains,}$$

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad \leftrightarrow \quad \text{One 1-chain and one 2-chain,}$$

$$J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad \leftrightarrow \quad \text{One 3-chains.}$$

Computing m-chains. Let us focus on an eigenvalue λ of the matrix A. Define $N = A - \lambda I$ and let $p = a(\lambda)$ denote the algebraic multiplicity of λ.
To compute an m-chain, begin with an eigenvector $\mathbf{v}_1$, and then iteratively solve the equation $N\mathbf{v}_{j+1} = \mathbf{v}_j$ using row-reduction (**rref**) methods until no further solutions exist. This process must be repeated for each possible choice of $\mathbf{v}_1$. The search for m-chains concludes when p independent generalized eigenvectors are found.

If A has a unique eigenpair $(\lambda, \mathbf{v}_1)$, the process terminates immediately with an m-chain of length $m = p$. This chain corresponds to a single Jordan block $B(\lambda, m)$, and the generalized eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ form the columns of the matrix P.

In the case where two eigenvectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, form a basis for the eigenspace of A corresponding to λ, the equation $N\mathbf{x} = \mathbf{0}$ will have two free variables. Here, the task is to find two chains, m_1 and m_2, such that $m_1 + m_2 = p$. These chains correspond to two Jordan blocks, $B(\lambda, m_1)$ and $B(\lambda, m_2)$.

To better understand the process, consider the matrix

$$\mathcal{N} = \operatorname{diag}\left(B(0, m_1), B(0, m_2), \ldots, B(0, m_k)\right),$$

where $\mathcal{N}$ is in row-reduced echelon form. The system $\mathcal{N}\mathbf{x} = \mathbf{0}$ has k free variables since $\mathcal{N}$ is already in row-reduced echelon form. This implies that if there exists a k-dimensional basis of eigenvectors of A corresponding to λ, the problem reduces to finding m_i-chains, where $1 \leq i \leq k$, such that $m_1 + \cdots + m_k = p$.

These chains correspond to k Jordan blocks, given by $B(\lambda, m_1), \ldots, B(\lambda, m_k)$.

A common naive approach for computing generalized eigenvectors can be illustrated using the following example

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

The matrix A has a single eigenvalue $\lambda = 1$ and two eigenpairs, $(1, \mathbf{u}_1)$ and $(1, \mathbf{u}_2)$. Starting a chain calculation with $\mathbf{v}_1$ equal to either $\mathbf{u}_1$ or $\mathbf{u}_2$ gives a 1-chain. This naive approach results in only two independent generalized eigenvectors. However, to compute the full set of generalized eigenvectors, the calculation must continue until three independent vectors are found.

To resolve this, we keep the 1-chain, say the one generated by $\mathbf{u}_1$, and start a new chain calculation by setting $\mathbf{v}_1 = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$. We adjust the values of a_1 and a_2 until a 2-chain is obtained. This gives the following augmented matrix

$$\text{aug}(A - \lambda I, \mathbf{v}_1) = \begin{pmatrix} 0 & 1 & 1 & a_1 \\ 0 & 0 & 0 & -a_1 + a_2 \\ 0 & 0 & 0 & a_1 - a_2 \end{pmatrix} \approx \begin{pmatrix} 0 & 1 & 1 & a_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where the condition $a_1 - a_2 = 0$ must hold. By choosing $a_1 = a_2 = 1$, we make $\mathbf{v}_1 = \mathbf{u}_1 + \mathbf{u}_2 \neq \mathbf{0}$, and we solve for $\mathbf{v}_2 = (0, 1, 0)^T$. Thus, $\mathbf{u}_1$ forms a 1-chain, and $\mathbf{v}_1, \mathbf{v}_2$ form a 2-chain. The generalized eigenvectors $\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2$ are independent and form the columns of the matrix P. The Jordan form of A is then

$$J = \mathbf{diag}(B(\lambda, 1), B(\lambda, 2)),$$

where $\lambda = 1$. To justify the relation $A = PJP^{-1}$, we check the condition $AP = PJ$ using the formulas

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Jordan blocks

An $m \times m$ matrix $B(\lambda, m)$ is called a Jordan block provided it is a Jordan form, all m diagonal elements are the same eigenvalue λ and all super-diagonal elements are

one:

$$B(\lambda, m) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \quad (m \times m \text{ matrix }).$$

The Jordan block form of J: Given a square matrix A, a Jordan form J for A is built from Jordan blocks, more precisely,

$$J = \operatorname{diag}\left(B\left(\lambda_1, m_1\right), B\left(\lambda_2, m_2\right), \ldots, B\left(\lambda_k, m_k\right)\right).$$

where $\lambda_1, \ldots, \lambda_k$ are eigenvalues of A and $m_1 + \cdots + m_k = n$. The $k-1$ zeros on the super-diagonal of J result from adjacent Jordan blocks. For example, no super-diagonal zeros means there is just one Jordan block. A complete specification of how to build J from A is done in generalized eigenanalysis.

Theorem 1.3.10. *(Exponential of a Jordan block matrix). If λ is real and*

$$B = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \quad (m \times m \text{ matrix })$$

then

$$e^{Bt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & t \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

The equality also holds if λ is a complex number, in which case both sides of the equation are complex.

The real Jordan form of A

Given a real matrix A, generalized eigenanalysis aims to find a real invertible matrix P and a real upper triangular block matrix R such that

$$A = PRP^{-1}.$$

This decomposition ensures that e^{At} can be expressed in terms of R, which is particularly useful when A is real, leading to a real-valued solution for e^{At}.
If λ is a real eigenvalue of A, then a real Jordan block is a matrix

$$B = \text{diag}(\lambda, \ldots, \lambda) + N, \quad \text{where} \quad N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

If $\lambda = a + ib$ is a complex eigenvalue of A, then the symbols $\lambda, 1$, and 0 are replaced respectively by the following 2×2 real matrices:

$$\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \quad \mathcal{I} = \text{diag}(1, 1), \quad \mathcal{O} = \text{diag}(0, 0).$$

The corresponding $2m \times 2m$ real Jordan block matrix is given by

$$B = \text{diag}(\Lambda, \ldots, \Lambda) + \mathcal{N},$$

where

$$\mathcal{N} = \begin{pmatrix} \mathcal{O} & \mathcal{I} & \mathcal{O} & \cdots & \mathcal{O} & \mathcal{O} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & \mathcal{O} & \mathcal{I} \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & \mathcal{O} & \mathcal{O} \end{pmatrix}.$$

Nilpotent matrices: A matrix N is called nilpotent if there exists a positive integer m such that $N^m = 0$. Similarly, for a block matrix $\mathcal{N}$, if $\mathcal{N}^m = 0$, then $\mathcal{N}$ is also nilpotent. The smallest integer m for which $N^m = 0$) is known as the nilpotency of N.

One important property of a nilpotent matrix N is that its matrix exponential can be expressed as a finite sum:

$$e^{Nt} = I + Nt + \frac{N^2t^2}{2!} + \cdots + \frac{N^{m-1}t^{m-1}}{(m-1)!}.$$

This truncated series holds because higher-order terms vanish beyond $m-1$ due to the nilpotency condition $N^m = 0$.

Computing P and R: Generalized eigenvectors corresponding to a real eigenvalue λ are arranged in the matrix P according to the order prescribed by the corresponding real Jordan block in R. In the case where $\lambda = a + ib$ is a complex eigenvalue with $b > 0$, the real and imaginary parts of each generalized eigenvector are grouped together in P, while the conjugate eigenvalue $\bar{\lambda} = a - ib$ is excluded. The result is a real matrix P and a real upper triangular block matrix R that satisfies the relation

$$nA = PRP^{-1}.$$

Theorem 1.3.11. *(Real block diagonal matrix, eigenvalue $a + ib$). Let* $\Lambda = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, $\mathcal{I} = \mathbf{diag}(1,1)$ *and* $\mathcal{O} = \mathbf{diag}(0,0)$. *Consider a real Jordan block matrix of dimension* $2m \times 2m$ *given by the formula*

$$B = \begin{pmatrix} \Lambda & \mathcal{I} & \mathcal{O} & \cdots & \mathcal{O} & \mathcal{O} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & \Lambda & \mathcal{I} \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & \mathcal{O} & \Lambda \end{pmatrix}.$$

If $R = \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}$, *then*

$$e^{Bt} = e^{at} \begin{pmatrix} R & tR & \frac{t^2}{2}R & \cdots & \frac{t^{m-2}}{(m-2)!}R & \frac{t^{m-1}}{(m-1)!}R \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & R & tR \\ \mathcal{O} & \mathcal{O} & \mathcal{O} & \cdots & \mathcal{O} & R \end{pmatrix}.$$

When solving $\mathbf{x}' = A\mathbf{x}$, the solution $\mathbf{x}(t) = e^{At}\mathbf{x}(0)$ must be real if A is real. The real solution can be expressed as $\mathbf{x}(t) = P\mathbf{y}(t)$ where $\mathbf{y}'(t) = R\mathbf{y}(t)$ and R is a real Jordan form of A, containing real Jordan blocks B down its diagonal. Theorems above provide explicit formulas for e^{Bt}, hence the resulting formula

$$\mathbf{x}(t) = Pe^{Rt}P^{-1}\mathbf{x}(0),$$

contains only real numbers, real exponentials, plus sine and cosine terms, which are possibly multiplied by polynomials in t.

1.3.2 Nonhomogeneous linear systems

Definition 1.3.5. *(Matrix notation for nonhomogeneous linear systems). A nonhomogeneous system of linear equations (1.3.3) is written as the equivalent vector-matrix system*

$$\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t),$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

The following theorem establishes the uniqueness of solutions for a class of nonhomogeneous linear systems and highlights the impossibility of solution crossings within the given interval.

Theorem 1.3.12. *(Uniqueness and Solution Crossings). Let $A(t)$ be an $m \times n$ matrix with continuous components on $a < t < b$ and assume that $\mathbf{f}(t)$ is also continuous on $a < t < b$. If $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are solutions of $\dot{\mathbf{u}} = A(t)\mathbf{u} + \mathbf{f}(t)$ on $a < t < b$ and $\mathbf{x}(t_0) = \mathbf{y}(t_0)$ for some $t_0, a < t_0 < b$, then $\mathbf{x}(t) = \mathbf{y}(t)$ for $a < t < b$.*

Remark 1.3.1. *(Superposition). Linear homogeneous systems exhibit a fundamental linear structure, and the solutions of nonhomogeneous systems satisfy the principle of superposition.*

Theorem 1.3.13. *(Superposition Principle). Let $\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$ have a particular solution $\mathbf{x}_p(t)$. If $\mathbf{x}(t)$ is any solution of $\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$, then $\mathbf{x}(t)$ can be*

decomposed as homogeneous plus particular solutions, i.e.

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_p(t).$$

The term $\mathbf{x}_h(t)$ *is a certain solution of the homogeneous differential equation* $\dot{\mathbf{x}} = A(t)\mathbf{x}$*, which means arbitrary constants* $c_1, c_2, \ldots$ *have been assigned certain values. The particular solution* $\mathbf{x}_p(t)$ *can be selected to be free of any unresolved or arbitrary constants.*

Theorem 1.3.14. *(Difference of Solutions). Let* $\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$ *have two solutions* $\mathbf{x} = \mathbf{u}(t)$ *and* $\mathbf{x} = \mathbf{v}(t)$*. Define* $\mathbf{y}(t) = \mathbf{u}(t) - \mathbf{v}(t)$*. Then* $\mathbf{y}(t)$ *satisfies the homogeneous equation* $\mathbf{y}' = A(t)\mathbf{y}$*.*

General solution. We explain a general solution by the following example
If a formula $x = c_1e^t + c_2e^{2t}$ is called a general solution, then it means that all possible solutions of the differential equation are expressed by this formula.
In particular, it means that a given solution can be represented by the formula, by specializing values for the constants c_1, c_2. We expect the number of arbitrary constants to be the least possible number.
The general solution of $\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$ is an expression involving arbitrary constants $c_1, c_2, \ldots$ and certain functions. The expression represents all solutions of the differential equation is given in the following sense:

1. Every assignment of constants produces a solution of the differential equation.
2. Every possible solution is uniquely obtained from the expression by specializing the constants.

Due to the superposition principle, the constants in the general solution are identified as multipliers against solutions of the homogeneous differential equation.

Theorem 1.3.15. *(General Solution) Let* $A(t)$ *be an* $n \times n$ *matrix and* $\mathbf{f}(t)$ *an* $n \times 1$ *vector, both continuous on an interval* $a < t < b$*. The general solution of the nonhomogeneous linear system*

$$\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$$

is given by

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_p(t).$$

Here, $\mathbf{x}_h(t)$ represents the general solution of the associated homogeneous system $\dot{\mathbf{y}} = A(t)\mathbf{y}$, containing n arbitrary constants $c_1, \ldots, c_n$. The term $\mathbf{x}_p(t)$ is a particular solution of the nonhomogeneous system, which contains no arbitrary constants.

Recognition of homogenous solutions terms An expression $\mathbf{x}$ for the general solution of a nonhomogeneous equation $\dot{\mathbf{x}} = A(t)\mathbf{x}+\mathbf{f}(t)$ involves arbitrary constants $c_1, \ldots, c_n$. It is possible to isolate both terms $\mathbf{x}_h$ and $\mathbf{x}_p$ by a simple procedure. **To find $\mathbf{x}_p$**, set to zero all arbitrary constants $c_1, c_2, \ldots$; the resulting expression is free of unresolved and arbitrary constants.
To find $\mathbf{x}_h$, we find first the vector solutions $\mathbf{y} = \mathbf{u}_k(t)$ of $\dot{\mathbf{y}} = A(t)\mathbf{y}$, which are multiplied by constants c_k. Then the general solution $\mathbf{x}_h$ of the homogeneous equation $\dot{\mathbf{y}} = A(t)\mathbf{y}$ is given by

$$\mathbf{x}_h(t) = c_1\mathbf{u}_1(t) + c_2\mathbf{u}_2(t) + \cdots + c_n\mathbf{u}_n(t).$$

Use partial derivatives on expression $\mathbf{x}$ to find the column vectors $\mathbf{u}_k(t) = \frac{\partial}{\partial c_k}\mathbf{x}$. This technique isolates the vector components of the homogeneous solution from any form of the general solution, including scalar formulas for the components of $\mathbf{x}$. In any case, the general solution $\mathbf{x}$ of the linear system $\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t)$ is represented by the expression

$$\mathbf{x} = c_1\mathbf{u}_1(t) + c_2\mathbf{u}_2(t) + \cdots + c_n\mathbf{u}_n(t) + \mathbf{x}_p(t).$$

In this expression, each assignment of the constants $c_1, \ldots, c_n$ produces a solution of the nonhomogeneous system, and conversely, each possible solution of the nonhomogeneous system is obtained by a unique specialization of the constants $c_1, \ldots, c_n$.
To illustrate the ideas, consider a 3×3 linear system $\mathbf{x}' = A(t)\mathbf{x}+\mathbf{f}(t)$ with general solution

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

given in scalar form by the expressions

$$\begin{aligned} x_1 &= c_1 e^t + c_2 e^{-t} + t, \\ x_2 &= (c_1 + c_2)\, e^t + c_3 e^{2t}, \\ x_3 &= (2c_2 - c_1)\, e^{-t} + (4c_1 - 2c_3)\, e^{2t} + 2t. \end{aligned}$$

To find the vector form of the general solution, we take partial derivatives $\mathbf{u}_k = \dfrac{\partial \mathbf{x}}{\partial c_k}$ with respect to the variable names c_1, c_2, c_3 :

$$\mathbf{u}_1 = \begin{pmatrix} e^t \\ e^t \\ -e^{-t} + 4e^{2t} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} e^{-t} \\ e^t \\ 2e^{-t} \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ e^{2t} \\ -2e^{2t} \end{pmatrix}.$$

To find $\mathbf{x}_p(t)$, set $c_1 = c_2 = c_3 = 0$:

$$\mathbf{x}_p(t) = \begin{pmatrix} t \\ 0 \\ 2t \end{pmatrix}.$$

Finally,

$$\mathbf{x} = c_1\mathbf{u}_1(t) + c_2\mathbf{u}_2(t) + c_3\mathbf{u}_3(t) + \mathbf{x}_p(t). \text{So}$$

$$\mathbf{x}(t) = c_1 \begin{pmatrix} e^t \\ e^t \\ -e^{-t} + 4e^{2t} \end{pmatrix} + c_2 \begin{pmatrix} e^{-t} \\ e^t \\ 2e^{-t} \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ e^{2t} \\ -2e^{2t} \end{pmatrix} + \begin{pmatrix} t \\ 0 \\ 2t \end{pmatrix}.$$

The expression $\mathbf{x} = c_1\mathbf{u}_1(t) + c_2\mathbf{u}_2(t) + c_3\mathbf{u}_3(t) + \mathbf{x}_p(t)$ satisfies required elements (i) and (ii) in the definition of general solution. We will develop now a way to routinely test the uniqueness requirement in (ii).

Independence: Constants $c_1, \ldots, c_n$ in the general solution x = $\mathbf{x}_h + \mathbf{x}_p$ appear exactly in the expression $\mathbf{x}_h$, which has the form $\mathbf{x}_h = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$.
A solution $\mathbf{x}$ uniquely determines the constants. In particular, the zero solution of the homogeneous equation is uniquely represented, which can be stated this way:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n = \mathbf{0} \quad \text{implies} \quad c_1 = c_2 = \cdots = c_n = 0.$$

This statement is equivalent to the statement that the vector-valued functions $\mathbf{u}_1(t), \ldots, \mathbf{u}_n(t)$ are **linearly independent**.
It is possible to write down a candidate general solution to some 3×3 linear system $\dot{\mathbf{x}} = A\mathbf{x}$ via equations like

$$\begin{aligned} x_1 &= c_1 e^t + c_2 e^t + c_3 e^{2t}, \\ x_2 &= c_1 e^t + c_2 e^t + 2c_3 e^{2t}, \\ x_3 &= c_1 e^t + c_2 e^t + 4c_3 e^{2t}. \end{aligned}$$

This example was constructed to contain a classic mistake, for purposes of illustration.

How can we detect a mistake, given only that this expression is supposed to represent the general solution? First of all, we can test that $\mathbf{u}_1 = \partial\mathbf{x}/\partial c_1, \mathbf{u}_2 = \partial\mathbf{x}/\partial c_2, \mathbf{u}_3 = \partial\mathbf{x}/\partial c_3$ are indeed solutions. But to ensure the unique representation requirement, the vector functions $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ must be linearly independent. We compute

$$\mathbf{u}_1 = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} e^{2t} \\ 2e^{2t} \\ 4e^{2t} \end{pmatrix}$$

Therefore, $\mathbf{u}_1 = \mathbf{u}_2$, which implies that the functions $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ fail to be independent
.

Initial value problems and the Reduced row echelon form method An initial value problem is the problem of solving for $\mathbf{x}$, given

$$\dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{f}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0.$$

If a general solution is known, $\mathbf{x} = c_1\mathbf{u}_1(t) + \cdots + c_n\mathbf{u}_n(t) + \mathbf{x}_p(t)$,
then the problem of finding $\mathbf{x}$ reduces to finding $c_1, \ldots, c_n$ in the relation

$$c_1\mathbf{u}_1(t_0) + \cdots + c_n\mathbf{u}_n(t_0) + \mathbf{x}_p(t_0) = \mathbf{x}_0.$$

This is a matrix equation for the unknown constants $c_1, \ldots, c_n$ of the form $B\mathbf{c} = \mathbf{d}$, where

$$B = \operatorname{aug}\left(\mathbf{u}_1\left(t_0\right), \ldots, \mathbf{u}_n\left(t_0\right)\right), \quad \mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}, \quad \mathbf{d} = \mathbf{x}_0 - \mathbf{x}_p\left(t_0\right).$$

The **reduced row echelon form** (rref) method applies to find $\mathbf{c}$. This method is used to perform swap, combination, and multiply operations to $C = \mathbf{aug}(B, \mathbf{d})$ until $\mathbf{rref}(C) = \operatorname{aug}(I, \mathbf{c})$.
To illustrate the method, consider the general solution

$$\begin{aligned} x_1 &= c_1 e^t + c_2 e^{-t} + t, \\ x_2 &= \left(c_1 + c_2\right) e^t + c_3 e^{2t}, \\ x_3 &= \left(2c_2 - c_1\right) e^{-t} + \left(4c_1 - 2c_3\right) e^{2t} + 2t. \end{aligned}$$

We shall solve for c_1, c_2, c_3, given the initial condition $x_1(0) = 1, x_2(0) = 0, x_3(0) = -1$. The above relations evaluated at $t = 0$ give the system

$$\begin{aligned} 1 &= c_1 e^0 + c_2 e^0 + 0 \\ 0 &= \left(c_1 + c_2\right) e^0 + c_3 e^0 \\ -1 &= \left(2c_2 - c_1\right) e^0 + \left(4c_1 - 2c_3\right) e^0 + 0. \end{aligned}$$

In standard scalar form, this is the 3×3 linear system

$$\begin{aligned} c_1 + c_2 &= 1, \\ c_1 + c_2 + c_3 &= 0, \\ 3c_1 + 2c_2 - 2c_3 &= -1. \end{aligned}$$

The augmented matrix C, to be reduced to rref form, is given by

$$C = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 2 & -2 & -1 \end{pmatrix}.$$

After the rref process is completed, we obtain

$$\text{rref}(C) = \begin{pmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

From this display, we read off the answer $c_1 = -5, c_2 = 6, c_3 = -1$ and report the final answer

$$\begin{aligned} x_1 &= -5e^t + 6e^{-t} + t \\ x_2 &= e^t - e^{2t} \\ x_3 &= 17e^{-t} - 18e^{2t} + 2t. \end{aligned}$$

Variation of Parameters

The method of variation of parameters is a general technique used to solve a linear nonhomogeneous system of the form $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$.
Historically, this method originated as a trial solution approach, where the nonhomogeneous system is solved by assuming a particular form for the solution, known as an ansatz. The ansatz takes the form

$$\mathbf{x}(t) = e^{At}\mathbf{c}(t),$$

where e^{At} is the solution to the corresponding homogeneous system, and $\mathbf{c}(t)$ is an unknown function to be determined through substitution and further analysis. This method effectively finds the particular solution of the nonhomogeneous system by varying the parameters of the homogeneous solution.
Conceptually, the method is derived by allowing $\mathbf{x}_0$ to vary in the general solution $\mathbf{x}(t) = e^{At}\mathbf{x}_0$ of the homogeneous system $\mathbf{x}' = A\mathbf{x}$, which gave rise to the terms **variation of parameters** and **variation of constants** .

Theorem 1.3.16. *(Variation of parameters for systems). Let A be a constant $n \times n$ matrix and $\mathbf{F}(t)$ a continuous function near $t = t_0$. The unique solution $\mathbf{x}(t)$ of the matrix initial value problem*

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{F}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0$$

is given by the variation of parameters formula (Duhamel's formula)

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 + e^{At}\int_{t_0}^{t} e^{-rA}\mathbf{F}(r)dr. \tag{1.3.12}$$

Proof. Define

$$\mathbf{u}(t) = \mathbf{x}_0 + \int_{t_0}^{t} e^{-rA}\mathbf{F}(r)\,dr.$$

To establish that (1.3.12) holds, i.e. $\mathbf{x}(t) = e^{At}\mathbf{u}(t)$, we verify that $\mathbf{x}(t)$ satisfies the given differential equation.

First, the function $\mathbf{u}(t)$ is differentiable, with a continuous derivative given by

$$\mathbf{u}'(t) = e^{-tA}\mathbf{F}(t),$$

which follows from the fundamental theorem of calculus applied componentwise. Applying the product rule, we obtain

$$\begin{aligned}\mathbf{x}'(t) &= \left(e^{At}\right)'\mathbf{u}(t) + e^{At}\mathbf{u}'(t)\\ &= Ae^{At}\mathbf{u}(t) + e^{At}e^{-At}\mathbf{F}(t)\\ &= A\mathbf{x}(t) + \mathbf{F}(t).\end{aligned}$$

Thus, $\mathbf{x}(t)$ satisfies the differential equation $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$. Furthermore, since $\mathbf{u}(t_0) = \mathbf{x}_0$, it follows that $\mathbf{x}(t_0) = \mathbf{x}_0$, ensuring that the initial condition is satisfied. This completes the proof. □

Undetermined coefficients

The method of undetermined coefficients, a trial solution approach, can be applied to vector-matrix systems of the form $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$, when the components of $\mathbf{F}(t)$ consist of sums of terms of the form

$$(\text{polynomial in } t)e^{at}(\cos(bt) \text{ or } \sin(bt)).$$

Such terms are referred to as atoms.

A practical approach is to express $\mathbf{F}(t)$ in terms of the columns $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of the

$n \times n$ identity matrix I, writing

$$\mathbf{F}(t) = \sum_{j=1}^{n} F_j(t)\mathbf{e}_j.$$

This allows the solution to be decomposed as

$$\mathbf{x}(t) = \sum_{j=1}^{n} \mathbf{x}_j(t),$$

where each $\mathbf{x}_j(t)$ satisfies the simpler equation

$$\mathbf{x}'(t) = A\mathbf{x}(t) + f(t)\mathbf{c}, \quad f = F_j, \quad \mathbf{c} = \mathbf{e}_j.$$

To construct an initial trial solution for $\mathbf{x}'(t) = A\mathbf{x}(t) + f(t)\mathbf{c}$, we follow the initial **trial solution rule** (assuming that $f(t)$ is a sum of atoms):

- Identify the independent functions whose linear combinations produce all derivatives of $f(t)$.
- Form an initial trial solution as a linear combination of these functions, with undetermined vector coefficients $\{\mathbf{c}_j\}$.

In the well-known scalar case, the trial solution must be modified if it overlaps with the general solution of the homogeneous equation. When $f(t)$ is a polynomial, this issue can be avoided by assuming that A is invertible. This assumption ensures that $r = 0$ is not a root of $\det(A - rI) = 0$, thereby preventing polynomial terms from appearing in the homogeneous solution.
Finally, substituting the initial vector trial solution into the differential equation allows us to determine the undetermined coefficients $\{\mathbf{c}_j\}$, thereby obtaining a particular solution.

Theorem 1.3.17. *(Polynomial solutions). Let*

$$f(t) = \sum_{j=0}^{k} p_j \frac{t^j}{j!}$$

be a polynomial of degree k, and assume that A is an $n \times n$ constant invertible matrix. Then the differential equation $\mathbf{u}' = A\mathbf{u} + f(t)\mathbf{c}$, admits a polynomial solution of degree

k of the form

$$\mathbf{u}(t) = \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!},$$

where the vector coefficients $\{\mathbf{c}_j\}$ are determined by the recurrence relation

$$\mathbf{c}_j = -\sum_{i=j}^{k} p_i A^{j-i-1}\mathbf{c}, \quad 0 \le j \le k.$$

Proof. Substituting $\mathbf{u}(t) = \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!}$ into the differential equation, we obtain

$$\sum_{j=0}^{k-1} \mathbf{c}_{j+1} \frac{t^j}{j!} = A \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!} + \sum_{j=0}^{k} p_j \frac{t^j}{j!}\mathbf{c}.$$

To satisfy this equation, the terms on the right-hand side for $j = k$ must sum to zero, while for other values of j, the coefficients of $\frac{t^j}{j!}$ on both sides must match, leading to the relations

$$A\mathbf{c}_k + p_k\mathbf{c} = \mathbf{0}, \quad \mathbf{c}_{j+1} = A\mathbf{c}_j + p_j\mathbf{c}.$$

Solving these relations recursively yields the formulas

$$\begin{aligned}
\mathbf{c}_k &= -p_k A^{-1}\mathbf{c}, \\
\mathbf{c}_{k-1} &= -\left(p_{k-1}A^{-1} + p_k A^{-2}\right)\mathbf{c}, \\
&\vdots \\
\mathbf{c}_0 &= -\left(p_0 A^{-1} + \cdots + p_k A^{-k-1}\right)\mathbf{c}.
\end{aligned}$$

These relations can be summarized as

$$\mathbf{c}_j = -\sum_{i=j}^{k} p_i A^{j-i-1}\mathbf{c}, \quad 0 \le j \le k.$$

Since substituting $\mathbf{u}(t) = \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!}$ into the differential equation yields matching left-hand and right-hand sides, the proof is complete. □

Theorem 1.3.18. *(Polynomial × exponential solutions). Let*

$$g(t) = \sum_{j=0}^{k} p_j \frac{t^j}{j!}$$

be a polynomial of degree k, and assume that A is an $n \times n$ constant matrix such that $B = A - aI$ is invertible. Then the differential equation

$$\mathbf{u}' = A\mathbf{u} + e^{at} g(t)\mathbf{c}$$

admits a polynomial-exponential solution of the form

$$\mathbf{u}(t) = e^{at} \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!},$$

where the vector coefficients $\{\mathbf{c}_j\}$ are determined by the recurrence relation

$$\mathbf{c}_j = -\sum_{i=j}^{k} p_i B^{j-i-1}\mathbf{c}, \quad 0 \le j \le k.$$

Proof. Let $\mathbf{u}(t) = e^{at}\mathbf{v}(t)$. Then, differentiating both sides, we obtain

$$\mathbf{u}' = ae^{at}\mathbf{v}(t) + e^{at}\mathbf{v}'(t) = e^{at}\left(\mathbf{v}'(t) + a\mathbf{v}(t)\right).$$

Substituting into the given differential equation

$$\mathbf{u}' = A\mathbf{u} + e^{at} g(t)\mathbf{c},$$

we obtain

$$e^{at}\left(\mathbf{v}' + a\mathbf{v}\right) = Ae^{at}\mathbf{v} + e^{at} g(t)\mathbf{c}.$$

Canceling e^{at} from both sides, this simplifies to $\mathbf{v}' = (A - aI)\mathbf{v} + g(t)\mathbf{c}$.
Setting $B = A - aI$, we recognize that this equation is of the form

$$\mathbf{v}' = B\mathbf{v} + g(t)\mathbf{c}.$$

By Theorem (1.3.17), the solution takes the form

$$\mathbf{v}(t) = \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!},$$

where the vector coefficients $\mathbf{c}_j$ satisfy

$$\mathbf{c}_j = -\sum_{i=j}^{k} p_i B^{j-i-1}\mathbf{c}, \quad 0 \leq j \leq k.$$

Thus, substituting $\mathbf{v}(t)$ back into $\mathbf{u}(t) = e^{at}\mathbf{v}(t)$, we obtain

$$\mathbf{u}(t) = e^{at} \sum_{j=0}^{k} \mathbf{c}_j \frac{t^j}{j!}.$$

Since the coefficients $\mathbf{c}_j$ satisfy the required formula

$$\mathbf{c}_j = -\sum_{i=j}^{k} p_i B^{j-i-1}\mathbf{c},$$

the proof is complete. □

1.4 Exercises

Exercice 1.4.1. *Explain the difference between an ordinary differential equation (ODE) and a partial differential equation (PDE). Give an example of each.*

Exercice 1.4.2. *Identify whether the following equations are ODEs or not. If it is an ODE, determine its order.*

1. $y' + y = x^2$.
2. $\dfrac{\partial u}{\partial x} + \dfrac{\partial u}{\partial y} = 0$
3. $y''' - xy' + y = 0$.

Exercice 1.4.3. *Determine whether the following ODEs are linear or nonlinear. If nonlinear, justify your answer.*

1. $y'' + y' + y = 0$.

2. $y' = y^2 + x$.

3. $y'' + \sin(y) = x^3$.

Exercice 1.4.4. *Identify whether the following ODEs are autonomous or non-autonomous.*

1. $y' = y^2 - 3y$.

2. $y' = x + y$.

Exercice 1.4.5. *Determine whether the following ODEs are homogeneous or non-homogeneous.*

1. $y'' - 3y' + 2y = 0$.

2. $y'' + 4y = \cos x$.

Exercice 1.4.6. *Verify that $y = e^{2x}$ is a solution to the equation*

$$y'' - 4y' + 4y = 0.$$

Exercice 1.4.7. *Solve the first-order linear differential equation*

$$y' - 2y = e^x.$$

Chapter 2

General Theory of ODEs

We begin with the general theory of ordinary differential equations (ODEs). Section 2.1 introduces ODEs, initial value problems (IVPs), and their solutions. In Section 2.2, we discuss the existence and uniqueness of solutions to IVPs. Section 2.3 addresses the continuation of solutions, focusing on conditions under which a local solution extends to a maximal interval of existence.

2.1 ODEs, IVPs, solutions

2.1.1 Ordinary differential equation, initial value problem

Definition 2.1.1. *An nth order ordinary differential equation (ODE) is a functional relationship taking the form*

$$F\left(t, x(t), \frac{d}{dt}x(t), \frac{d^2}{dt^2}x(t), \ldots, \frac{d^n}{dt^n}x(t)\right) = 0,$$

that involves an independent variable $t \in \mathcal{I} \subset \mathbb{R}$, an unknown function $x(t) \in \mathcal{D} \subset \mathbb{R}^n$ of the independent variable, its derivative and derivatives of order up to n. For simplicity, the time dependence of x is often omitted, and we in general write equations as

$$F\left(t, x, x', x'', \ldots, x^{(n)}\right) = 0, \tag{2.1.1}$$

where $x^{(n)}$ denotes the nth order derivative of x. An equation such as (2.1.1) is said to be in general (or implicit) form.

An equation is said to be in normal (or explicit) form when it is written as

$$x^{(n)} = f\left(t, x, x', x'', \ldots, x^{(n-1)}\right).$$

Note that it is not always possible to write a differential equation in normal form, as it can be impossible to solve $F\left(t, x, \ldots, x^{(n)}\right) = 0$ in terms of $x^{(n)}$.

Definition 2.1.2. *(First-order ODE). In the following, we consider for simplicity the more restrictive case of a first-order ordinary differential equation in normal form*

$$x' = f(t, x). \tag{2.1.2}$$

Note that the theory developed here usually holds for nth order equations. The function f is assumed continuous and real-valued on a set $\mathcal{U} \subset \mathbb{R} \times \mathbb{R}^n$.

Definition 2.1.3. *(Initial value problem). An initial value problem (IVP) for equation (2.1.2) is given by*

$$\begin{cases} x' = f(t, x), \\ x\left(t_0\right) = x_0, \end{cases} \tag{2.1.3}$$

where f is continuous and real-valued on a set $\mathcal{U} \subset \mathbb{R} \times \mathbb{R}^n$, with $(t_0, x_0) \in \mathcal{U}$.

Remark 2.1.1. *An IVP for an nth order differential equation takes the form*

$$\begin{cases} x^{(n)} = f\left(t, x, x', \ldots, x^{(n-1)}\right) \\ x\left(t_0\right) = x_0, x'\left(t_0\right) = x_0', \ldots, x^{(n-1)}\left(t_0\right) = x_0^{(n-1)} \end{cases}$$

i.e., initial conditions have to be given for derivatives up to order $n-1$.

We have already seen in the previous sections that the order of an ODE is the order of the highest derivative involved in the equation, and an equation is then classified as a function of its linearity. A linear equation is one in which the vector field f takes the form

$$f(t, x) = a(t)x(t) + b(t).$$

If $b(t) = 0$ for all t, the equation is linear homogeneous; otherwise, it is linear nonhomogeneous. If the vector field f depends only on x, i.e., $f(t, x) = f(x)$ for all

t, then the equation is autonomous; otherwise, it is nonautonomous. Thus, a linear equation is autonomous if $a(t) = a$ and $b(t) = b$ for all t. Nonlinear equations are those that are not linear. They too, can be autonomous or nonautonomous.

2.1.2 Solutions of an ODE

Definition 2.1.4. *A function $\phi(t)$ is a solution to the ODE (2.1.2) if it satisfies this equation, that is, if*

$$\phi'(t) = f(t, \phi(t)),$$

for all $t \in \mathcal{I} \subset \mathbb{R}$, an open interval such that $(t, \phi(t)) \in \mathcal{U}$ for all $t \in \mathcal{I}$.

Definition 2.1.5. *(Integral form of the solution). The function*

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds, \tag{2.1.4}$$

is called the integral form of the solution to the IVP (2.1.3).

Theorem 2.1.1. *If $\mathbf{x}(t)$ is a solution of the IVP (2.1.3) in an interval $|t - t_0| < \alpha$, then*

$$\|\phi(t_1) - \phi(t_2)\| \leq M|t_1 - t_2|,$$

whenever t_1, t_2 are in the interval $|t - t_0| < \alpha$.

Proof. Let us begin by considering $t \geq t_0$. On $t_0 \leq t \leq t_0 + \alpha$,

$$\phi(t_1) - \phi(t_2) = x_0 + \int_{t_0}^{t_1} f(s, \phi(s))\, ds - x_0 - \int_{t_0}^{t_2} f(s, \phi(s))\, ds.$$

This simplifies to

$$\phi(t_1) - \phi(t_2) = -\int_{t_2}^{t_1} f(s, \phi(s))\, ds \quad \text{if} \quad t_2 > t_1,$$

and

$$\phi(t_1) - \phi(t_2) = \int_{t_1}^{t_2} f(s, \phi(s))\, ds \quad \text{if} \quad t_1 > t_2.$$

□

Theorem 2.1.2. *Suppose f is continuous on an open set $\mathcal{U} \subset \mathbb{R}\times\mathbb{R}^n$. Let $(t_0, x_0) \in \mathcal{U}$, and ϕ be a function defined on an open set $\mathcal{I}$ of $\mathbb{R}$ such that $t_0 \in \mathcal{I}$. Then ϕ is a solution of the IVP (2.1.3) if, and only if,*

(i) $\forall t \in \mathcal{I}, (t, \phi(t)) \in \mathcal{U}$.

(ii) ϕ *is continuous on* $\mathcal{I}$.

(iii) $\forall t \in \mathcal{I}, \phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds$.

Proof. ($\Rightarrow$) Let us suppose that $\phi' = f(t, \phi)$ for all $t \in \mathcal{I}$ and that $\phi(t_0) = x_0$. Then for all $t \in \mathcal{I}, (t, \phi(t)) \in \mathcal{U}$
. (i). Also, ϕ is differentiable and thus continuous on $\mathcal{I}$
(ii). Finally,

$$\phi'(s) = f(s, \phi(s))$$

so integrating both sides from t_0 to t,

$$\phi(t) - \phi(t_0) = \int_{t_0}^{t} f(s, \phi(s))ds$$

and thus

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds$$

hence (iii).
($\Leftarrow$) Assume $(i), (ii)$ and (iii). Then ϕ is differentiable on $\mathcal{I}$ and $\phi'(t) = f(t, \phi(t))$ for all $t \in \mathcal{I}$. From (iii), $\phi(t_0) = x_0 + \int_{t_0}^{t_0} f(s, \phi(s))ds = x_0$. □

Note that Theorem 2.1.2 states that ϕ should be continuous, whereas the solution should of course be C^1, for its derivative needs to be continuous. However, this is implied by point (iii). In fact, more generally, the following result holds about the regularity of solutions.

Theorem 2.1.3. *(Regularity). Let $f : \mathcal{U} \to \mathbb{R}^n$, with $\mathcal{U}$ an open set of $\mathbb{R} \times \mathbb{R}^n$. Suppose that $f \in C^k$. Then all solutions of (2.1.2) are of class C^{k+1}.*

Proof. The proof is obvious, since a solution ϕ is such that $\phi' \in C^k$. □

Geometric interpretation

The function f is the vector field of the equation. At every point in (t, x) space, a solution ϕ is tangent to the value of the vector field at that point. A particular consequence of this fact is the following theorem.

Theorem 2.1.4. *Let $x' = f(x)$ be a scalar autonomous differential equation. Then the solutions of this equation are monotone.*

Proof. The direction field of an autonomous scalar differential equation consists of vectors that are parallel for all t (since $f(t, x) = f(x)$ for all t). Suppose that a solution ϕ of $x' = f(x)$ is non monotone. Then this means that, given an initial point (t_0, x_0), one the following two occurs, as illustrated in Figure 2.1.
i) $f(x_0) \neq 0$ and there exists t_1 such that $\phi(t_1) = x_0$.
ii) $f(x_0) = 0$ and there exists t_1 such that $\phi(t_1) \neq x_0$.

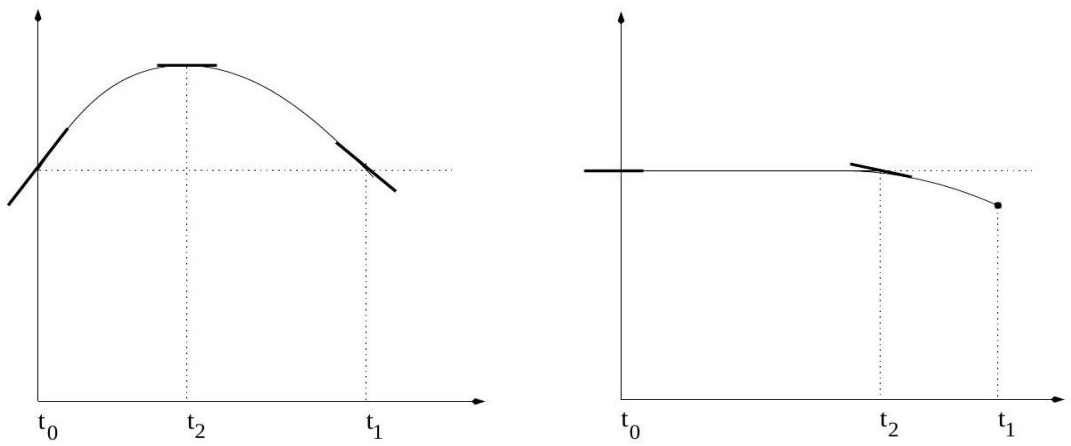


Figure 2.1: Situations that would lead to a scalar autonomous differential equation having nonmonotone solutions.

Suppose we are in case i), and assume we are in the case $f(x_0) > 0$. Thus, the solution curve ϕ is increasing at (t_0, x_0), i.e., $\phi'(t_0) > 0$. As ϕ is continuous, i) implies that there exists $t_2 \in (t_0, t_1)$ such that $\phi(t_2)$ is a maximum, with ϕ increasing for $t \in [t_0, t_2)$ and ϕ decreasing for $t \in (t_2, t_1]$. It follows that $\phi'(t_1) < 0$, which is a contradiction with $\phi'(t_0) > 0$.
Now assume that we are in case ii). Then there exists $t_2 \in (t_0, t_1)$ with $\phi(t_2) = x_0$ but such that $\phi'(t_2) < 0$. This is a contradiction. □

Remark 2.1.2. *If we have uniqueness of solutions, it follows from this theorem that if ϕ_1 and ϕ_2 are two solutions of the scalar autonomous differential equation $x' = f(x)$, then $\phi_1(t_0) < \phi_2(t_0)$ implies that $\phi_1(t) < \phi_2(t)$ for all t.*

Remark 2.1.3. *Theorem 2.1.4 is only true for scalar equations.*

2.2 Existence and uniqueness theorems

Several approaches can be used to show the existence and/or uniqueness of solutions.

2.2.1 Successive approximations method

Picard's successive approximation method consists of using the integral form (2.1.4) of the solution to the IVP (2.1.3) to construct a sequence of approximation of the solution, that converges to the solution. The steps followed in constructing this approximating sequence are the following.

Step 1. Start with an initial estimate of the solution, say, the constant function $\phi_0(t) = \phi_0 = x_0$, for $|t - t_0| \leq h$. Evidently, this function satisfies the IVP.

Step 2. Use ϕ_0 in (1.4) to define the second element in the sequence:

$$\phi_1(t) = x_0 + \int_{t_0}^{t} f\left(s, \phi_0(s)\right) ds.$$

Step 3. Use ϕ_1 in (1.4) to define the third element in the sequence

$$\phi_2(t) = x_0 + \int_{t_0}^{t} f\left(s, \phi_1(s)\right) ds.$$

...

Step n. Use ϕ_{n-1} in (1.4) to define the nth element in the sequence:

$$\phi_n(t) = x_0 + \int_{t_0}^{t} f\left(s, \phi_{n-1}(s)\right) ds.$$

At this stage, there are two major ways to tackle the problem, which use the same idea: if we can prove that the sequence $\{\phi_n\}$ converges, and that the limit happens to satisfy the differential equation, then we have the solution to the IVP (2.1.3).

The first method uses a fixed point approach. The second method studies explicitly the limit.

2.2.2 Local existence and uniqueness - Proof by fixed point theorem

Here, we present two slightly different formulations of the same theorem, which establishes that if the vector field is continuous and Lipschitz, then the solutions exist and are unique. We provide the proof for the second formulation.

Definition 2.2.1. *(Uniform Continuity). Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $E \subset X$ and $F \subset Y$. A function $f : E \to F$ is said to be uniformly continuous on the set $E \subset X$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$d_Y(f(x), f(y)) < \varepsilon \quad \text{whenever} \quad x, y \in E \text{ and } d_X(x, y) < \delta.$$

In other words,

$$f : E \subset (X, d_X) \to F \subset (Y, d_Y) \text{ is uniformly continuous on } E \quad \text{if and only if}$$

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in E,\ d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon.$$

Definition 2.2.2. *(Lipschitz function). A map $f : \mathcal{U} \subset \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is Lipschitz in x if there exists a real number K such that for all $(t, x_1) \in \mathcal{U}$ and $(t, x_2) \in \mathcal{U}$,*

$$\|f(t, x_1) - f(t, x_2)\| \leq K\|x_1 - x_2\|,$$

where K is independent of x_1 and x_2.

The following definition is a weaker version of Lipschitz functions.

Definition 2.2.3. *(Locally Lipschitz function). A map $f : \mathcal{U} \subset \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is locally Lipschitz in x if for all $(t_0, x_0) \in \mathcal{U}$, there exists a neighborhood $V \subset \mathcal{U}$ of (t_0, x_0) and a real number K such that for all (t, x_1) and $(t, x_2) \in V$,*

$$\|f(t, x_1) - f(t, x_2)\| \leq K\|x_1 - x_2\|.$$

In other words, f is locally Lipschtiz if the restriction of f to V is Lipschitz.

Thus, a locally Lipschitz function is Lipschitz if there exists a uniform Lipschitz constant L for all points in U. Another equivalent definition of a locally Lipschitz function is given below.

Definition 2.2.4. *A function $f : \mathcal{U} \subset \mathbb{R}^{n+1} \to \mathbb{R}^n$ is locally Lipschitz continuous if, for every compact set $V \subset \mathcal{U}$, the number*

$$L = \sup_{(t,x)\neq(t,y)\in V} \frac{\|f(t,x) - f(t,y)\|}{\|x - y\|}$$

is finite, with L depending on V.

Property. Let $f(t,x)$ be a function. The following properties hold.

1. f is Lipschitz $\Rightarrow$ f is uniformly continuous x.

2. f is uniformly continuous $\not\Rightarrow$ f is Lipschitz

3. $f(t,x)$ has continuous partial derivative $\dfrac{\partial f}{\partial x}$ on a bounded closed domain D $\Rightarrow$ f is locally Lipschitz on D.

Proof. **1.** Suppose that f is Lipschitz, meaning there exists a constant $L > 0$ such that

$$\|f(t,x_1) - f(t,x_2)\| \leq L\|x_1 - x_2\|, \quad \forall x_1, x_2.$$

Recall that f is uniformly continuous if, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$\|x_1 - x_2\| < \delta \quad \Rightarrow \quad \|f(t,x_1) - f(t,x_2)\| < \epsilon.$$

Given $\epsilon > 0$, we choose $\delta = \frac{\epsilon}{L}$. Then, for $\|x_1 - x_2\| < \delta$, we obtain

$$\|f(t,x_1) - f(t,x_2)\| \leq L\|x_1 - x_2\| < L\delta = \epsilon.$$

Hence, f is uniformly continuous.

2. This case is left as an exercise. Consider, for instance, the function

$$f(x) = \frac{1}{\ln x}, \quad x \in \left(0, \frac{1}{2}\right], \quad \text{with } f(0) = 0.$$

3. If $\dfrac{\partial f}{\partial x}$ is continuous on D, then the function $\left\|\dfrac{\partial f}{\partial x}\right\|$ is also continuous on the compact domain D. Consequently, it attains a finite upper bound on D. Define

$$L = \sup_{(t,x)\in D} \left\|\frac{\partial f}{\partial x}(t,x)\right\|.$$

Applying the mean-value theorem, for any $(t, x_1), (t, x_2) \in U$, there exists $\xi \in [x_1, x_2]$ such that

$$f(t, x_2) - f(t, x_1) = (x_2 - x_1)\frac{\partial f}{\partial x}(t, \xi).$$

Since $\xi \in \mathcal{U}$, it follows that

$$\left\|\frac{\partial f}{\partial x}(t, \xi)\right\| \leq L,$$

leading to

$$\|f(t, x_2) - f(t, x_1)\| \leq L\|x_2 - x_1\|.$$

This confirms that f is Lipschitz on $\mathcal{U}$. □

Lemma 2.2.1. *(Gronwall's lemma). Suppose that*
(i) *The function $g(t)$ is continuous for $t_0 \leq t \leq t_1$.*
(ii) *For $t_0 \leq t \leq t_1$, the function $g(t)$ satisfies the inequality*

$$0 \leq g(t) \leq K + L\int_{t_0}^{t} g(s)\, ds.$$

Then, it follows that

$$0 \leq g(t) \leq Ke^{L(t-t_0)}, \quad \textit{for all} \quad t_0 \leq t \leq t_1.$$

Proof. Define

$$v(t) = \int_{t_0}^{t} g(s)\, ds.$$

Then, differentiating both sides, we get

$$v'(t) = g(t),$$

which allows us to rewrite the given inequality as

$$0 \leq v'(t) \leq K + Lv(t).$$

The right-hand side forms a linear differential inequality. Using the integrating factor

$$\mu(t) = e^{-L(t-t_0)},$$

we multiply both sides by $\mu(t)$ to obtain

$$\frac{d}{dt}\Big(e^{-L(t-t_0)}v(t)\Big) \leq Ke^{-L(t-t_0)}.$$

Integrating from t_0 to t and using $v(t_0) = 0$, we get

$$e^{-L(t-t_0)}v(t) \leq \frac{K}{L}\Big(1 - e^{-L(t-t_0)}\Big).$$

Multiplying by $e^{L(t-t_0)}$,

$$v(t) \leq \frac{K}{L}\Big(e^{L(t-t_0)} - 1\Big).$$

Since $g(t) \leq K + Lv(t)$, it follows that

$$g(t) \leq K + L \cdot \frac{K}{L}(e^{L(t-t_0)} - 1) = Ke^{L(t-t_0)}.$$

Thus, the result is proven. □

Definition 2.2.5. *(Contraction Mapping). Let (X, d) be a metric space, and let $S \subset X$. A mapping $f : S \to S$ is called a contraction on S if there exists $K < 1$ such that, for all $x, y \in S$,*

$$d(f(x), f(y)) \leq Kd(x, y).$$

Every contraction is uniformly continuous on X.

Theorem 2.2.1. *(Contraction Mapping Principle). Let (X, d) be a complete metric space. Every contraction mapping $f : X \to X$ has a unique fixed point $x \in X$ such that $f(x) = x$.*

Proof. **Existence:** We use successive approximations. Let $x_0 \in X$. Define the sequence $x_1 = f(x_0), x_2 = f(x_1), \ldots, x_n = f(x_{n-1}), \ldots$, which defines an infinite sequence of elements of X. Since f is a contraction, we have

$$d(x_2, x_1) = d(f(x_1), f(x_0)) \leq Kd(x_1, x_0).$$

Similarly,

$$d(x_3, x_2) = d(f(x_2), f(x_1)) \leq Kd(x_2, x_1) \leq K^2 d(x_1, x_0).$$

Iterating this process gives

$$d(x_{n+1}, x_n) \leq K^n d(x_1, x_0).$$

Therefore,

$$\begin{aligned} d(x_{n+p}, x_n) &\leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (K^{p-1} + K^{p-2} + \cdots + K + 1)K^n d(x_1, x_0) \\ &\leq \frac{K^n}{1-K} d(x_1, x_0). \end{aligned}$$

Thus, $d(x_{n+p}, x_n)$ tends to 0 as $n \to \infty$, so $\{x_n\}$ is a Cauchy sequence. Since X is a complete space, it follows that $\{x_n\}$ converges to some limit $\ell \in X$. As $\lim_{n\to\infty} x_n = \ell$, and since f is continuous, we have

$$\lim_{n\to\infty} x_{n+1} = f(\ell).$$

But x_{n+1} also converges to ℓ, so $f(\ell) = \ell$, meaning that ℓ is a fixed point of f.
Uniqueness: Suppose ℓ_1 and ℓ_2 are two fixed points of f. Then we must have

$$d(\ell_1, \ell_2) \leq Kd(\ell_1, \ell_2) < d(\ell_1, \ell_2),$$

if $d(\ell_1, \ell_2) \neq 0$. Therefore, $d(\ell_1, \ell_2) = 0$, so $\ell_1 = \ell_2$. □

Theorem 2.2.2. *(Cauchy-Lipschitz-Picard theorem). Assume $f : \mathcal{U} \subset \mathbb{R} \times \mathbb{R}^n \to \mathcal{D} \subset \mathbb{R}^n$ is continuous, and that $f(t, x)$ satisfies a Lipschitz condition in $\mathcal{U}$ with respect to x. Then, given any point $(t_0, x_0) \in \mathcal{U}$, there exists a unique solution of (2.1.3) on some interval containing t_0 in its interior.*

Theorem 2.2.3. *(Picard's local existence and uniqueness theorem). Consider the IVP (2.1.3), and assume f is continuous in t and satisfies the Lipschitz condition*

$$\|f(t, x_1) - f(t, x_2)\| \leq L \|x_1 - x_2\|,$$

for all $x_1, x_2 \in D = \{x : \|x - x_0\| \leq b\}$ *and all* t *such that* $|t - t_0| \leq a$. *Then there exists* $0 < \delta \leq \alpha = \min\left(a, \frac{b}{M}\right)$ *such that (2.1.3) has a unique solution in* $|t - t_0| \leq \delta$.

To set up the proof, we proceed as follows.
Define the operator F by

$$F : x \mapsto x_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

Note that the function $(F\phi)(t)$ is a continuous function of t. Then Picard's successive approximations takes the form $\phi_1 = F\phi_0, \phi_2 = F\phi_1 = F^2\phi_0$, where F^2 represents $F \circ F$. Iterating, the general term is given by

$$\phi_k = F^k\phi_0, \quad \text{for } k = 0, \ldots$$

Therefore, finding the limit $\lim_{k\to\infty} \phi_k$ is equivalent to finding the function ϕ, solution of the fixed point problem

$$x = Fx,$$

with x a continuously differentiable function. Thus, a solution of (2.1.3) is a fixed point of F, and we aim to use the contraction mapping principle to verify the existence (and uniqueness) of such a fixed point.

Proof. (See [11], p. 56-58). We now proceed to demonstrate the result for the interval $t - t_0 \leq \delta$, with the proof for the interval $t_0 - t \leq \delta$ being analogous. Let X denote the space of continuous functions defined on the interval $[t_0, t_0 + \delta]$, i.e., $X = C([t_0, t_0 + \delta])$, which is endowed with the sup norm, defined for $x \in X$ as

$$\|x\|_c = \max_{t\in[t_0,t_0+\delta]} \|x(t)\|.$$

This norm represents the uniform convergence norm. Next, define

$$S = \{x \in X : \|x - x_0\|_c \leq b\}.$$

Clearly, $S \subset X$. Furthermore, S is closed, and since X is a complete metric space with the sup norm, we have the necessary properties for the application of further

results. It is important to note that we have transformed the problem into one involving the space of continuous functions, which places us in the infinite-dimensional setting. The proof unfolds in three steps.

Step 1. We begin by showing that $F : S \to S$. From equation (2.1.4),

$$(F\phi)(t) - x_0 = \int_{t_0}^{t} f(s, \phi(s))\, ds = \int_{t_0}^{t} ((f(s, \phi(s)) - f(s, x_0)) + f(s, x_0))\, ds.$$

By the triangle inequality, we get

$$\|F\phi - x_0\| \leq \int_{t_0}^{t} (\|f(s, \phi(s)) - f(s, x_0)\| + \|f(s, x_0)\|)\ ds.$$

Since f is (piecewise) continuous, it is bounded on $[t_0, t_1]$, and there exists a constant $M = \max_{t\in[t_0,t_1]} \|f(t, x_0)\|$. Thus, we obtain

$$\|F\phi - x_0\| \leq \int_{t_0}^{t} (\|f(s, \phi(s)) - f(s, x_0)\| + M)\, ds \leq \int_{t_0}^{t} (L\|\phi(s) - x_0\| + M)\, ds,$$

since f is Lipschitz. For all $\phi \in S$, we have $\|\phi - x_0\| \leq b$, so

$$\|F\phi - x_0\| \leq \int_{t_0}^{t} (Lb + M)\ ds \leq (t - t_0)(Lb + M).$$

Since $t \in [t_0, t_0 + \delta]$, we have $t - t_0 \leq \delta$, and thus

$$\|F\phi - x_0\|_c = \max_{t\in[t_0,t_0+\delta]} \|F\phi - x_0\| \leq (Lb + M)\delta.$$

Choose δ such that $\delta \leq \dfrac{b}{Lb + M}$, i.e., for t sufficiently close to t_0. Then we have

$$\|F\phi - x_0\|_c \leq b,$$

which implies that for $\phi \in S$, $F\phi \in S$. Hence, $F : S \to S$.

Step 2. We now show that F is a contraction. Let $\phi_1, \phi_2 \in S$. We have

$$\|(F\phi_1)(t) - (F\phi_2)(t)\| = \left\| \int_{t_0}^{t} (f(s, \phi_1(s)) - f(s, \phi_2(s)))\ ds \right\|.$$

Using the triangle inequality and the Lipschitz condition, we get

$$\|(F\phi_1)(t) - (F\phi_2)(t)\| \leq \int_{t_0}^{t} \|f(s,\phi_1(s)) - f(s,\phi_2(s))\|\, ds \leq \int_{t_0}^{t} L\|\phi_1(s) - \phi_2(s)\|\, ds.$$

Thus,

$$\|F\phi_1 - F\phi_2\|_c \leq L\delta\|\phi_1 - \phi_2\|_c \leq \rho\|\phi_1 - \phi_2\|_c \quad \text{for} \quad \delta \leq \frac{\rho}{L}.$$

Therefore, by choosing $\rho < 1$ and $\delta \leq \dfrac{\rho}{L}$, we see that F is a contraction. Since, by Step 1, $F : S \to S$, the contraction mapping principle (Theorem 2.2.1) implies that F has a unique fixed point in S, and equation (2.1.3) has a unique solution in S.

Step 3. It remains to show that any solution in X is actually in S, as the result is to be shown in X. Consider a solution starting at x_0 at time t_0. The solution leaves S if there exists a $t > t_0$ such that $\|\phi(t) - x_0\| = b$, i.e., the solution crosses the boundary of D. Let $\tau > t_0$ be the first such time. For all $t_0 \leq t \leq \tau$, we have

$$\begin{aligned}
\|\phi(t) - x_0\| &\leq \int_{t_0}^{t} \left(\|f(s,\phi(s)) - f(s,x_0)\| + \|f(s,x_0)\|\right) ds \\
&\leq \int_{t_0}^{t} \left(L\|\phi(s) - x_0\| + M\right) ds \\
&\leq \int_{t_0}^{t} (Lb + M)\, ds.
\end{aligned}$$

Therefore,

$$b = \|\phi(\tau) - x_0\| \leq (Lb + M)(\tau - t_0).$$

Let $\tau = t_0 + \mu$, for some $\mu > 0$. It follows that if

$$\mu > \frac{b}{Lb + M},$$

then the solution ϕ remains within D. □

Note that the condition $x_1, x_2 \in D = \{x : \|x - x_0\| \leq b\}$ in the statement of the theorem refers to a local Lipschitz condition. If the function f is Lipschitz, then the following theorem holds.

Theorem 2.2.4. *(Global existence). Suppose that f is continuous in t and is Lipschitz on $\mathcal{U} = \mathcal{I} \times D$. Then (2.1.3) admits a unique solution on $\mathcal{I}$.*

2.2.3 Local existence and uniqueness - Proof by successive approximations method

Using the method of successive approximations, we can prove the following theorem.

Theorem 2.2.5. *(Existence and Uniqueness via Picard Iteration). Suppose that f is continuous on a domain $\mathcal{R}$ of the (t,x)-plane defined, for $a,b>0$, by $\mathcal{R}=\{(t,x):|t-t_0|\le a,\|x-x_0\|\le b\}$, and that f is locally Lipschitz in x on $\mathcal{R}$. Let then, as previously defined,*

$$M=\sup_{(t,x)\in\mathcal{R}}\|f(t,x)\|<\infty \quad \textit{and} \quad \alpha=\min\left(a,\frac{b}{M}\right).$$

Then the sequence defined by

$$\begin{aligned}\phi_0&=x_0, \quad |t-t_0|\le\alpha\\ \phi_i(t)&=x_0+\int_{t_0}^{t} f\left(s,\phi_{i-1}(s)\right)ds, \quad i\ge 1, \quad |t-t_0|\le\alpha\end{aligned}$$

converges uniformly on the interval $|t-t_0|\le\alpha$ to ϕ, unique solution of (2.1.3).

Proof. See ([16], p. 3-6).

Existence:

Assume that $|t-t_0|\le\alpha$. Then, we have

$$\|\phi_1-\phi_0\|=\left\|\int_{t_0}^{t} f\left(s,\phi_0(s)\right)ds\right\|.$$

Applying the bound on f, it follows that

$$\|\phi_1-\phi_0\|\le M\,|t-t_0|\le\alpha M\le b.$$

By the definitions of M and α, we conclude that $\|\phi_1-\phi_0\|\le b$, ensuring that the integral $\int_{t_0}^{t} f\left(s,\phi_1(s)\right)ds$ is well-defined for $|t-t_0|\le\alpha$. Consequently,

$$\|\phi_2(t)-\phi_0\|=\left\|\int_{t_0}^{t} f\left(s,\phi_1(s)\right)ds\right\|\le\int_{t_0}^{t}\|f\left(s,\phi_1(s)\right)\|\,ds\le\alpha M\le b.$$

By induction, it follows that for all $k = 1, \ldots, n$ and for $|t - t_0| \leq \alpha$,

$$\|\phi_k(t) - \phi_0\| \leq \alpha M \leq b.$$

Now, For $|t - t_0| \leq \alpha$, we obtain

$$\begin{aligned}\|\phi_{k+1}(t) - \phi_k(t)\| &= \left\| x_0 + \int_{t_0}^{t} f\left(s, \phi_k(s)\right) ds - x_0 - \int_{t_0}^{t} f\left(s, \phi_{k-1}(s)\right) ds \right\| \\ &= \left\| \int_{t_0}^{t} \left(f\left(s, \phi_k(s)\right) - f\left(s, \phi_{k-1}(s)\right)\right) ds \right\|.\end{aligned}$$

By the Lipschitz continuity of f in x on $\mathcal{R}$, we conclude that

$$\|\phi_{k+1}(t) - \phi_k(t)\| \leq L \int_{t_0}^{t} \|\phi_k(s) - \phi_{k-1}(s)\| \, ds.$$

We now prove that, for all k,

$$\|\phi_{k+1} - \phi_k\| \leq b \frac{\left(L\,|t - t_0|\right)^k}{k!} \text{ for } |t - t_0| \leq \alpha. \tag{2.2.5}$$

Indeed, (2.2.5) holds for $k = 1$, as previously established. Assume that (2.2.5) holds for $k = n$. Then

$$\begin{aligned}\|\phi_{n+2} - \phi_{n+1}\| &= \left\| \int_{t_0}^{t} f\left(s, \phi_{n+1}(s)\right) - f\left(s, \phi_n(s)\right) ds \right\| \\ &\leq \int_{t_0}^{t} L \, \|\phi_{n+1}(s) - \phi_n(s)\| \, ds \\ &\leq \int_{t_0}^{t} Lb \frac{\left(L\,|s - t_0|\right)^n}{n!} ds \text{ for } |t - t_0| \leq \alpha \\ &\leq b \frac{L^{n+1}}{n!} \frac{|t - t_0|^{n+1}}{n+1} \Bigg|_{s=t_0}^{s=t} \\ &\leq b \frac{\left(L\,|t - t_0|\right)^{n+1}}{(n+1)!}\end{aligned}$$

and thus (2.2.5) holds for $k = 1, \ldots$.

Thus, for $N > n$ we have

$$\|\phi_N(t) - \phi_n(t)\| \leq \sum_{k=n}^{N-1} \|\phi_{k+1}(t) - \phi_k(t)\| \leq \sum_{k=n}^{N-1} b\frac{(L\,|t-t_0|)^k}{k!} \leq b\sum_{k=n}^{N-1}\frac{(L\alpha)^k}{k!}.$$

The rightmost term in this expression tends to zero as $n \to \infty$. Therefore, the sequence $\{\phi_k(t)\}$ converges uniformly to a function $\phi(t)$ on the interval $|t - t_0| \leq \alpha$. Since the convergence is uniform, the limit function is continuous. Moreover, $\phi(t_0) = x_0$. In fact, we can write

$$\phi_N(t) = \phi_0(t) + \sum_{k=1}^{N} \left(\phi_k(t) - \phi_{k-1}(t)\right),$$

which implies that

$$\phi(t) = \phi_0(t) + \sum_{k=1}^{\infty} \left(\phi_k(t) - \phi_{k-1}(t)\right).$$

The fact that ϕ is a solution to (2.1.3) follows from the following result. If a sequence of functions $\{\phi_k(t)\}$ converges uniformly, and each $\phi_k(t)$ is continuous on the interval $|t - t_0| \leq \alpha$, then

$$\lim_{n\to\infty} \int_{t_0}^{t} \phi_n(s)\,ds = \int_{t_0}^{t} \lim_{n\to\infty} \phi_n(s)\,ds.$$

Thus, we have

$$\begin{aligned}
\phi(t) &= \lim_{n\to\infty} \phi_n(t), \\
&= x_0 + \lim_{n\to\infty} \int_{t_0}^{t} f\left(s, \phi_{n-1}(s)\right)\,ds, \\
&= x_0 + \int_{t_0}^{t} \lim_{n\to\infty} f\left(s, \phi_{n-1}(s)\right)\,ds, \\
&= x_0 + \int_{t_0}^{t} f(s, \phi(s))\,ds,
\end{aligned}$$

which means that

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\,ds \quad \text{for} \quad |t - t_0| \leq \alpha.$$

Since the integrand $f(t, \phi)$ is continuous, ϕ is differentiable with respect to t, and

its derivative is given by

$$\phi'(t) = f(t, \phi(t)),$$

which shows that ϕ is a solution to the initial value problem (IVP) (2.1.3).

Uniqueness:

Let ϕ and ψ be two solutions of (2.1.3), i.e., for $|t - t_0| \leq \alpha$,

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds,$$
$$\psi(t) = x_0 + \int_{t_0}^{t} f(s, \psi(s))ds.$$

Then, for $|t - t_0| \leq \alpha$,

$$\begin{aligned} \|\phi(t) - \psi(t)\| &= \left\| \int_{t_0}^{t} f(s, \phi(s)) - f(s, \psi(s))ds \right\| \\ &\leq L \int_{t_0}^{t} \|\phi(s) - \psi(s)\|ds. \end{aligned} \tag{2.2.6}$$

We now apply Gronwall's Lemma to this inequality, using $K = 0$ and $g(t) = \|\phi(t) - \psi(t)\|$. First, applying the lemma for $t_0 \leq t \leq t_0 + \alpha$, we get $0 \leq \|\phi(t) - \psi(t)\| \leq 0$, that is,

$$\|\phi(t) - \psi(t)\| = 0$$

and thus $\phi(t) = \psi(t)$ for $t_0 \leq t \leq t_0 + \alpha$. Similarly, for $t_0 - \alpha \leq t \leq t_0$, $\|\phi(t) - \psi(t)\| = 0$. Therefore, $\phi(t) = \psi(t)$ on $|t - t_0| \leq \alpha$. □

Example 2.2.1. *Let us consider the initial value problem (IVP)*

$$x' = -x, \quad x(0) = x_0 = c, \quad c \in \mathbb{R}.$$

For the initial solution, we choose $\phi_0(t) = c$. Then,

$$\begin{aligned} \phi_1(t) &= x_0 + \int_0^t f\left(s, \phi_0(s)\right) ds \\ &= c + \int_0^t -\phi_0(s)\, ds \\ &= c - c \int_0^t 1\, ds \\ &= c - ct. \end{aligned}$$

To find ϕ_2, we use ϕ_1 in (2.1.4):

$$\begin{aligned}\phi_2(t) &= x_0 + \int_0^t f\left(s, \phi_1(s)\right)\, ds \\ &= c - \int_0^t (c - cs)\, ds \\ &= c - ct + c\frac{t^2}{2}.\end{aligned}$$

Continuing this method, we find a general term of the form

$$\phi_n(t) = \sum_{i=0}^{n} c\frac{(-1)^i t^i}{i!}.$$

This is the power series expansion of ce^{-t}, so we conclude that

$$\phi_n(t) \to \phi(t) = ce^{-t} \quad \textit{(and the approximation is valid on } \mathbb{R}\textit{)},$$

which is the solution to the initial value problem.

It is important to note that the method of successive approximations is highly versatile and can be extended to much broader contexts; see ([4], p. 264-269).

2.2.4 Local existence (non Lipschitz case)

The following theorem is frequently referred to as Peano's Existence Theorem. Since the vector field is not required to satisfy the Lipschitz condition, it is important to note that uniqueness of the solution may not hold.

Theorem 2.2.6. *(Peano's theorem). Suppose that f is continuous on some region*

$$\mathcal{R} = \{(t, x) : |t - t_0| \leq a, \|x - x_0\| \leq b\}$$

with $a, b > 0$, and let $M = \max_{\mathcal{R}} \|f(t, x)\|$. Then there exists a continuous function $\phi(t)$, differentiable on $\mathcal{R}$, such that

(i) $\phi\left(t_0\right) = x_0$,

(ii) $\phi'(t) = f(t, \phi)$ *on* $|t - t_0| \leq \alpha$, *where*

$$\alpha = \begin{cases} a & \text{if } M = 0 \\ \min\left(a, \frac{b}{M}\right) & \text{if } M > 0. \end{cases}$$

Before proving this result, we must introduce a number of preliminary notations and results. The definition of equicontinuity and a statement of the Ascoli lemma are provided below. In order to construct a solution without the Lipschitz condition, we approximate the differential equation by another one that satisfies the Lipschitz condition. The unique solution of this approximate problem serves as an ε-approximate solution, which is formally defined in [4].

Definition 2.2.6. *(Equicontinuous Set). A set of functions* $\mathcal{F} = \{f\}$ *defined on a real interval* I *is said to be equicontinuous on* I *if, for every* $\varepsilon > 0$, *there exists* $\delta_\varepsilon > 0$, *independent of* $f \in \mathcal{F}$ *and also of* $t, \tilde{t} \in I$, *such that*

$$\|f(t) - f(\tilde{t})\| < \varepsilon \quad \text{whenever} \quad |t - \tilde{t}| < \delta_\varepsilon.$$

An interpretation of equicontinuity is that a sequence of functions is equicontinuous if all the functions are continuous and exhibit uniform variation over a given neighborhood.

Theorem 2.2.7. *Let* $\{f_n\}$ *be an equicontinuous sequence of functions. If* $f_n(x) \to f(x)$ *for every* $x \in X$, *then the function* f *is continuous.*

Lemma 2.2.2. *(Ascoli). Let* I *be a bounded interval, and let* $F = \{f\}$ *be an infinite, uniformly bounded, equicontinuous set of functions. Then* F *contains a sequence* $\{f_n\}$, *where* $n = 1, 2, \ldots$, *that is uniformly convergent on* I.

Definition 2.2.7. *(ε-approximate solution). A differentiable mapping* u *of an open ball* $J \in \mathcal{I}$ *into* $\mathcal{U}$ *is an approximate solution of* $x' = f(t, x)$ *with approximation* ε *(or an* ε*-approximate solution) if we have*

$$\|u'(t) - f(t, u(t))\| \leq \varepsilon \quad \text{for any } t \in J.$$

Lemma 2.2.3. *Suppose that $f(t,x)$ is continuous on a region*

$$\mathcal{R} = \{(t,x) : |t-t_0| \leq a, \|x-x_0\| \leq b\}.$$

Then, for every positive number ε, there exists a function $F_\varepsilon(t,x)$ such that

(i) *F_ε is continuous for $|t-t_0| \leq a$ and all x,*

(ii) *F_ε has continuous partial derivatives of all orders with respect to $x_1, \ldots, x_n$ for $|t-t_0| \leq a$ and all x,*

(iii) *$\|F_\varepsilon(t,x)\| \leq \max_{\mathcal{R}} \|f(t,x)\| = M$ for $|t-t_0| \leq a$ and all x,*

(iv) *$\|F_\varepsilon(t,x) - f(t,x)\| \leq \varepsilon$ on $\mathcal{R}$.*

For a proof of this lemma, see [9].
We now prove Peano's Theorem 2.2.6.

Proof. (Peano's Theorem proof). The proof takes four steps.

1. We construct, for every positive number ε, a function $F_\varepsilon(t,x)$ that satisfies the requirements given in Lemma 2.2.3. Using an existence-uniqueness result in the Lipschitz case (such as Theorem 2.2.3), we construct a function $\phi_\varepsilon(t)$ such that it holds the following properties:

   - (P1) $\phi_\varepsilon(t_0) = x_0$,
   - (P2) $\phi'_\varepsilon(t) = F_\varepsilon(t, \phi_\varepsilon(t))$ on $|t-t_0| < \alpha$,
   - (P3) $(t, \phi_\varepsilon(t)) \in \mathcal{R}$ on $|t-t_0| \leq \alpha$.

2. The set $\mathcal{F} = \{\phi_\varepsilon : \varepsilon > 0\}$ is bounded and equicontinuous on $|t-t_0| \leq \alpha$. Indeed, property (P3) of ϕ_ε implies that $\mathcal{F}$ is bounded on $|t-t_0| \leq \alpha$, and that
   $\|F_\varepsilon(t, \phi_\varepsilon(t))\| \leq M$ on $|t-t_0| \leq \alpha$. Hence, property (P2) of ϕ_ε implies that

   $$\|\phi_\varepsilon(t_1) - \phi_\varepsilon(t_2)\| \leq M|t_1 - t_2|,$$

   if $|t_1-t_0| \leq \alpha$ and $|t_2-t_0| \leq \alpha$ (this is a consequence of Theorem 2.1.1). Therefore, for a given positive number μ, we have $\|\phi_\varepsilon(t_1) - \phi_\varepsilon(t_2)\| \leq \mu$ whenever $|t_1-t_0| \leq \alpha$, $|t_2-t_0| \leq \alpha$, and $|t_1-t_2| \leq \dfrac{\mu}{M}$.

3. Using Lemma 2.2.2, choose a sequence $\{\varepsilon_k : k = 1, 2, \ldots\}$ of positive numbers such that $\lim_{k\to\infty} \varepsilon_k = 0$, and that the sequence $\{\phi_{\varepsilon_k} : k = 1, 2, \ldots\}$ converges uniformly on $|t - t_0| \le \alpha$ as $k \to \infty$. Then set

$$\phi(t) = \lim_{k\to\infty} \phi_{\varepsilon_k}(t) \quad \text{on} \quad |t - t_0| \le \alpha.$$

4. Observe that

$$\phi_\varepsilon(t) = x_0 + \int_{t_0}^{t} F_\varepsilon(s, \phi_\varepsilon(s))\, ds$$

and

$$\phi_\varepsilon(t) = x_0 + \int_{t_0}^{t} f(s, \phi_\varepsilon(s))\, ds + \int_{t_0}^{t} \left(F_\varepsilon(s, \phi_\varepsilon(s)) - f(s, \phi_\varepsilon(s))\right)\, ds.$$

It follows from (iv) in Lemma 2.2.3 that

$$\left\| \int_{t_0}^{t} \left(F_\varepsilon(s, \phi_\varepsilon(s)) - f(s, \phi_\varepsilon(s))\right)\, ds \right\| \le \varepsilon |t - t_0|$$

on $|t - t_0| \le \alpha$. This is true for all $\varepsilon \ge 0$, so letting $\varepsilon \to 0$, we obtain

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds.$$

□

A proof, by Hartman ([7], p. 10-11), is as follows.

Proof. Let $\delta > 0$ and let $\phi_\ell(t)$ be a continuously differentiable n-dimensional vector-valued function defined on the interval $[t_0 - \delta, t_0]$ that satisfies the following conditions:

$$\phi_\ell(t_0) = x_0, \quad \phi_\ell'(t_0) = f(t_0, x_0), \quad \|\phi_\ell(t) - x_0\| \le b, \quad \|\phi_\ell'(t)\| \le M.$$

For $0 < \varepsilon \le \delta$, define the function $\phi_\varepsilon(t)$ on the interval $[t_0 - \delta, t_0 + \alpha]$ by setting

$$\phi_\varepsilon(t) = \phi_\ell(t) \quad \text{for} \quad t \in [t_0 - \delta, t_0].$$

and

$$\phi_\varepsilon(t) = x_0 + \int_{t_0}^{t} f\left(s, \phi_\varepsilon(s-\varepsilon)\right) ds \text{ on } [t_0, t_0+\alpha]. \tag{2.2.7}$$

The function ϕ_ε can indeed be defined on the interval $[t_0-\delta, t_0+\alpha]$. To show this, note that the formula for $\phi_\varepsilon(t)$ is well-defined for $t_0 \le t \le t_0+\alpha_1$, where $\alpha_1 = \min(\alpha, \varepsilon)$. Thus, $\phi_\varepsilon(t)$ is continuously differentiable on the interval $[t_0-\delta, t_0+\alpha_1]$ and, On this interval,

$$\|\phi_\varepsilon(t) - x_0\| \le b, \quad \|\phi_\varepsilon(t) - \phi_\varepsilon(s)\| \le M|t-s|. \tag{2.2.8}$$

It then follows that (2.2.7) can be used to extend $\phi_\varepsilon(t)$ as a C^1 function over $[t_0-\delta, t_0+\alpha_2]$, where $\alpha_2 = \min(\alpha, 2\varepsilon)$, satisfying relation (2.2.8). Continuing in this fashion, (2.2.7) serves to define $\phi_\varepsilon(t)$ over $[t_0, t_0+\alpha]$ so that $\phi_\varepsilon(t)$ is a C^0 function on $[t_0-\delta, t_0+\alpha]$, satisfying relation (2.2.8).
Since $\|\phi'_\varepsilon(t)\| \le M, M$ can be used as a Lipschitz constant for ϕ_ε, giving uniform continuity of ϕ_ε. It follows that the family of functions, $\phi_\varepsilon(t), 0 < \varepsilon \le \delta$, is equicontinuous. Thus, using Ascoli's Lemma, there exists a sequence $\varepsilon(1) > \varepsilon(2) > \ldots$, such that $\varepsilon(n) \to 0$ as $n \to \infty$ and

$$\phi(t) = \lim_{n\to\infty} \phi_{\varepsilon(n)}(t) \text{ exists uniformly on } [t_0-\delta, t_0+\alpha].$$

The continuity of f implies that $f\left(t, \phi_{\varepsilon(n)}(t-\varepsilon(n))\right.$ tends uniformly to $f(t, \phi(t))$ as $n \to \infty$; thus term-by-term integration of (2.2.7) where $\varepsilon = \varepsilon(n)$ gives

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s)) ds,$$

and thus $\phi(t)$ is a solution of (2.1.3).

□

An important corollary follows.

Corollary 2.2.1. *Let $f(t,x)$ be continuous on an open set E and satisfy $\|f(t,x)\| \le M$. Let E_0 be a compact subset of E. Then there exists an $\alpha = \alpha(E, E_0, M) > 0$ with the property that if $(t_0, x_0) \in E_0$, then the IVP (2.1.3) has a solution and every solution exists on $|t-t_0| \le \alpha$.*

In fact, the hypotheses can be slightly relaxed. In [3] an ε-approximate solution is

defined as follows.

Definition 2.2.8. *An ε-approximate solution of the differential equation (2.1.2), where f is continuous, on a t-interval I is a function $\phi \in C$ on I such that:*
1. *$(t, \phi(t)) \in \mathcal{U}$ for $t \in I$;*
2. *$\phi \in C^1$ on I, except possibly for a finite set of points S on I, where ϕ' may have simple discontinuities (i.e., g has finite discontinuities at c if the left and right limits of g at c exist but are not equal);*
3. *$\|\phi'(t) - f(t, \phi(t))\| \leq \varepsilon$ for $t \in I \setminus S$.*

Hence it is assumed that ϕ has a continuous derivative on I, which is denoted by $\phi \in C_p^1(I)$.

Theorem 2.2.8. *Let $f \in C$ on the rectangle*

$$\mathcal{R} = \{(t, x) : |t - t_0| \leq a, \|x - x_0\| \leq b\}.$$

Given any $\varepsilon > 0$, there exists an ε-approximate solution ϕ of (2.1.3) on $|t - t_0| \leq \alpha$ such that $\phi(t_0) = x_0$.

Proof. Let $\varepsilon > 0$ be given. We will construct an ε-approximate solution on the interval $[t_0, t_0 + \varepsilon]$; the construction is similar for the interval $[t_0 - \alpha, t_0]$. The ε-approximate solution that we construct will be a polygonal path, starting at (t_0, x_0). Since f is continuous on $\mathcal{R}$, it is uniformly continuous on $\mathcal{R}$. Therefore, for the given value of ε, there exists $\delta_\varepsilon > 0$ such that

$$\|f(t, \phi) - f(\tilde{t}, \tilde{\phi})\| \leq \varepsilon, \tag{2.2.9}$$

whenever

$$(t, \phi), (\tilde{t}, \tilde{\phi}) \in \mathcal{R}, \quad |t - \tilde{t}| \leq \delta_\varepsilon, \quad \|\phi - \tilde{\phi}\| \leq \delta_\varepsilon.$$

Now, divide the interval $[t_0, t_0 + \alpha]$ into n subintervals: $t_0 < t_1 < \cdots < t_n = t_0 + \alpha$, such that

$$\max |t_k - t_{k-1}| \leq \min\left(\delta_\varepsilon, \frac{\delta_\varepsilon}{M}\right). \tag{2.2.10}$$

Starting at (t_0, x_0), construct a line segment with slope $f(t_0, x_0)$, which intercepts the line $t = t_1$ at (t_1, x_1). By the definition of α and M, it is clear that this line

segment lies inside the triangular region T, which is bounded by the line segments with slopes $\pm M$ from (t_0, x_0) to their intercept at $t = t_0 + \alpha$, and the line $t = t_0 + \alpha$. In particular, $(t_1, x_1) \in T$.

At the point (t_1, x_1), construct a line segment with slope $f(t_1, x_1)$ until the line $t = t_2$, yielding the point (t_2, x_2). Continue this process iteratively. The polygonal path ϕ is thus constructed, which intersects the line $t = t_0 + \alpha$ after a finite number of steps and lies entirely within the region T.

The function ϕ, which can be expressed as

$$\begin{aligned} \phi(t_0) &= x_0, \\ \phi(t) &= \phi(t_{k-1}) + f(t_{k-1}, \phi(t_{k-1}))(t - t_{k-1}), \quad t \in [t_{k-1}, t_k], \quad k = 1, \ldots, n, \end{aligned} \tag{2.2.11}$$

is the desired ε-approximate solution. Clearly, $\phi \in C_p^1([t_0, t_0 + \alpha])$, and

$$\|\phi(t) - \phi(\tilde{t})\| \leq M|t - \tilde{t}| \quad \text{for} \quad t, \tilde{t} \in [t_0, t_0 + \alpha]. \tag{2.2.12}$$

For $t \in [t_{k-1}, t_k]$, from the above inequality (2.2.12) and the condition on the partition in equation (2.2.10), we deduce that

$$\|\phi(t) - \phi(t_{k-1})\| \leq \delta_\varepsilon.$$

Moreover, from (2.2.11) and (2.2.9), we obtain

$$\|\phi'(t) - f(t, \phi(t))\| = \|f(t_{k-1}, \phi(t_{k-1})) - f(t, \phi(t))\| \leq \varepsilon.$$

Thus, ϕ is an ε-approximate solution. □

2.2.5 Some examples of existence and uniqueness

Example 2.2.2. *Consider the IVP*

$$\begin{aligned} x' &= 3|x|^{\frac{2}{3}} \\ x(t_0) &= x_0 \end{aligned} \tag{2.2.13}$$

Here, Theorem 2.2.6 applies, since $f(t, x) = 3x^{2/3}$ is continuous. However, Theorem 2.2.3 does not apply, since $f(t, x)$ is not locally Lipschitz in $x = 0$ (or, f is not

Lipschitz on any interval containing 0). This means that we have existence of solutions to this IVP, but not uniqueness of the solution.

The fact that f is not Lipschitz on any interval containing 0 is established using the following argument. Suppose that f is Lipschitz on an interval $\mathcal{I} = (-\varepsilon, \varepsilon)$, with $\varepsilon > 0$. Then, there exists $L > 0$ such that for all $x_1, x_2 \in \mathcal{I}$,

$$\|f(t, x_1) - f(t, x_2)\| \leq L |x_1 - x_2|$$

that is,

$$3 \left| |x_1|^{\frac{2}{3}} - |x_2|^{\frac{2}{3}} \right| \leq L |x_1 - x_2|$$

Since this has to hold true for all $x_1, x_2 \in \mathcal{I}$, it must hold true in particular for $x_2 = 0$. Thus

$$3 |x_1|^{\frac{2}{3}} \leq L |x_1| .$$

Given an $\varepsilon > 0$, it is possible to find $N_\varepsilon > 0$ such that $\frac{1}{n} < \varepsilon$ for all $n \geq N_\varepsilon$. Let $x_1 = \frac{1}{n}$. Then for $n \geq N_\varepsilon$, if f is Lipschitz there must hold

$$3 \left(\frac{1}{n} \right)^{\frac{2}{3}} \leq \frac{L}{n}.$$

So, for all $n \geq N_\varepsilon$,

$$n^{\frac{1}{3}} \leq \frac{L}{3}.$$

This is a contradiction, since $\lim_{n \to \infty} n^{1/3} = \infty$, and so f is not Lipschitz on $\mathcal{I}$.
Let us consider the set

$$E = \{t \in \mathbb{R} : x(t) = 0\}.$$

The set E can have several forms, depending on the situation.

1. *$E = \emptyset$.*
2. *$E = [a, b]$, (closed since x is continuous and thus reaches its bounds).*
3. *$E = (-\infty, b)$.*
4. *$E = (a, +\infty)$.*
5. *$E = \mathbb{R}$.*

Note that case (2) includes the case of a single intersection point, when $a = b$, giving $E = \{a\}$. Let us now consider the nature of x in these different situations. Recall

that from Theorem 2.1.4, since (2.2.13) is defined by a scalar autonomous equation, its solutions are monotone. For simplicity, we consider here the case of monotone increasing solutions. The case of monotone decreasing solutions can be treated in a similar fashion.

1. Here, there is no intersection with the $x = 0$ axis. Thus if follows that

$$x(t) \text{ is } \begin{cases} > 0, \text{ if } x_0 > 0 \\ < 0, \text{ if } x_0 < 0. \end{cases}$$

2. In this case,

$$x(t) \text{ is } \begin{cases} < 0, \text{ if } t < a \\ = 0, \text{ if } t \in [a, b] \\ > 0, \text{ if } t > b. \end{cases}$$

3. Here,

$$x(t) \text{ is } \begin{cases} = 0, \text{ if } t < b \\ > 0, \text{ if } t > b. \end{cases}$$

4. In this case,

$$x(t) \text{ is } \begin{cases} < 0, \text{ if } t < a \\ = 0, \text{ if } t > a. \end{cases}$$

5. In this last case, $x(t) = 0$ for all $t \in \mathbb{R}$.

Now, depending on the sign of x, we can integrate the equation. First, if $x > 0$, then $|x| = x$ and so

$$\begin{aligned} & x' = 3x^{2/3} \\ \Leftrightarrow \quad & \frac{1}{3} x^{-2/3} x' = 1 \\ \Leftrightarrow \quad & x^{1/3} = t + k_1 \\ \Leftrightarrow \quad & x(t) = (t + k_1)^3, \end{aligned}$$

for $k_1 \in \mathbb{R}$. Then, if $x < 0$, then $|x| = -x$, and

$$x' = 3(-x)^{2/3}$$
$$\Leftrightarrow \frac{1}{3}(-x)^{-2/3}(-x') = -1$$
$$\Leftrightarrow (-x)^{1/3} = -t + k_2$$
$$\Leftrightarrow x(t) = -(-t + k_2)^3,$$

for $k_2 \in \mathbb{R}$. We can now use these computations with the different cases that were discussed earlier, depending on the value of t_0 and x_0. We begin with the case of $t_0 > 0$ and $x_0 > 0$.

1. The case $E = \emptyset$ is impossible, for all initial conditions (t_0, x_0). Indeed, as $x_0 > 0$, we have $x(t) = (t + k_1)^3$. Using the initial condition, we find that $x(t_0) = x_0 = (t_0 + k_1)^3$, i.e., $k_1 = x_0^{1/3} - t_0$, and $x(t) = \left(t + x_0^{1/3} - t_0\right)^3$.

2. If $E = [a, b]$, then

$$x(t) = \begin{cases} -(-t + k_2)^3 \text{ if } t < a \\ 0 \text{ if } t \in [a, b] \\ (t + k_1)^3 \text{ if } t > b. \end{cases}$$

Since $x_0 > 0$, we have to be in the $t > b$ region, so $t_0 > b$, and $(t_0 + k_1)^3 = x_0$, which implies that $k_1 = x_0^{1/3} - t_0$. Thus

$$x(t) = \begin{cases} -(-t + k_2)^3 \text{ if } t < a \\ 0 \text{ if } t \in [a, b] \\ \left(t + x_0^{1/3} - t_0\right)^3 \text{ if } t > b. \end{cases}$$

Since x is continuous,

$$\lim_{t \to b, t > b} \left(t + x_0^{1/3} - t_0\right)^3 = 0$$

and

$$\lim_{t \to a, t < a} -(-t + k_2)^3 = 0.$$

This implies that $b = t_0 - x_0^{1/3}$ and $k_2 = a$. So finally,

$$x(t) = \begin{cases} -(-t + a)^3 \text{ if } t < a \\ 0 \text{ if } t \in \left[a, t_0 - x_0^{\frac{1}{3}}\right] \quad \left(a \leq t_0 - x_0^{\frac{1}{3}}\right) \\ \left(t + x_0^{1/3} - t_0\right)^3 \text{ if } t > t_0 - x_0^{\frac{1}{3}}. \end{cases}$$

Thus, choosing $a \leq t_0 - x_0^{1/3}$, *we have solutions of the form shown in Figure 2.2. Indeed, any* a_i *satisfying this property yields a solution.*

3. The case $[a, +\infty)$ *is impossible. Indeed, there does not exist a solution through* (t_0, x_0) *such that* $x(t) = 0$ *for all* $t \in [a, +\infty)$*; since we are in the case of monotone increasing functions, if* $x_0 > 0$ *then* $x(t) \geq x_0$ *for all* $t \geq t_0$.

4. $E = \mathbb{R}$ *is also impossible, for the same reason.*

5. For the case $E = (-\infty, b]$, *we have*

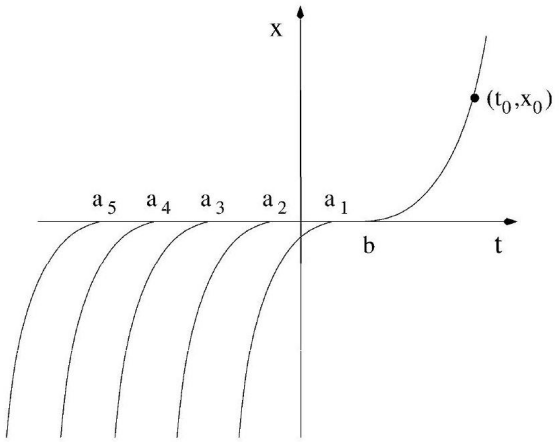


Figure 2.2: Case $t_0, x_0 > 0$, subcase 2 , in the resolution of (2.2.13).

$$x(t) = \begin{cases} 0 \text{ if } t \in (-\infty, b] \\ (t + k_1)^3 \text{ if } t > b. \end{cases}$$

Since $x(t_0) = x_0, k_1 = x_0^{1/3} - t_0$, *and since* x *is continuous,* $b = -k_1 = t_0 - x_0^{1/3}$. *So,*

$$x(t) = \begin{cases} 0 \text{ if } t \in \left(-\infty, t_0 - x_0^{1/3}\right] \\ \left(t + x_0^{1/3} - t_0\right)^3 \text{ if } t > t_0 - x_0^{1/3}. \end{cases}$$

The other cases are left as an exercise.

Example 2.2.3. *Consider the IVP*

$$\begin{cases} x' = 2tx^2 \\ x(0) = 0. \end{cases} \tag{2.2.14}$$

Here, we have existence and uniqueness of the solutions to (2.2.14). Indeed, $f(t,x) = 2tx^2$ *is continuous and locally Lipschitz on* $\mathbb{R}$.

2.3 Continuation of solutions

The results obtained so far concern the local existence (and uniqueness) of solutions to an initial value problem (IVP), meaning that solutions are guaranteed to exist within a neighborhood of the initial data. The study of solution continuation focuses on identifying criteria that enable the extension of solutions to potentially larger intervals.

Consider the IVP

$$\begin{cases} x' = f(t,x) \\ x(t_0) = x_0, \end{cases} \tag{2.3.15}$$

with f continuous on a domain $\mathcal{U}$ of the (t,x) space, and the initial point $(t_0,x_0) \in \mathcal{U}$.

Lemma 2.3.1. *Let* $f(t,x)$ *be a continuous function defined in an open set* $\mathcal{U}$ *in the* (t,x)*-space. Suppose that a function* $\phi(t)$ *satisfies the differential equation*

$$\phi'(t) = f(t,\phi(t)),$$

and that $(t,\phi(t)) \in \mathcal{U}$ *for all* t *in an open interval* $\mathcal{I} = (t_1,t_2)$.
Under these assumptions, if there exists a sequence $\{\tau_j\}_{j=1}^{\infty} \subset \mathcal{I}$ *such that*

$$\lim_{j\to\infty}(\tau_j,\phi(\tau_j)) = (t_1,\eta) \in \mathcal{U},$$

then it follows that

$$\lim_{\tau\to t_1}(\tau,\phi(\tau)) = (t_1,\eta).$$

Similarly, if there exists a sequence $\{\tau_j\}_{j=1}^{\infty} \subset \mathcal{I}$ *such that*

$$\lim_{j\to\infty}(\tau_j,\phi(\tau_j)) = (t_2,\eta) \in \mathcal{U},$$

then we have

$$\lim_{\tau\to t_2}(\tau,\phi(\tau)) = (t_2,\eta).$$

Proof. For a proof, see [1]. □

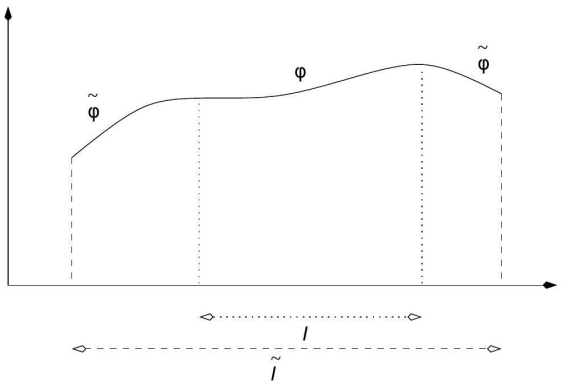


Figure 2.3: The extension $\tilde{\phi}$ on the interval $\tilde{\mathcal{I}}$ of the solution ϕ (defined on the interval $\mathcal{I}$).

From the previous result, we can deduce a statement regarding the maximal interval over which a solution can be extended. To emphasize that the solution ϕ of a differential equation exists within a specific interval $\mathcal{I}$, we denote it as $(\phi, \mathcal{I})$.
To proceed, we introduce the notion of an extension of a solution, which is defined in the classical sense (see Figure 2.3).

Definition 2.3.1. *(Extension). Let $(\phi, \mathcal{I})$ and $(\tilde{\phi}, \tilde{\mathcal{I}})$ be two solutions of the same ODE. We say that $(\tilde{\phi}, \tilde{\mathcal{I}})$ is an extension of $(\phi, \mathcal{I})$ if, and only if,*

$$\mathcal{I} \subset \tilde{\mathcal{I}}, \quad \tilde{\phi}_{|\mathcal{I}} = \phi,$$

where $_{|\mathcal{I}}$ denotes the restriction to $\mathcal{I}$.

Theorem 2.3.1. *Let $f(t, x)$ be continuous in an open set $\mathcal{U}$ in (t, x)-space, and the function $\phi(t)$ be a function satisfying the condition $\phi'(t) = f(t, \phi(t))$ and $(t, \phi(t)) \in \mathcal{U}$, in an open interval $\mathcal{I} = \{t_1 < t < t_2\}$. If the following two conditions are satisfied:*
(i) *$\phi(t)$ cannot be extended to the left of t_1 (or, respectively, to the right of t_2),*
(ii) *$\lim_{j\to\infty} (\tau_j, \phi(\tau_j)) = (t_1, \eta)$ (or, respectively, (t_2, η)) exists for some sequence $\{\tau_j : j = 1, 2, \ldots\}$ of points in the interval $\mathcal{I}$,*
then the limit point (t_1, η) (or, respectively, (t_2, η)) must be on the boundary of $\mathcal{U}$.

Proof. Suppose that the hypotheses of the theorem are satisfied, and that $(t_1, \eta) \in \mathcal{U}$

(respectively, $(t_2, \eta) \in \mathcal{U}$). Then, from Lemma 2.3.1, it follows that

$$\lim_{\tau \to t_1} (\tau, \phi(\tau)) = (t_1, \eta)$$

(or, respectively, $\lim_{\tau \to t_2} (\tau, \phi(\tau)) = (t_2, \eta)$). Thus we can apply Theorem 2.2.6 (Peano's Theorem) to the IVP

$$\begin{cases} x' = f(t, x) \\ x(t_1) = \eta, \end{cases}$$

(or, respectively, $x' = f(t, x), x(t_2) = \eta$). This implies that the solution ϕ can be extended to the left of t_1 (respectively, to the right of t_2), since Theorem 2.2.6 implies existence in a neighborhood of t_1. This is a contradiction. □

A particularly important consequence of the previous theorem is the following corollary.

Corollary 2.3.1. *Assume that $f(t, x)$ is continuous for $t_1 < t < t_2$ and all $x \in \mathbb{R}^n$. Also, assume that there exists a function $\phi(t)$ satisfying the following conditions:*
1. *ϕ and ϕ' are continuous in a subinterval $\mathcal{I}$ of the interval $t_1 < t < t_2$,*
2. *$\phi'(t) = f(t, \phi(t))$ in $\mathcal{I}$. Then, either*
(i) *$\phi(t)$ can be extended to the entire interval $t_1 < t < t_2$ as a solution of the differential equation $x' = f(t, x)$, or*
(ii) *$\lim_{t \to \tau} \|\phi(t)\| = \infty$ for some τ in the interval $t_1 < t < t_2$.*

2.3.1 Maximal interval of existence

Another way to express these results is through the concept of the maximal interval of existence. Given the differential equation

$$x' = f(t, x). \tag{2.3.16}$$

Let $x = x(t)$ be a solution of (2.3.16) on an interval $\mathcal{I}$.

Definition 2.3.2. *(Right maximal interval of existence). The interval $\mathcal{I}$ is said to be a right maximal interval of existence for x if there is no extension of $x(t)$ over an interval $\mathcal{I}_1$ such that x remains a solution of equation (2.3.16), with $\mathcal{I} \subset \mathcal{I}_1$ and*

$\mathcal{I}$ and $\mathcal{I}_1$ having different right endpoints. A left maximal interval of existence is defined similarly.

Definition 2.3.3. *(Maximal interval of existence). An interval that is both a left and a right maximal interval of existence is called a maximal interval of existence.*

Theorem 2.3.2. *Let $f(t,x)$ be continuous on an open set $\mathcal{U}$, and let $\phi(t)$ be a solution of (2.3.16) on some interval. Then $\phi(t)$ can be extended (as a solution) over a maximal interval of existence (ω_-,ω_+). Moreover, if (ω_-,ω_+) is a maximal interval of existence, then $\phi(t)$ tends to the boundary $\partial\mathcal{U}$ of $\mathcal{U}$ as $t\to\omega_-$ and $t\to\omega_+$.*

Remark 2.3.1. *The extension of the solution need not be unique, and the values of $\omega_\pm$ depend on the specific extension. Moreover, to say that $\phi\to\partial\mathcal{U}$ as $t\to\omega_+$ means that either $\omega_+=\infty$, or that $\omega_+<\infty$ and, for any compact subset $\mathcal{U}^0$ of $\mathcal{U}$, there exists a t near ω_+ such that $(t,\phi(t))\notin\mathcal{U}^0$.*

The following two corollaries are stated in [7].

Corollary 2.3.2. *Let $f(t,x)$ be continuous on the strip $t_0\le t\le t_0+a$ (with $a<\infty$) for all $x\in\mathbb{R}^n$. Suppose that ϕ is a solution of (2.1.3) on a right maximal interval J. Then one of the following holds:*

(i) $J=[t_0,t_0+a]$.

(ii) $J=[t_0,\delta)$, *with* $\delta\le t_0+a$, *and* $\|\phi(t)\|\to\infty$ *as* $t\to\delta$.

Corollary 2.3.3. *Let $f(t,x)$ be continuous on the closure $\overline{\mathcal{U}}$ of an open set $\mathcal{U}$ in the (t,x)-space, and let (2.1.3) admit a solution ϕ on a right maximal interval J. Then one of the following holds:*

(i) $J=[t_0,\infty)$.

(ii) $J=[t_0,\delta)$, *with* $\delta<\infty$ *and* $(\delta,\phi(\delta))\in\partial\mathcal{U}$.

(iii) $J=[t_0,\delta)$, *with* $\delta<\infty$ *and* $\|\phi(t)\|\to\infty$ *as* $t\to\delta$.

2.3.2 Maximal and global solutions

The concept of maximal intervals of existence of solutions is closely related to the notions of maximal and global solutions.

Definition 2.3.4. *(Maximal solution). Let $\mathcal{I}_1 \subset \mathbb{R}$ and $\mathcal{I}_2 \subset \mathbb{R}$ be two intervals satisfying $\mathcal{I}_1 \subset \mathcal{I}_2$. A solution $(\phi, \mathcal{I}_1)$ is said to be maximal in $\mathcal{I}_2$ if there exists no extension $(\tilde{\phi}, \tilde{\mathcal{I}})$ that remains a solution of the ODE and satisfies $\mathcal{I}_1 \subsetneq \tilde{\mathcal{I}} \subset \mathcal{I}_2$.*

Definition 2.3.5. *(Global solution). A solution $(\phi, \mathcal{I}_1)$ is said to be global on $\mathcal{I}_2$ if there exists an extension $\tilde{\phi}$ that remains a solution of the ODE and is defined on the entire interval $\mathcal{I}_2$.*

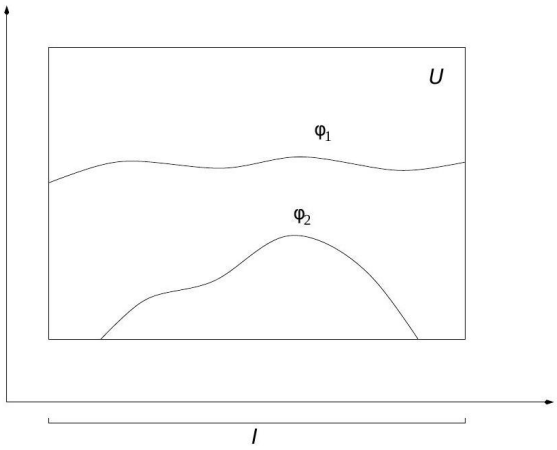


Figure 2.4: ϕ_1 is a global and maximal solution on $\mathcal{I}$; ϕ_2 is a maximal solution on $\mathcal{I}$, but it is not global on $\mathcal{I}$.

Every global solution on a given interval $\mathcal{I}$ is also a maximal solution on $\mathcal{I}$. However, the converse does not necessarily hold.

Example 2.3.1. *Consider the differential equation*

$$x' = -2tx^2$$

defined on $\mathbb{R}$. If $x \neq 0$, integrating the equation

$$x'x^{-2} = -2t,$$

leads to the general solution

$$x(t) = \frac{1}{t^2 - c}, \quad c \in \mathbb{R}.$$

Depending on the value of c, we distinguish the following cases:

- *If $c < 0$, then $x(t)$ is defined for all $t \in \mathbb{R}$, making it a global solution.*
- *If $c > 0$, the denominator $t^2 - c$ vanishes at $t = \pm\sqrt{c}$, causing singularities. The solutions are then defined on the intervals $(-\infty, -\sqrt{c})$, $(-\sqrt{c}, \sqrt{c})$, and $(\sqrt{c}, \infty)$. Each of these solutions is maximal on its domain but not global on $\mathbb{R}$.*
- *If $c = 0$, the function simplifies to $x(t) = 1/t^2$, which is defined on $(-\infty, 0)$ and $(0, \infty)$, leading to maximal but non-global solutions on $\mathbb{R}$.*

Additionally, the constant function $x \equiv 0$ is a global solution on $\mathbb{R}$.

Lemma 2.3.2. *Every solution ϕ of the differential equation $x' = f(t, x)$ can be extended to a maximal solution $\tilde{\phi}$, meaning that ϕ is always contained within some maximal solution.*

The following theorem extends the uniqueness property to an interval of existence of the solution.

Theorem 2.3.3. *Let $\phi_1, \phi_2 : \mathcal{I} \to \mathbb{R}^n$ be two solutions of the equation $x' = f(t, x)$, with f being locally Lipschitz in x on $\mathcal{U}$. If ϕ_1 and ϕ_2 coincide at a point $t_0 \in \mathcal{I}$, then $\phi_1 = \phi_2$ on the entire interval $\mathcal{I}$.*

Proof. Under the assumptions of the theorem, suppose that $\phi_1(t_0) = \phi_2(t_0)$. Let us assume, for simplicity, that there exists a $t_1 \neq t_0$ such that $\phi_1(t_1) \neq \phi_2(t_1)$, with $t_1 > t_0$.
By the local uniqueness of the solution, it follows from $\phi_1(t_0) = \phi_2(t_0)$ that there exists a neighborhood $\mathcal{N}$ of t_0 such that $\phi_1(t) = \phi_2(t)$ for all $t \in \mathcal{N}$. Let us define the set

$$E = \{t \in [t_0, t_1] : \phi_1(t) \neq \phi_2(t)\}.$$

Since $t_1 \in E$, we have $E \neq \emptyset$. Let $\alpha = \inf(E)$, where $\alpha \in (t_0, t_1]$. For all $t \in [t_0, \alpha)$, we have $\phi_1(t) = \phi_2(t)$.

By the continuity of ϕ_1 and ϕ_2, it follows that $\phi_1(\alpha) = \phi_2(\alpha)$. This implies there exists a neighborhood $\mathcal{W}$ of α where $\phi_1(t) = \phi_2(t)$. However, this contradicts the assumption that $\phi_1(t) \neq \phi_2(t)$ for $t > \alpha$. Thus, there can be no such t_1, and therefore $\phi_1 = \phi_2$ on the entire interval $\mathcal{I}$. □

Corollary 2.3.4. *(Global uniqueness). Let $f(t, x)$ be locally Lipschitz in x on $\mathcal{U}$. Then, for any point $(t_0, x_0) \in \mathcal{U}$, there exists a unique maximal solution $\phi : \mathcal{I} \to \mathbb{R}^n$. Furthermore, if there exists a global solution on $\mathcal{I}$, it is unique.*

2.4 Exercises

Exercice 2.4.1. *Consider the initial value problem (IVP):*

$$y' + 2y = e^{-x}, \quad y(0) = 1.$$

Find the explicit solution.

Exercice 2.4.2. *Verify that $y = x^2 + 2x + 1$ is a solution to the differential equation:*

$$y' = 2x + 2.$$

Exercice 2.4.3. *Consider the problem*

$$y' = 1 - 2xy, \quad y(0) = 0.$$

(a) Since the differential equation is linear, an expression can be found for the solution. Find it.

(b) Consider the above problem on the region

$$R: \quad |x| \leq \frac{1}{2}, \quad |y| \leq 1.$$

If $f(x, y) = 1 - 2xy$, show that

$$|f(x, y)| \leq 2, \quad \text{for } (x, y) \in R,$$

and that all the successive approximations to the solution exist on $|x| \leq \frac{1}{2}$*, and their graphs remain in* R*.*

(c) Show that f *satisfies a Lipschitz condition on* R *with Lipschitz constant* $K = 1$*, and therefore by Theorem 2.2.5, the successive approximations converge to a solution* ϕ *of the initial value problem on* $|x| \leq \frac{1}{2}$*.*

(d) Show that the approximation ϕ_3 *satisfies*

$$|\phi(x) - \phi_3(x)| < 0.01 \quad for\ |x| \leq \frac{1}{4}.$$

(e) Compute ϕ_3*.*

Exercice 2.4.4. *Consider the initial value problem (IVP):*

$$y' = y^2, \quad y(0) = 1.$$

Determine the interval of existence for the solution.

Exercice 2.4.5. *Consider the initial value problem:*

$$y' = \frac{y}{x}, \quad y(1) = 2.$$

Find the solution and determine its maximal interval of existence.

Chapter 3

Stability in Linear and Nonlinear Systems

In this chapter, we study the stability of solutions to differential equations, focusing on the behavior of systems near their equilibrium points. Stability analysis is crucial in understanding how systems respond to small disturbances. Stability can provide insight into the long-term behavior of equilibrium states and periodic solutions for both linear and nonlinear systems.

3.1 Linear stability analysis

Consider the linear differential system

$$\dot{x} = f(x), \tag{3.1.1}$$

where $f : D \longrightarrow \mathbb{R}^n$, a C^1 function and D is an open subset of $\mathbb{R}^n$.

We can also define a linear differential system by

$$\dot{x} = Ax + b, \tag{3.1.2}$$

where A is a constant $n \times n$ matrix and b is a constant vector in $\mathbb{R}^n$.

3.1.1 Equilibria

Equilibria are points in the phase space where the system does not change over time. Stability describes the system's behavior in response to small perturbations near these equilibrium points.

Definition 3.1.1. *In a dynamical system defined by the equation (3.1.1), an equilibrium point x^* is a point where $f(x^*) = 0$.*

Remark 3.1.1. *An equilibrium point is often referred to as a singular point, fixed point, or critical point.*

Remark 3.1.2. *If the point x^* is an equilibrium point for the system (3.1.2) then, with the change of variables $y = x - x^*$, (3.1.2) can be rewritten as*

$$\dot{y} = Ay. \tag{3.1.3}$$

Notice that $y = 0$ is an equilibrium point for the equation (3.1.3).

3.1.2 Stability of Equilibria

Definition 3.1.2. *We define the stability of the equilibrium point x^* as follows*

- *The equilibrium point x^* is stable if,*

$$\forall \epsilon > 0, \quad \exists \delta(\epsilon) > 0 \mid \left\| \overrightarrow{x(0) - x^*} \right\| < \delta \implies \left\| \overrightarrow{x(t) - x^*} \right\| < \epsilon, \ \forall t \geq 0.$$

- *The equilibrium point x^* is asymptotically stable if it is stable and, $\left\| \overrightarrow{x(t) - x^*} \right\| \to 0$ as $t \to \infty$.*

- *The equilibrium point x^* is unstable if it is not stable.*

Remark 3.1.3. *Asymptotic stability indicates that the trajectories of the system converge to the equilibrium point as $t \to \infty$. In contrast, neutral stability (stable but not asymptotically stable) means that while the trajectories remain near the equilibrium point, they do not necessarily converge to it over time.*

Consider the linear system (3.1.3), where A is a constant $n \times n$ matrix, and let $y = 0$ be an equilibrium point. The behavior near an equilibrium point (at the origin in this case) is determined by the eigenvalues of A. Specifically, the characteristic equation $\det(A - \lambda I) = 0$ gives the eigenvalues λ_i of A ,for $i = 1, \cdots, n$.

Theorem 3.1.1.

- *If all eigenvalues of A have negative real parts, then the equilibrium $y = 0$ is asymptotically stable.*
- *If at least one eigenvalue of A has a positive real part, then the equilibrium $y = 0$ is unstable.*
- *If all eigenvalues of A have non-positive real parts and at least one eigenvalue has a real part of zero, the stability of $y = 0$ cannot be determined solely from the eigenvalues of A and requires further analysis.*

Proof. The general solution to the system $\dot{y} = Ay$ is given by

$$y(t) = e^{At}y(0),$$

where e^{At} denotes the matrix exponential of A, and $y(0)$ is the initial condition.
To analyze the stability of $y = 0$, we examine the behavior of $\|y(t)\|$ as $t \to \infty$. The properties of e^{At} depend on the eigenvalues of A.
1. **Case 1: All eigenvalues of A have negative real parts.**
If all eigenvalues of A have negative real parts, then e^{At} decays to zero as $t \to \infty$. Specifically, since the real parts of the eigenvalues determine the rate of decay, we have:

$$\|e^{At}\| \leq Ce^{-\alpha t},$$

for some constants $C > 0$ and $\alpha > 0$ (where α is related to the smallest magnitude of the negative real parts of the eigenvalues). It follows that

$$\|y(t)\| = \|e^{At}y(0)\| \leq Ce^{-\alpha t}\|y(0)\| \to 0 \quad \text{as } t \to \infty.$$

Thus, $y = 0$ is **asymptotically stable**.
2. **Case 2: At least one eigenvalue of A has a positive real part.**
If A has at least one eigenvalue with a positive real part, there exists a component of the solution that grows exponentially as $t \to \infty$. Let λ be an eigenvalue of A with

$\text{Re}(\lambda) > 0$, and let v be the corresponding eigenvector. Then the term $e^{\lambda t}v$ appears as part of the solution $y(t)$, resulting in $\|y(t)\| \to \infty \quad \text{as } t \to \infty$. Hence, $y = 0$ is **unstable**.

3. **Case 3: All eigenvalues of A have non-positive real parts, with at least one zero real part.**

If all eigenvalues have non-positive real parts and at least one eigenvalue has zero real part, the stability analysis is more subtle. The matrix A may have both decaying and non-decaying modes. In this case:

- If A is **diagonalizable**, the solution can be expressed as a combination of terms that either decay (from eigenvalues with negative real parts) or remain constant (from eigenvalues with zero real parts). Thus, $y = 0$ is **stable** but **not asymptotically stable**, as certain components of $y(t)$ do not decay to zero but remain bounded.

- If A is **not diagonalizable**, the matrix e^{At} may exhibit terms that grow linearly in t, due to terms involving $te^{\lambda t}$ when $\lambda = 0$. In this situation, the solution $y(t)$ can grow unboundedly as $t \to \infty$, rendering $y = 0$ **unstable**. This completes the proof. □

Remark 3.1.4. *To analyze the stability of the equilibrium point x^* in the original system $\dot{x} = Ax + b$, we equivalently study the stability of $y = 0$ in the shifted system $\dot{y} = Ay$, where $y = x - x^*$. In this approach, by shifting the system to center around the equilibrium x^*, we reduce the analysis to examining the stability of the zero equilibrium in the transformed system, simplifying the calculations.*

3.1.3 Classification of fixed points in $2D$

Consider the linear differential system (3.1.3), where A is a 2×2 matrix. The nature of fixed points (equilibria) can be classified based on the eigenvalues of the matrix A. The eigenvalues λ_1 and λ_2 of A can lead to different types of fixed points:

- **Nodes**: Both eigenvalues are real and have the same sign i.e. $\lambda_1\lambda_2 > 0$ and $\text{Im}(\lambda_{1,2}) = 0$. If negative, the node is a stable (or attracting) node; if positive, it is an unstable (or repelling) node.

- **Saddles**: Eigenvalues are real with opposite signs i.e. $\lambda_1\lambda_2 < 0$ and $\text{Im}(\lambda_{1,2}) = 0$. Saddles are always unstable.

- **Spirals (Foci)**: Eigenvalues are complex conjugates with non-zero real parts i.e. $\mathrm{Re}(\lambda_{1,2}) \neq 0$. If the real parts are negative, it is a stable spiral (focus); if positive, it is an unstable spiral.

- **Centers**: Eigenvalues are purely imaginary i.e. $\mathrm{Re}(\lambda_{1,2}) = 0$, leading to closed orbits around the equilibrium. Centers are typically stable but not asymptotically stable.

Example 3.1.1. *Consider the dynamical system given by:*

$$\dot{x} = \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix} x.$$

To analyze the stability of the equilibrium point at the origin, we first determine the characteristic equation. This is achieved by computing the determinant of the matrix obtained by subtracting λ (the eigenvalue) from the diagonal entries:

$$\det \begin{pmatrix} -2-\lambda & 1 \\ 1 & -3-\lambda \end{pmatrix} = 0.$$

Calculating the determinant, we have:

$$(-2-\lambda)(-3-\lambda) - (1)(1) = 0.$$

Expanding this yields:

$$(\lambda+2)(\lambda+3) - 1 = 0.$$

Simplifying, we get:

$$\lambda^2 + 5\lambda + 5 = 0.$$

Next, we can solve this quadratic equation using the quadratic formula:

$$\lambda = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{-5 \pm \sqrt{25-20}}{2} = \frac{-5 \pm \sqrt{5}}{2}.$$

This leads to the eigenvalues:

$$\lambda_1 = -1 \quad \textit{and} \quad \lambda_2 = -4.$$

Since both eigenvalues are negative ($\lambda_1 < 0$ and $\lambda_2 < 0$), we conclude that the equilibrium point at the origin is a stable node.

3.2 Nonlinear stability analysis

Consider the nonlinear differential system

$$\dot{x} = f(x), \tag{3.2.4}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a sufficiently smooth function.

Definition 3.2.1. *(Flow of a nonlinear differential system). The flow associated with system (3.2.4) is a function*

$$\varphi : D \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n,$$

that assigns to each initial condition x_0 and time t the unique solution $\varphi(t, x_0)$ of (3.2.4) satisfying the initial condition

$$\varphi(0, x_0) = x_0.$$

If the solution exists for all t in some interval containing zero, the function $\varphi(t, x_0)$ describes the trajectory of the system starting from x_0. The flow satisfies the semi-group property:

$$\varphi(t + s, x_0) = \varphi(t, \varphi(s, x_0)),$$

for all t, s where the flow is defined.

Definition 3.2.2. *An equilibrium point x_0 is called a hyperbolic equilibrium point of (3.2.4) if none of the eigenvalues of the matrix $Df(x_0)$ have zero real part, where $Df(x_0)$ is the $n \times n$ Jacobian matrix of (3.2.4) at x_0.*

Definition 3.2.3. *The linear system*

$$\dot{x} = Ax, \tag{3.2.5}$$

with the matrix $A = Df(x_0)$ is called the linearized system of (3.2.4) at x_0.

Definition 3.2.4. *Two autonomous systems of differential equations, such as (3.2.4) and (3.2.5) are said to be topologically equivalent in a neighborhood of the origin, or to have the same qualitative structure near the origin, if there exists a homeomorphism H mapping an open set U containing the origin onto an open set V containing the origin, such that H maps trajectories of (3.2.4) in U onto trajectories of (3.2.7) in V while preserving their orientation in time. That is, if a trajectory is directed from x_1 to x_2 in U, then its image is directed from $H(x_1)$ to $H(x_2)$ in V.*
If, in addition, the homeomorphism H preserves the parameterization by time, then the systems (3.2.4) and (3.2.5) are said to be topologically conjugate in a neighborhood of the origin.

In the following, we present a simple example of two topologically conjugate linear systems.

Example 3.2.1. *Consider the linear systems*

$$\dot{x} = Ax, \quad \dot{y} = By, \tag{3.2.6}$$

where the matrices A and B are given by

$$A = \begin{bmatrix} -1 & -3 \\ -3 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}.$$

Let $H(x) = Rx$, where the matrix R is

$$R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad R^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

It follows that $B = RAR^{-1}$ and setting $y = H(x) = Rx$ and equivalently $x = R^{-1}y$, we obtain

$$\dot{y} = RAR^{-1}y = By.$$

Thus, if $x(t) = e^{At}x_0$ is the solution of the first system through x_0, then

$$y(t) = H(x(t)) = Rx(t) = Re^{At}x_0 = e^{Bt}Rx_0,$$

is the solution of the second system through Rx_0. That is, H maps trajectories of

the first system onto trajectories of the second system while preserving their parameterization, since

$$He^{At} = e^{Bt}H.$$

The mapping $H(x) = Rx$ corresponds to a rotation by 45° and is clearly a homeomorphism.

The following theorem is a fundamental result in the local qualitative theory of ordinary differential equations. It establishes that near a hyperbolic equilibrium point x_0, the nonlinear system (3.2.4) exhibits the same qualitative behavior as the corresponding linear system

$$\dot{x} = Ax, \tag{3.2.7}$$

where $A = Df(x_0)$.
In what follows, we assume that the equilibrium point x_0 has been translated to the origin.

Theorem 3.2.1. *(The Hartman-Grobman Theorem). Let E be an open subset of $\mathbb{R}^n$ containing the origin, let $f \in C^1(E)$, and let φ_t be the flow of the nonlinear system (3.2.4). Suppose that $f(0) = 0$ and that the matrix $A = Df(0)$ has no eigenvalue with zero real part. Then there exists a homeomorphism H of an open set U containing the origin onto an open set V containing the origin such that for each $x_0 \in U$, there is an open interval $I_0 \subset \mathbb{R}$ containing zero, such that for all $x_0 \in U$ and $t \in I_0$,*

$$H \circ \varphi_t(x_0) = e^{At}H(x_0).$$

That is, H maps trajectories of the nonlinear system (3.2.4) near the origin onto trajectories of the linear system (3.2.7) near the origin and preserves the parameterization by time.

Proof. See ([12], p. 121).

3.2.1 Linearization Around Equilibria

To analyze the stability of nonlinear systems, one common approach is to linearize the system around its equilibrium points. Consider a nonlinear system described by

the ordinary differential equation:

$$\dot{x} = f(x), \tag{3.2.8}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function. An equilibrium point x^* satisfies the condition $f(x^*) = 0$.
To investigate the stability, we perform a Taylor expansion of $f(x)$ around the equilibrium point x^* given by

$$f(x) = f(x^*) + Df(x^*)(x - x^*) + \mathcal{O}(\|x - x^*\|^2), \tag{3.2.9}$$

where $Df(x^*)$ is the Jacobian matrix of f evaluated at the equilibrium point x^*. Since $f(x^*) = 0$, we have

$$f(x) \approx Df(x^*)(x - x^*). \tag{3.2.10}$$

Thus, the linearized system can be expressed as

$$\dot{x} = Df(x^*)(x - x^*). \tag{3.2.11}$$

Letting $y = x - x^*$, we rewrite the system as

$$\dot{y} = Df(x^*)y. \tag{3.2.12}$$

In a similar way to the linear case, the stability of the equilibrium point x^* is determined by the eigenvalues of the Jacobian matrix $Df(x^*)$ as follows:

- If all eigenvalues have negative real parts, then x^* is asymptotically stable.
- If at least one eigenvalue has a positive real part, then x^* is unstable.
- If eigenvalues have zero real parts, the linearization does not provide conclusive information about stability, and further analysis is required.

Remark 3.2.1. *This linearization method is powerful and widely used for determining local stability properties of equilibrium points in nonlinear dynamical systems. However, it is important to note that the conclusions drawn from linearization are only valid in a neighborhood of the equilibrium point.*

3.2.2 Stability and Lyapunov's method

The stability of an equilibrium point x_0 in a dynamical system (3.2.4) depends significantly on the nature of the eigenvalues of the Jacobian matrix $Df(x_0)$ at that point. For hyperbolic equilibrium points —where none of the eigenvalues have zero real parts— stability is straightforward to assess: if all eigenvalues of $Df(x_0)$ have negative real parts, then x_0 is a **sink**, making it asymptotically stable. This means that trajectories close to x_0 will converge to it over time. In contrast, if any eigenvalue has a positive real part, the equilibrium x_0 becomes unstable. Such instability is observed in points classified as **sources** (all real parts are positive) or **saddles** (a mix of positive and negative real parts among the eigenvalues).
When dealing with nonhyperbolic equilibrium points—those with at least one eigenvalue whose real part is zero—the stability analysis becomes more complex. In these cases, linearization fails to provide a definitive answer, and other techniques are necessary. One powerful approach is Lyapunov's method, which relies on constructing a special function, called a Lyapunov function. This function is designed to decrease along trajectories, indicating stability if it satisfies certain positivity conditions. Lyapunov's method is particularly effective in cases where traditional eigenvalue analysis is inconclusive, offering insight into the stability of more challenging nonhyperbolic points.

Definition 3.2.5. *Let φ_t represent the flow of the differential equation (3.2.4), defined for all $t \in \mathbb{R}$. An equilibrium point x_0 of (3.2.4) is stable if, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that for every point $x \in N_\delta(x_0)$ (a neighborhood of x_0) and for all $t \geq 0$, the flow satisfies $\varphi_t(x) \in N_\varepsilon(x_0)$.*
The equilibrium point x_0 is unstable if it is not stable. Additionally, x_0 is asymptotically stable if it is stable and there exists a $\delta > 0$ such that for any $x \in N_\delta(x_0)$, we have $\lim_{t\to\infty} \varphi_t(x) = x_0$, meaning that nearby trajectories approach x_0 as time progresses.

Equilibrium points in a linear system in $\mathbb{R}^2$ exhibit distinct stability properties. A stable node or focus is an asymptotically stable equilibrium, meaning nearby trajectories converge to this point over time. Conversely, an unstable node, focus, or saddle is inherently unstable, as trajectories diverge from these points. Meanwhile, a center is stable but not asymptotically stable, with trajectories forming closed

orbits around it instead of converging. By the Hartman-Grobman Theorem, any sink in the system (3.2.4) is asymptotically stable, while any source or saddle is unstable. Consequently, for any hyperbolic equilibrium point of (3.2.4), stability is characterized as either asymptotically stable or unstable.
The following theorem provides more information about the local dynamics of trajectories near a sink.

Theorem 3.2.2. *If x_0 is a sink of the nonlinear system (3.2.4) and $\mathrm{Re}(\lambda_j) < -\alpha < 0$ for all of the eigenvalues λ_j of the matrix $Df(x_0)$, then given $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in N_\delta(x_0)$, the flow $\varphi_t(x)$ of (3.2.4) satisfies*

$$|\varphi_t(x) - x_0| < \varepsilon e^{-\alpha t} \quad \textit{for all } t > 0.$$

Since hyperbolic equilibrium points are either asymptotically stable or unstable, an equilibrium point x_0 of (3.2.4) can only be stable, but not asymptotically stable, if $Df(x_0)$ has a zero eigenvalue or a pair of complex-conjugate, purely imaginary eigenvalues $\lambda = \pm i\beta$. The following theorem implies that for x_0 to be stable, all other eigenvalues λ_j of $Df(x_0)$ must satisfy $\mathrm{Re}(\lambda_j) \leq 0$.

Theorem 3.2.3. *If x_0 is a stable equilibrium point of (3.2.4), then no eigenvalue of $Df(x_0)$ has a positive real part.*

We conclude that stable equilibrium points that are not asymptotically stable can only exist at nonhyperbolic points. However, determining whether a nonhyperbolic equilibrium is stable, asymptotically stable, or unstable is a subtle issue. The following method, developed by Lyapunov in his 1892 doctoral thesis, is very useful in addressing this question.

Definition 3.2.6. *Let $V \in C^1(E)$ with E an open subset of $\mathbb{R}^n$, and φ_t is the flow of the differential equation (3.2.4), then for $x \in E$, the derivative of the function $V(x)$ along the solution $\varphi_t(x)$ is given by:*

$$\frac{d}{dt} V(\varphi_t(x))\Big|_{t=0} = DV(x)f(x).$$

A function $V : \mathbb{R}^n \to \mathbb{R}$ satisfying the hypotheses of the following theorem is called a Lyapunov function.

Theorem 3.2.4. *Let E be an open subset of $\mathbb{R}^n$ containing x_0. Suppose that $f \in C^1(E)$ and that $f(x_0) = 0$. Suppose further that there exists a real-valued function $V \in C^1(E)$ satisfying $V(x_0) = 0$ and $V(x) > 0$ if $x \neq x_0$. Then:*

1. *If $\dot{V}(x) \leq 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is stable.*

2. *If $\dot{V}(x) < 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is asymptotically stable.*

3. *If $\dot{V}(x) > 0$ for all $x \in E \setminus \{x_0\}$, then x_0 is unstable.*

Proof. Without loss of generality, assume the equilibrium point is $x_0 = 0$.

1. Choose $\varepsilon > 0$ small enough such that $N_\varepsilon(0) \subset E$. Let m_ε be the minimum of the continuous function $V(x)$ on the compact set $S_\varepsilon = \{x \in \mathbb{R}^n : |x| = \varepsilon\}$. Since $V(x) > 0$ for $x \neq 0$ (it follows that $m_\varepsilon > 0$) and $V(0) = 0$, there exists $\delta > 0$ such that for $|x| < \delta$, $V(x) < m_\varepsilon$. Since $V(x)$ is decreasing along trajectories of (3.2.4) (i.e $\dot{V}(x) \leq 0$), and if φ_t is the flow of the equation (3.2.4), then for all $x_0 \in N_\delta(0)$ and $t \geq 0$, we have

$$V(\varphi_t(x_0)) \leq V(x_0) < m_\varepsilon.$$

   Now assume that for $|x_0| < \delta$ there is a $t_1 > 0$ such that $|\varphi_{t_1}(x_0)| = \varepsilon$ i.e., such that $\varphi_{t_1}(x_0) \in S_\varepsilon$. Then since m_ε is the minimum of $V(x)$ on S_ε, this would lead to $V(\varphi_{t_1}(x_0)) \geq m_\varepsilon$ which contradicts the above inequality. Hence, for $|x_0| < \delta$ and $t \geq 0$ it follows that $\varphi_{t_1}(x_0) < \varepsilon$, i.e., 0 is a stable equilibrium point.

2. If $\dot{V}(x) < 0$ for all $x \in E$, $V(x)$ is strictly decreasing along trajectories of (3.2.4). Let φ_t be the flow of (3.2.4) and $x_0 \in N_\delta(0)$. By part (i), if $\|x_0\| < \delta$, then $\varphi_t(x_0) \subset N_\varepsilon(0)$ for all $t \geq 0$. Let $\{t_k\}$ be any sequence with $t_k \to \infty$. Since $\overline{N_\delta(0)}$ is compact, there is a subsequence of $\{\varphi_{t_k}(x_0)\}$ that converges to a point in $\overline{N_\delta(0)}$. However, for any subsequence $\{t_{k_j}\}$ of $\{t_k\}$ such that $\{\varphi_{t_{k_j}}(x_0)\}$ converges, we show below that the limit is zero. It follows that $\varphi_{t_k}(x_0) \to 0$ as $t_k \to \infty$, and thus $\varphi_t(x_0) \to 0$ as $t \to \infty$, i.e., 0 is asymptotically stable. It remains to show that if $\varphi_{t_j}(x_0) \to y_0$, then $y_0 = 0$. Since $V(x)$ is strictly decreasing along trajectories of (3.2.4) and since $V(\varphi_{t_j}(x_0)) \to V(y_0)$ by the

continuity of V, it follows that

$$V(\varphi_t(x_0)) > V(y_0),$$

for all $t > 0$. But if $y_0 \neq 0$, then for some $s > 0$, we have $V(\varphi_s(y_0)) < V(y_0)$, and by continuity, it follows that for all y sufficiently close to y_0, we have $V(\varphi_s(y)) < V(y_0)$ for $s > 0$. But then for $y = \varphi_{t_j} x_0$ and j sufficiently large, we would have

$$V(\varphi_{s+t_j}(x_0)) < V(y_0),$$

which contradicts the above inequality. Therefore, $y_0 = 0$, and it follows that 0 is asymptotically stable.

3. Let M be the maximum of $V(x)$ on the compact $\overline{N_\delta(0)}$. Since $\dot{V}(x) > 0$, $V(x)$ is strictly increasing along trajectories of (3.2.4). Hence, if φ_t is the flow of (3.2.4), then for any $\delta > 0$ and $x_0 \in N_\delta(0) \setminus \{0\}$, we have

$$V(\varphi_t(x_0)) > V(x_0) > 0,$$

for all $t > 0$. Since $V(x)$ is positive definite, this implies that

$$\inf_{t \geq 0} \dot{V}(\varphi_t(x_0)) = m > 0.$$

Thus, for all $t > 0$, we have

$$V(\varphi_t(x_0)) - V(x_0) \geq mt.$$

Therefore, for t sufficiently large,

$$V(\varphi_t(x_0)) > mt > M,$$

i.e., $\varphi_t(x_0)$ lies outside the closed set $\overline{N_\delta(0)}$. Hence, 0 is unstable.

□

Remark 3.2.2. *If $\dot{V}(x) = 0$ for all $x \in E$, then the trajectories of system (3.2.4) lie on the surfaces in $\mathbb{R}^n$ (or curves in $\mathbb{R}^2$) defined by $V(x) = c$, with c constant.*

Example 3.2.2. *Consider the system*

$$\begin{cases} \dot{x}_1 &= -x_2^3, \\ \dot{x}_2 &= x_1^3. \end{cases}$$

The origin is a nonhyperbolic equilibrium point of this system, and the function

$$V(x) = x_1^4 + x_2^4,$$

serves as a Lyapunov function for this system. Indeed, we have

$$\dot{V}(x) = 4x_1^3\dot{x}_1 + 4x_2^3\dot{x}_2 = 0.$$

Thus, the solution trajectories lie on the closed curves defined by

$$x_1^4 + x_2^4 = c^2,$$

which encircles the origin. This implies that the origin is a stable equilibrium point of the system, though it is not asymptotically stable. Note that in this example, $Df(0) = 0$*, meaning that* $Df(0)$ *has two zero eigenvalues.*

Example 3.2.3. *Consider the system*

$$\begin{cases} \dot{x}_1 &= -2x_2 + x_2x_3, \\ \dot{x}_2 &= x_1 - x_1x_3, \\ \dot{x}_3 &= x_1x_2. \end{cases}$$

The origin is an equilibrium point for this system, and the Jacobian at the origin is

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

This matrix has eigenvalues $\lambda_1 = 0$ *and* $\lambda_{2,3} = \pm 2i$*, so* $x = 0$ *is a nonhyperbolic equilibrium point. To assess stability, we use Lyapunov's method.*

We try a Lyapunov function of the form

$$V(x) = c_1x_1^2 + c_2x_2^2 + c_3x_3^2,$$

where c_1, c_2, and c_3 are positive constants. Calculating $\dot{V}(x) = DV(x)f(x)$, we find

$$\frac{1}{2}\dot{V}(x) = (c_1 - c_2 + c_3)x_1x_2x_3 + (-2c_1 + c_2)x_1x_2.$$

Setting $c_2 = 2c_1$ and $c_3 = c_1 > 0$ ensures $\dot{V}(x) > 0$ for $x \neq 0$ and $\dot{V}(x) = 0$ for all $x \in \mathbb{R}^3$. Therefore, by Theorem 3.2.4, the origin is stable.
Moreover, choosing $c_1 = c_3 = 1$ and $c_2 = 2$, we get that the trajectories of this system lie on ellipsoids given by

$$x_1^2 + 2x_2^2 + x_3^2 = c^2.$$

As we noted previously, all sinks are asymptotically stable. However, the following example shows that not all asymptotically stable equilibrium points are sinks. (A hyperbolic equilibrium point is asymptotically stable if and only if it is a sink.)

Example 3.2.4. *Consider the following modification of the system in the previous example*

$$\begin{cases} \dot{x}_1 &= -2x_2 + x_2x_3 - x_1^3, \\ \dot{x}_2 &= x_1 - x_1x_3 - x_2^3, \\ \dot{x}_3 &= x_1x_2 - x_3^3. \end{cases}$$

The Lyapunov function from the previous example,

$$V(x) = x_1^2 + 2x_2^2 + x_3^2,$$

satisfies $V(x) > 0$ and

$$\dot{V}(x) = -2(x_1^4 + 2x_2^4 + x_3^4) < 0,$$

for $x \neq 0$. Therefore, by Theorem 3.2.4, the origin is asymptotically stable. However, it is not a sink since the eigenvalues $\lambda_1 = 0$ and $\lambda_{2,3} = \pm 2i$ do not have negative real parts.

3.3 First integral

Consider a system of first-order ordinary differential equations in the unknown function $x : I \subset \mathbb{R} \to \mathbb{R}^n$, given by

$$\Phi(t, x(t), \dot{x}(t)) = 0, \tag{3.3.13}$$

where Φ is a smooth function. A first integral of the system is a continuously differentiable function $\Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ that is locally constant along any solution of (3.3.13), i.e.,

$$\frac{d}{dt}\Psi(t, x(t)) = 0,$$

for any $x : J \to \mathbb{R}^n$ that solves (3.3.13). The domain of definition of Ψ must be adjusted appropriately when the domain of Φ does not cover the entire space, and often, one considers Ψ defined in a smaller domain, typically in a neighborhood of some specific point.

Example 3.3.1. *Consider a scalar equation of the form*

$$\dot{x}(t) = f(t, x(t)),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$. It can be easily seen that F is a first integral if and only if F satisfies the partial differential equation

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f = 0. \tag{3.3.14}$$

A solution to equation (3.3.14) always exists in a neighborhood of a point where f is Lipschitz. In fact, if f is Lipschitz on $\mathbb{R} \times \mathbb{R}$, then there is a unique global solution to equation (3.3.14) subject to any initial condition of the Cauchy type $f(0, x) = f_0(x)$ for all x.

A particularly interesting case occurs when the system is completely integrable, i.e., when there are n functionally independent first integrals $\Psi_1, \ldots, \Psi_n$. This condition is equivalent to the existence of a general formula for the solutions of equation (3.3.13) in implicit form. The knowledge of n functionally independent first integrals

also ensures that any other first integral Λ must take the form

$$\Lambda(t,x) = F(\Psi_1(t,x),\dots,\Psi_n(t,x)).$$

First integrals of motion are particularly studied in the theory of Hamiltonian systems. For physically relevant cases, the first integrals are also called "constants of motion," and some of them correspond to conservation laws for physically relevant quantities. The primary example is the system of equations governing the motion of a particle in a potential field

$$\ddot{x}(t) = -\nabla U(x(t)). \tag{3.3.15}$$

Introducing the new variables $v(t) := \dot{x}(t)$, we can transform equation (3.3.15) into a first-order system:

$$\begin{cases} \dot{x}(t) = v(t), \\ \dot{v}(t) = -\nabla U(x(t)). \end{cases} \tag{3.3.16}$$

Well-known first integrals of motion are then the total energy

$$\Psi(x,v) = \frac{1}{2}|v|^2 + U(x)$$

and the components of the angular momentum

$$\Psi_{ij}(x,v) = x_j v_i - x_i v_j.$$

For more details, see [2].

3.4 Conservative and dissipative systems

In the study of dynamical systems, a key distinction can be made between conservative systems and dissipative systems. These classifications are based on the behavior of the system concerning energy conservation and the nature of the trajectories in the phase space.

3.4.1 Conservative Systems

A dynamical system is said to be *conservative* if it preserves a certain quantity, typically referred to as energy, throughout its evolution. Mathematically, such systems are characterized by the absence of energy loss or dissipation. A common example of conservative systems arises in Hamiltonian mechanics, where the system's dynamics are governed by a Hamiltonian function H that remains constant along the trajectories.

Definition 3.4.1 (Conservative System)**.** *A system described by a differential equation* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ *is conservative if there exists a differentiable scalar function* $H(\mathbf{x})$ *such that:*

$$\frac{dH}{dt} = \nabla H \cdot \mathbf{f}(\mathbf{x}) = 0,$$

for all $\mathbf{x}$ *in the domain of* H*.*

Energy Conservation in Newtonian Mechanics: Newton's second law, $F = ma$, serves as the foundation for many second-order systems. For instance, consider a particle of mass m moving along the x-axis, subject to a nonlinear force $F(x)$. The equation of motion is

$$m\ddot{x} = F(x).$$

Here, F is assumed to be independent of both $\dot{x}$ and t, implying no damping or external time-dependent driving force. Energy conservation in this context can be demonstrated as follows:
Let $V(x)$ denote the potential energy, defined by $F(x) = -\dfrac{dV}{dx}$. Then the equation becomes:

$$m\ddot{x} + \frac{dV}{dx} = 0.$$

Multiplying through by $\dot{x}$ and applying the chain rule, we find:

$$m\dot{x}\ddot{x} + \frac{dV}{dx}\dot{x} = 0 \quad \Rightarrow \quad \frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2 + V(x)\right) = 0.$$

Thus, the total energy:

$$E = \frac{1}{2}m\dot{x}^2 + V(x),$$

is constant over time. This conserved quantity is often referred to as the total

energy, a constant of motion, or a first integral. Systems with a conserved quantity are classified as *conservative systems*.

General Definition: More generally, given a system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, a conserved quantity is a real-valued continuous function $E(\mathbf{x})$ that remains constant along trajectories, i.e., $\dfrac{dE}{dt} = 0$. To avoid trivial examples, $E(\mathbf{x})$ must also be nonconstant on every open set. Otherwise, constant functions like $E(\mathbf{x}) \equiv 0$ would qualify, and every system would be conservative.

Exercice 3.4.1. *Show that a conservative system cannot have any attracting fixed points.*

Solution: Suppose x_0 were an attracting fixed point. By definition, an attracting fixed point is one where all trajectories in a neighborhood of x_0 converge to x_0 as time approaches infinity. In other words, there exists a neighborhood U of x_0 such that for all initial conditions $x(0) \in U$, the trajectories $x(t)$ satisfy $\lim_{t\to\infty} x(t) = x_0$. Now, consider the conserved quantity $E(x)$ for the system. Since $E(x)$ is constant along trajectories, all points within the neighborhood U must have the same energy $E(x_0)$, as trajectories originating in U flow toward x_0. This implies that $E(x)$ is constant throughout U.
However, this contradicts the definition of a conservative system, which requires that the conserved quantity $E(x)$ be nonconstant on every open set. Thus, x_0 cannot be an attracting fixed point in a conservative system.

Conclusion: Conservative systems cannot have any attracting fixed points because the existence of an attracting fixed point would violate the requirement that the conserved quantity $E(x)$ be nonconstant on open sets.

Example 3.4.1. *If attracting fixed points cannot occur in a conservative system, the fixed points that can arise are typically of other types, such as saddles and centers. These are illustrated in the following problem.*

Problem: *Consider a particle of mass $m = 1$ moving in a double-well potential*

$$V(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4.$$

Find and classify all the equilibrium points for the system. Then, plot the phase portrait and interpret the results physically.

Solution: *The force associated with the potential is given by* $F = -\frac{dV}{dx}$. *Computing the derivative, we have*

$$-\frac{dV}{dx} = x - x^3.$$

Thus, the equation of motion becomes

$$\ddot{x} = x - x^3.$$

This can be rewritten as a first-order system:

$$\begin{cases} \dot{x} = y, \\ \dot{y} = x - x^3, \end{cases}$$

where y *represents the particle's velocity.*

Equilibrium points occur when $\dot{x} = 0$ *and* $\dot{y} = 0$. *Solving these equations yields the equilibria* $(x*, y*) = (0, 0) \quad and \quad (x_0, y_0) = (\pm 1, 0)$.

To classify these fixed points, we compute the Jacobian matrix of the system

$$A = \begin{pmatrix} 0 & 1 \\ 1 - 3x^2 & 0 \end{pmatrix}.$$

1. At $(0, 0)$, *the Jacobian becomes*

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

The determinant $\delta = -1$, *indicating a saddle point.*

2. At $(\pm 1, 0)$, *the Jacobian becomes*

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

Here, the trace $\tau = 0$ *and the determinant* $\delta = 2$, *predicting centers.*

Nonlinear Behavior: *While small nonlinear terms can disrupt centers predicted by linear analysis, this is not the case here due to energy conservation. The system's trajectories correspond to contours of constant energy:*

$$E = \frac{1}{2}y^2 - \frac{1}{2}x^2 + \frac{1}{4}x^4 = \textit{constant}.$$

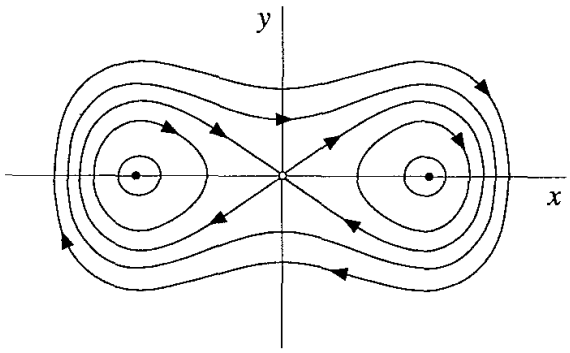


Figure 3.1: Phase portrait with a saddle, centers, closed orbits, and homoclinic trajectories

Phase Portrait: *The phase portrait, in Figure 3.1, is composed of:*

- *A saddle point at* $(0, 0)$.
- *Centers at* $(1, 0)$ *and* $(-1, 0)$.
- *Closed orbits around the centers, corresponding to periodic motion.*
- *Larger closed orbits encircling all three fixed points.*
- *Two special trajectories (homoclinic orbits) that approach the saddle point* $(0, 0)$ *as* $t \to \pm\infty$.

Physical Interpretation: *This system models the motion of an undamped particle in a double-well potential. The saddle at* $(0, 0)$ *corresponds to the unstable equilibrium at the top of the potential barrier. The centers at* $(\pm 1, 0)$ *represent stable equilibria at the minima of the potential wells. The closed orbits indicate periodic oscillations of the particle around these minima, while the homoclinic orbits describe trajectories where the particle transitions between the wells in infinite time.*
The neutrally stable equilibria at $(\pm 1, 0)$ *represent the particle at rest at the bottoms of the wells, with the small closed orbits around them corresponding to low-energy*

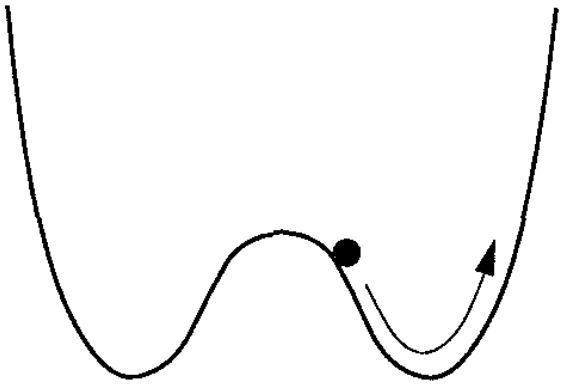


Figure 3.2: The motion of an undamped particle in a double-well potential

oscillations. The larger closed orbits represent higher-energy oscillations where the particle repeatedly crosses the potential barrier. The saddle point at $(0, 0)$ *corresponds to the unstable equilibrium at the top of the barrier, while the homoclinic orbits describe trajectories where the particle approaches the saddle point asymptotically, representing borderline cases between oscillations confined to a single well and transitions between both wells, see Figures 3.1 and 3.2.*

3.4.2 Dissipative Systems

Dissipative systems are characterized by the decrease of a specific function, often representing energy, over time along solution trajectories. Formally, consider a dynamical system

$$\frac{dx}{dt} = f(x, y), \quad \frac{dy}{dt} = g(x, y).$$

If there exists a function $H(x, y)$, often called a Lyapunov function or energy function, such that

$$\frac{d}{dt}H(x(t), y(t)) \leq 0 \quad \text{for all solution trajectories } (x(t), y(t)),$$

then the system is said to be dissipative.

Properties of Dissipative Systems:

- **Energy Dissipation:** Dissipative systems are distinguished by the continuous decrease of energy (or a related measure) over time.

- **Lyapunov Function:** The Lyapunov function $H(x, y)$ serves as a mathematical tool to demonstrate the dissipation property, satisfying $\frac{d}{dt}H \leq 0$.

Examples:

1. **Physical Interpretation:** A damped pendulum is a classic example of a dissipative system. Due to friction, its oscillations decrease in amplitude over time until the pendulum comes to rest at its equilibrium position.

2. **Linear Systems:**

   - A spiral sink is a typical feature of dissipative dynamics, where trajectories spiral inward as energy dissipates.

Qualitative Dynamics

In dissipative systems, solutions often tend towards an equilibrium state or attractor. For instance, in the case of a damped pendulum, the oscillations gradually decrease until the pendulum rests at the lowest energy position.

3.5 Exercises

Exercice 3.5.1. *Find the equilibrium points of the following system and determine their nature.*

$$\begin{cases} \dot{x} = x(2 - x - y), \\ \dot{y} = y(1 - x). \end{cases} \tag{3.5.17}$$

Exercice 3.5.2. *Consider the linear system:*

$$\begin{cases} \dot{x} = 3x + 4y, \\ \dot{y} = -2x - y. \end{cases} \tag{3.5.18}$$

1. *Find the equilibrium point(s).*

2. *Compute the Jacobian matrix and its eigenvalues.*

3. *Determine the stability of the equilibrium point(s).*

Exercice 3.5.3. *Classify the fixed points of the system:*

$$\begin{cases} \dot{x} = x + y, \\ \dot{y} = x - y. \end{cases} \tag{3.5.19}$$

using eigenvalue analysis. Identify whether they are stable, unstable, or saddle points.

Exercice 3.5.4. *Consider the nonlinear system:*

$$\begin{cases} \dot{x} = x - y + x^2, \\ \dot{y} = x + y^2. \end{cases} \tag{3.5.20}$$

1. *Find the equilibrium points.*
2. *Compute the Jacobian matrix and evaluate it at the equilibrium points.*
3. *Determine the local stability of each equilibrium using eigenvalues.*

Exercice 3.5.5. *Consider the system:*

$$\begin{cases} \dot{x} = -x + xy, \\ \dot{y} = -y + x^2. \end{cases} \tag{3.5.21}$$

1. *Find the equilibrium points.*
2. *Propose a Lyapunov function candidate $V(x, y)$.*
3. *Analyze stability using Lyapunov's method.*

Exercice 3.5.6. *Show that the system*

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -\sin x, \end{cases}$$

has a first integral and find it.

Chapter 4

Periodic Solutions and Their Stability

This chapter examines periodic solutions in dynamical systems, focusing on their existence, stability, and characterization. We introduce limit sets and invariant sets, followed by periodic solutions, limit cycles, and Poincaré maps. Stability analysis is then discussed, along with Bendixson's and Dulac's criteria for ruling out limit cycles. Finally, we present the Poincaré-Bendixson theorem, which provides a key result for planar systems.

4.1 Limit sets and invariant sets

Consider the autonomous nonlinear system

$$\dot{x} = f(x) \tag{4.1.1}$$

with $f \in C^1(E)$, where E is an open subset of $\mathbb{R}^n$. Without loss of generality, we assume that the system (4.1.1) defines a dynamical system $\phi(t, x)$ on E (i.e. a C^1-map $\phi : \mathbb{R} \times E \longrightarrow E$) where the family of maps $\phi_t(x) = \phi(t, x)$ have the properties of a flow.

Let $x_0 \in E$, and consider the function $\phi(\cdot, x) : \mathbb{R} \longrightarrow E$, which defines the solution curve, or trajectory, or orbit of system (4.1.1) starting from the initial point x_0 in E. By identifying the function $\phi(\cdot, x)$ with its graph, we can interpret the trajectory

through x_0 as the set of points traced by the motion along the curve

$$\Gamma_{x_0} = \{x \in E |\ x = \phi(t, x_0),\ t \in \mathbb{R}\},$$

defined by the equation (4.1.1).

Remark 4.1.1. *If the specific point x_0 is not relevant to the discussion, we simply denote the trajectory by Γ and represent the curve Γ within the subset E of the phase space $\mathbb{R}^n$, marking the direction of motion along Γ with an arrow to indicate the progression over time.*

The positive half-trajectory through the point $x_0 \in E$ refers to the motion along the curve

$$\Gamma_+ = \{x \in E \mid x = \phi(t, x_0),\ t \geq 0\}.$$

Similarly, the negative half-trajectory Γ_- is defined for $t < 0$. Any complete trajectory Γ can then be expressed as $\Gamma = \Gamma_+ \cup \Gamma_-$.

Definition 4.1.1. *A point $p_1 \in E$ is an ω-limit point of the trajectory $\phi(\cdot, x)$ of system (4.1.1) if there exists a sequence $t_n \to \infty$ such that*

$$\lim_{n\to\infty} \phi(t_n, x) = p_1.$$

Similarly, if there exists a sequence $t_n \to -\infty$ such that

$$\lim_{n\to\infty} \phi(t_n, x) = p_2,$$

for some point $p_2 \in E$, then p_2 is called an α-limit point of the trajectory $\phi(\cdot, x)$ of (4.1.1). The set of all ω-limit points of a trajectory Γ is called the ω-limit set of Γ and is denoted by $\omega(\Gamma)$. The set of all α-limit points of a trajectory Γ is called the α-limit set of Γ and is denoted by $\alpha(\Gamma)$. The union of these sets, $\alpha(\Gamma) \cup \omega(\Gamma)$, is called the limit set of Γ.

Corollary 4.1.1. *The α- and ω-limit sets of a trajectory Γ of (4.1.1), denoted $\alpha(\Gamma)$ and $\omega(\Gamma)$, are closed subsets of E.*

The proof of this corollary is left as an exercise for the student.

Theorem 4.1.1. *If p is an ω-limit point of a trajectory Γ of (4.1.1), then all other points of the trajectory $\phi(\cdot, p)$ of (4.1.1) through the point p are also ω-limit points of Γ; that is, if $p \in \omega(\Gamma)$, then $\phi(\mathbb{R}, p) \subset \omega(\Gamma)$. Similarly, if $p \in \alpha(\Gamma)$, then $\phi(\mathbb{R}, p) \subset \alpha(\Gamma)$.*

Proof. Let $p \in \omega(\Gamma)$ where Γ is the trajectory $\phi(\cdot, x_0)$ of (4.1.1) through the point $x_0 \in E$. Let q be a point on the trajectory $\phi(\cdot, p)$ of (4.1.1) through the point p; that is, $q = \phi(\bar{t}, p)$ for some $\bar{t} \in \mathbb{R}$. Since p is an ω-limit point of the trajectory $\phi(\cdot, x_0)$, there is a sequence $t_n \to \infty$ such that $\phi(t_n, x_0) \to p$. Hence, by the theorem of continuity with respect to initial conditions, and by the second property of dynamical systems,

$$\phi(t_n + \bar{t}, x_0) = \phi(\bar{t}, \phi(t_n, x_0)) \to \phi(\bar{t}, p) = q.$$

Since $t_n + \bar{t} \to \infty$, the point q is also an ω-limit point of Γ. A similar argument applies when p is an α-limit point of Γ, and this completes the proof of the theorem. □

Definition 4.1.2. *Let E be an open subset of $\mathbb{R}^n$, let $f \in C^1(E)$, and let $\phi_t : E \to E$ be the flow of the differential equation (4.1.1) defined for all $t \in \mathbb{R}$. A set $S \subset E$ is called invariant with respect to the flow ϕ_t if $\phi_t(S) \subset S$ for all $t \in \mathbb{R}$. Furthermore, S is called positively invariant (or negatively invariant) with respect to the flow ϕ_t if $\phi_t(S) \subset S$ for all $t \geq 0$ (or $t \leq 0$).*

It follows from the last theorem that for every point $p \in \omega(\gamma)$, we have $\phi_t(p) \in \omega(\gamma)$ for all $t \in \mathbb{R}$. In other words, $\phi_t(\omega(\gamma)) \subset \omega(\gamma)$. Therefore, according to the previous definition, we can state the following result.

Corollary 4.1.2. *$\alpha(\gamma)$ and $\omega(\gamma)$ are invariant with respect to the flow ϕ_t of the differential equation (4.1.1).*

The α- and ω-limit sets of a trajectory γ of (4.1.1) are thus closed, invariant subsets of E. In the next definition, a neighborhood of a set A is any open set U containing A, and we say that $x(t) \to A$ as $t \to \infty$ if the distance $d(x(t), A) \to 0$ as $t \to \infty$.

Definition 4.1.3. *A closed invariant set $A \subset E$ is called an attracting set of (4.1.1) if there exists a neighborhood U of A such that for all $x \in U$, $\phi_t(x) \in U$ for all $t \geq 0$ and $\phi_t(x) \to A$ as $t \to \infty$. An attractor of (4.1.1) is an attracting set which contains a dense orbit.*

Note that any equilibrium point x_0 of (4.1.1) is its own α- and ω-limit set since $\phi(t, x_0) = x_0$ for all $t \in \mathbb{R}$. And if a trajectory γ of (4.1.1) has a unique ω-limit point x_0, then by the above corollary, x_0 is an equilibrium point of (4.1.1). A stable node or focus is the ω-limit set of every trajectory in some neighborhood of the point; and a stable node or focus of (4.1.1) is an attractor of (4.1.1). However, not every ω-limit set of a trajectory of (4.1.1) is an attracting set of (4.1.1); for example, a saddle x_0 of a planar system (4.1.1) is the ω-limit set of three trajectories in a neighborhood $N(x_0)$, but no other trajectories through points in $N(x_0)$ approach x_0 as $t \to \infty$.

If q is any regular point in $\alpha(\gamma)$ or $\omega(\gamma)$, then the trajectory through q is called a limit orbit of γ. Thus, by Theorem 2, we see that $\alpha(\gamma)$ and $\omega(\gamma)$ consist of equilibrium points and limit orbits of (4.1.1).

Example 4.1.1. *Consider the system*

$$\begin{cases} \dot{x} &= -y + x(1 - x^2 - y^2), \\ \dot{y} &= x + y(1 - x^2 - y^2). \end{cases} \tag{4.1.2}$$

In polar coordinates, this system can be expressed as

$$\begin{cases} \dot{r} &= r(1 - r^2), \\ \dot{\theta} &= 1. \end{cases}$$

We observe that the origin is an equilibrium point of this system. The flow spirals around the origin in a counter-clockwise direction, moving outward when $0 < r < 1$ *(since* $\dot{r} > 0$ *for* $0 < r < 1$*) and inward when* $r > 1$ *(since* $\dot{r} < 0$ *for* $r > 1$*). The counter-clockwise motion on the unit circle describes a trajectory* γ_0 *of (1) since* $\dot{r} = 0$ *when* $r = 1$.

The trajectory through the point $(\cos(\theta_0), \sin(\theta_0))$ *on the unit circle at* $t = 0$ *is given by* $x(t) = (\cos(t + \theta_0), \sin(t + \theta_0))^T$. *The phase portrait for this system is shown in Figure 4.1.*

The trajectory γ_0 *is known as a stable limit cycle. A formal definition of a limit cycle is provided in the following section.*

The stable limit cycle γ_0 of the system in Example 4.1.1, shown in Figure 4.1, serves

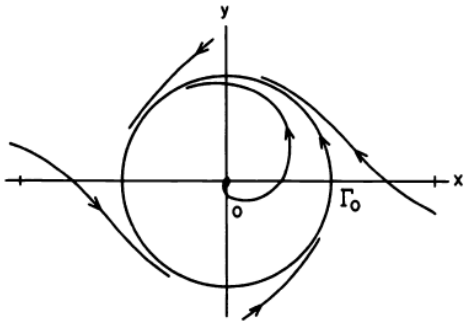


Figure 4.1: A stable limit cycle γ_0.

as the ω-limit set of every trajectory of this system except the equilibrium point at the origin. γ_0 consists of one limit orbit and acts as both its own α- and ω-limit set. This example illustrates that a trajectory or orbit γ of the system (4.1.1) refers to an equivalence class of solution curves $\phi(\cdot, x)$, where $x \in \gamma$. We usually select a representative solution $\phi(\cdot, x_0)$, with $x_0 \in \gamma$, to describe the trajectory and refer to it as the trajectory γ_{x_0} passing through the point x_0 at time $t = 0$. In the following section, we demonstrate that any stable limit cycle of (4.1.1) qualifies as an attractor of (4.1.1).

Example 4.1.2. *Consider the system*

$$\begin{cases} \dot{x} &= -y + x(1 - z^2 - x^2 - y^2), \\ \dot{y} &= x + y(1 - z^2 - x^2 - y^2), \\ \dot{z} &= 0. \end{cases} \tag{4.1.3}$$

In this case, the unit two-dimensional sphere S^2 (given by $x^2 + y^2 + z^2 = 1$) together with the portion of the z-axis outside S^2 forms an attracting set. Each plane $z = z_0$ is an invariant set and for $|z_0| < 1$ the ω-limit set of any trajectory not on the z-axis is a stable cycle (as defined in the next section) on S^2. Refer to Figure 4.2 for an illustration.

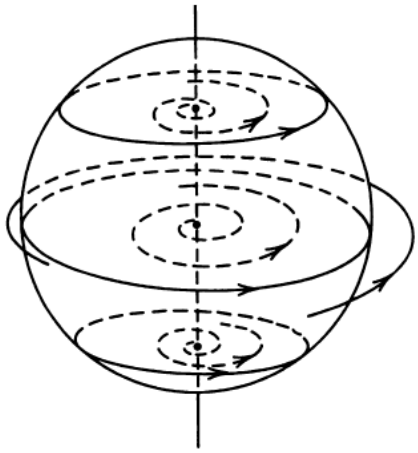


Figure 4.2: A dynamical system with S^2 as part of its attracting set.

Example 4.1.3. *Now consider the system*

$$\begin{cases} \dot{x} &= -y + x(1 - x^2 - y^2), \\ \dot{y} &= x + y(1 - x^2 - y^2), \\ \dot{z} &= a. \end{cases} \tag{4.1.4}$$

Here, the z-axis and the cylinder defined by $x^2 + y^2 = 1$ are invariant sets. The cylinder acts as an attracting set, as shown in Figure 4.3 for $a > 0$.

If in Example 4.1.3 we identify the points $(x, y, 0)$ and $(x, y, 2\pi)$ in the planes $z = 0$ and $z = 2\pi$, we obtain a flow in $\mathbb{R}^3$ with a two-dimensional invariant torus T^2 as an attracting set. The z-axis is mapped onto an unstable cycle γ (as defined in the next section). Furthermore, if a is an irrational multiple of π, then the torus T^2 becomes an attractor, and it is the ω-limit set of every trajectory except the cycle γ. Refer to Figure 4.4 for illustration.

In the following sections, we establish the Poincaré-Bendixson Theorem, demonstrating that the α and ω-limit sets of any two-dimensional system are relatively simple objects. Specifically, it is shown that they are either equilibrium points, limit cycles, or a union of separatrix cycles (defined in the next section). However, for higher-dimensional systems, the α and ω-limit sets can be much more complex, as the next example illustrates. One of the strange types of limit sets that can arise in higher-dimensional systems is called the "strange attractor" of the Lorenz system.

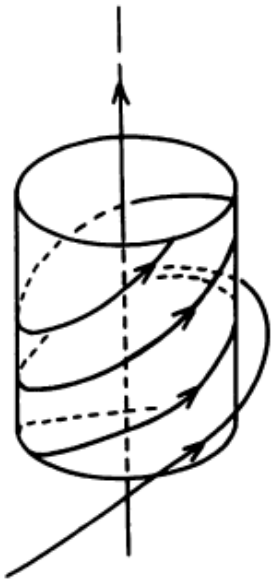


Figure 4.3: A dynamical system with a cylinder as its attracting set.

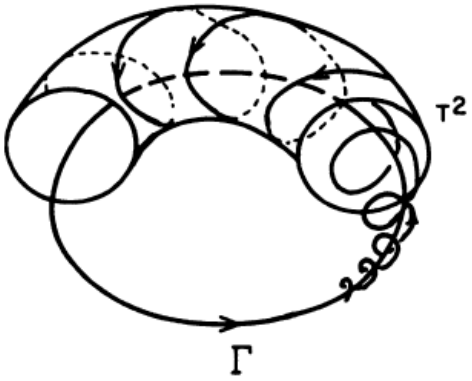


Figure 4.4: A dynamical system with an invariant torus T^2 as its attracting set.

Example 4.1.4. ***(The Lorenz System)***. *The original work of Lorenz in 1963, as well as the more recent work of Sparrow [13], indicates that for certain values of the parameters α, β, and ρ, the system*

$$\begin{cases} \dot{x} &= r(y-x), \\ \dot{y} &= \rho x - y - xz, \\ \dot{z} &= -bz + xy, \end{cases}$$

has a strange attracting set. For example, for $r = 10$, $\rho = 28$, and $b = \frac{8}{3}$, a single trajectory of this system is shown in Figure 6 along with a "branched surface" S. The attractor A of this system consists of an infinite number of branched surfaces S that are interleaved and intersect; however, the trajectories of this system within A do not intersect but instead, move from one branched surface to another as they circulate through the apparent branch.
The numerical results in [13] and the related theoretical work in [6] indicate that the closed invariant set A contains:

- *A countable set of periodic orbits with arbitrarily large periods,*
- *An uncountable set of nonperiodic motions, and*
- *A dense orbit.*

The attracting set A with these properties is referred to as a strange attractor.

4.2 Periodic solutions and limit cycles

As in earlier sections, we shall consider here a dynamical system $\phi(t, x)$ defined by

$$\dot{x} = f(x) \tag{4.2.5}$$

A periodic solution in a dynamical system is a solution where the system's state repeats itself after a constant period of time.

Definition 4.2.1. *A solution $x = \mathcal{E}(t)$ of the system (4.2.5) is called a periodic solution if there exists a positive number T such that $\mathcal{E}(t+T) = \mathcal{E}(t)$ for all $t \in \mathbb{R}$.*

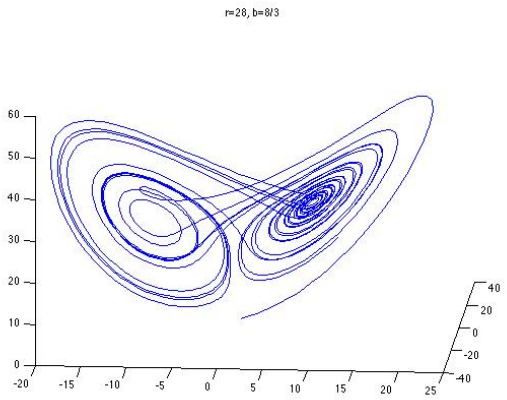


Figure 4.5: The Lorenz attractor.

In other words, a periodic orbit of (4.2.5) is any closed solution curve of (4.2.5) that is not an equilibrium point.

Remarks 4.2.1.

- *If the solution $\mathcal{E}(t)$ has a period T, then it also has periods $2T$, $3T$, and so on. If T is the smallest period among all the periods of $\mathcal{E}(t)$, then the solution is called T-periodic.*

- *Note that $\mathcal{E}(t) = \phi(t, x_0)$ for the system (4.2.5) where $\phi(t, x)$ is its flow. Thus, a cycle of (4.2.5) correspond to a periodic solution of (4.2.5) since $\phi(\cdot, x_0)$ defines a closed solution curve of (4.2.5) if and only if for all $t \in \mathbb{R}$ $\phi(t + T, x_0) = \phi(t, x_0)$ for some $T > 0$.*

Definition 4.2.2. *A periodic orbit $\mathcal{E}$ is called stable if, for each $\epsilon > 0$, there exists a neighborhood U of $\mathcal{E}$ such that for all $x \in U$, $d(\mathcal{E}_x^+, \mathcal{E}) < \epsilon$. That is, if for all $x \in U$ and $t \geq 0$, $d(\phi(t, x), \mathcal{E}) < \epsilon$. The orbit $\mathcal{E}$ is called unstable if it does not meet the previous condition. Furthermore, $\mathcal{E}$ is asymptotically stable if it is stable and if, for all points x within some neighborhood U of $\mathcal{E}$,*

$$\lim_{t \to \infty} d(\phi(t, x), \mathcal{E}) = 0.$$

Remark 4.2.1. *A center is an equilibrium point surrounded by a continuous family of cycles. Typically, the period T of these cycles varies continuously as one moves along a continuous path intersecting this family. However, for a center in a linear system, each periodic orbit within this family has the same period. Each orbit in this family surrounding a center is stable, but not asymptotically stable.*

Periodic orbits are important in various fields. For example:

- Understanding the motion of celestial bodies where they interact with each other by mutual gravitational force which may lead to periodic trajectories.
- In nonlinear dynamics, when analyzing systems that have complex behaviors, periodic solutions can help in understanding the structure of chaotic attractors.
- In mathematical physics, one studies Hamiltonian systems in which periodic solutions become involved in energy levels and stability of the system.

Example 4.2.1. *Consider the equation of the pendulum governed by the equation*

$$\ddot{x} + \sin(x) = 0, \tag{4.2.6}$$

which can be transformed into the system

$$\begin{cases} \dot{x_1} &= x_2, \\ \dot{x_2} &= -\sin(x_1), \end{cases}$$

where x_1 represents the angular displacement and x_2 represents the angular velocity. In the phase plane, we see a family of closed orbits corresponding with periodic solutions, see Figure 4.6.

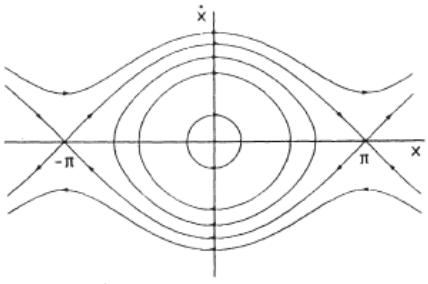


Figure 4.6: Periodic solutions of the pendulum equation (4.2.6).

Definition 4.2.3. *A limit cycle is a periodic orbit in the phase space that has the following properties:*

- *The limit cycle is isolated, meaning no other periodic orbits are arbitrarily close to it.*
- *All nearby trajectories either approach the limit cycle as time tends to infinity (stable limit cycle) or diverge away from it (unstable limit cycle).*

Definition 4.2.4. *Consider the nonlinear system (4.2.5) with $x \in \mathbb{R}^2$ and we shall refer to it as (4.2.5)*. A limit cycle $\mathcal{E}$ of a planar system (4.2.5)* is a cycle of (4.2.5)* that serves as the α- or ω-limit set of a trajectory of (4.2.5)* other than $\mathcal{E}$ itself. If a cycle $\mathcal{E}$ is the ω-limit set of every trajectory in some neighborhood of $\mathcal{E}$, then $\mathcal{E}$ is called an ω-limit cycle or stable limit cycle. If $\mathcal{E}$ is the α-limit set of every trajectory in some neighborhood of $\mathcal{E}$, then it is called an α-limit cycle or unstable limit cycle. Finally, if $\mathcal{E}$ is the ω-limit set of one trajectory other than itself and the α-limit set of another trajectory other than itself, then $\mathcal{E}$ is called a semi-stable limit cycle.*

Note that a stable limit cycle is an asymptotically stable cycle in the sense of Definition 4.2.2, and any stable limit cycle is an attractor. A stable limit cycle is shown in Example 4.1.1 in the previous section, and replacing t with $-t$ in that example (reversing the flow's direction) would yield an unstable limit cycle.

Example 4.2.2. *Consider the system*

$$\begin{cases} \dot{x} &= -y + x(x^2+y^2)\sin(\frac{1}{x^2+y^2}) \\ \dot{y} &= x + y(x^2+y^2)\sin(\frac{1}{x^2+y^2}), \end{cases}$$

for $x^2+y^2 \neq 0$, with $\dot{x} = \dot{y} = 0$ at $(0,0)$. This defines a C^1-system on $\mathbb{R}^2$, which can be rewritten in polar coordinates as

$$\dot{r} = r^3 \sin(\frac{1}{r}), \quad \dot{\theta} = 1.$$

Here, the origin is an equilibrium point, and the system has limit cycles $\mathcal{E}_n$ on the circles $r = \dfrac{1}{n\pi}$. These limit cycles accumulate at the origin, i.e.,

$$\lim_{n\to\infty} d(\mathcal{E}_n, 0) = 0.$$

The limit cycles $\mathcal{E}_{2n}$ are stable, while $\mathcal{E}_{2n+1}$ are unstable.

The following theorem of Dulac establishes that a planar analytic system, i.e. a system (4.2.5) where $f(x)$ is analytic on $E \subset \mathbb{R}^2$, cannot possess an infinite number of limit cycles accumulating at a critical point, as observed in the previous example.

Theorem 4.2.1. *In any bounded region of the plane, a planar analytic system has at most a finite number of limit cycles. Furthermore, any polynomial system has at most a finite number of limit cycles in* $\mathbb{R}^2$.

4.3 Poincaré maps

Discrete-time dynamical systems (maps) often naturally arise in the study of continuous-time systems governed by differential equations. The use of these maps facilitates the application of established results for discrete systems to continuous ones. This approach proves particularly advantageous when the derived map is defined in a space of lower dimension than the original differential system. A notable instance of such maps is the Poincaré map, which we define as the map constructed from ordinary differential equations to study the qualitative dynamics of the system. Poincaré maps are useful for studying swirling flows, such as those near periodic orbits, by reducing continuous dynamics to a discrete framework.
Consider the autonomous differential system

$$\dot{x} = f(x), \tag{4.3.7}$$

where $f : E \to \mathbb{R}^n$, with $E \subseteq \mathbb{R}^n$.

Definition 4.3.1. *Let S be an $n-1$ dimensional surface of section (Poincaré section) and be transversal to the flow i.e., all trajectories starting on S flow through it, not parallel to it. The Poincaré map $P : S \to S$ is a mapping obtained by following trajectories from one intersection with S to the next. If $x_k \in S$ represents the kth intersection point, then the Poincaré map is defined by*

$$x_{k+1} = P(x_k).$$

Remarks 4.3.1.

- *If x_0 is a fixed point of the Poincaré map P, i.e., $P(x_0) = x_0$, then:*

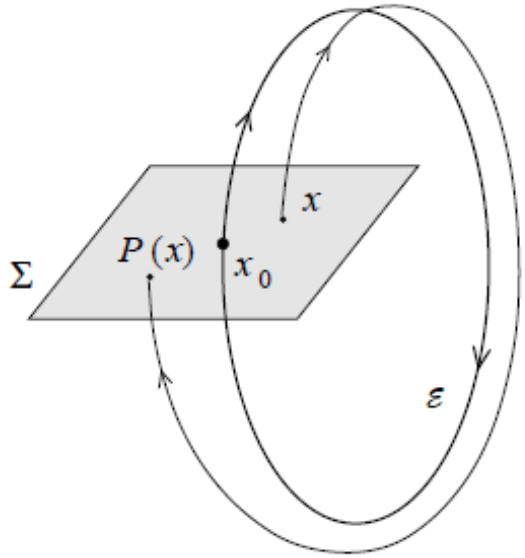


Figure 4.7: The Poincaré map associated with a cycle.

  - *A trajectory starting at x_0 returns to x_0 after some time T, forming a closed orbit in the original system $\dot{x} = f(x)$.*
  - *The stability of the closed orbit can be determined by examining the behavior of P near x_0.*

- *The Poincaré map simplifies the study of closed orbits by converting the problem into one of analyzing fixed points of P, which is generally easier.*

- *The challenge lies in the fact that finding an explicit formula for P is typically not feasible. However, there are cases where P can be computed explicitly, as shown in some examples below.*

- *The Poincaré map can also be defined when S is a hyperplane perpendicular to a periodic orbit $\mathcal{E}$ of the system (4.3.7) at x_0, then for any point $x \in S$ sufficiently near x_0, the solution $\phi_t(x)$ of (4.3.7) through x at $t = 0$ will cross S again at a point $P(x)$ near x_0. The mapping $x \to P(x)$ represents The Poincaré map.*

Theorem 4.3.1. *Let E be an open subset of $\mathbb{R}^n$, and let $f \in C^1(E)$. Suppose that*

$\phi_t(x_0)$ is a periodic solution of (4.3.7) with period T, and that the cycle

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid x = \phi_t(x_0),\, 0 \le t \le T\},$$

is contained in E. Let S be the hyperplane orthogonal to $\mathcal{E}$ at x_0, defined as:

$$S = \{x \in \mathbb{R}^n \mid (x - x_0) \cdot f(x_0) = 0\},$$

Then there exists a $\delta > 0$ and a unique function $\tau(x)$, defined and continuously differentiable for $x \in N_\delta(x_0)$, such that $\tau(x_0) = T$ and

$$\phi_{\tau(x)}(x) \in S, \quad \text{for all } x \in N_\delta(x_0).$$

Definition 4.3.2. *Let $\mathcal{E}$, S, δ, and $\tau(x)$ be as defined in Theorem 4.3.1. For $x \in N_\delta(x_0) \cap S$, the Poincaré map associated with $\mathcal{E}$ at x_0 is given by*

$$P(x) = \phi_{\tau(x)}(x).$$

Example 4.3.1. *Consider the vector field given in polar coordinates by:*

$$\begin{cases} \dot{r} = r(1 - r^2), \\ \dot{\theta} = 1. \end{cases}$$

Let S represent the positive x-axis. To compute the Poincaré map, note that the first return to S occurs after a time of flight $t = 2\pi$, since $\dot{\theta} = 1$. Let r_0 be the initial condition on S. The return value $r_1 = P(r_0)$ is determined by solving:

$$\frac{dr}{r(1 - r^2)} = dt \quad \text{for } t = 2\pi.$$

Evaluating this integral gives:

$$P(r) = \left[1 + e^{-2\pi}(r^{-2} - 1)\right]^{-1/2}.$$

The fixed point $r^* = 1$ satisfies $P(r^*) = r^*$, as the graph of $P(r)$ intersects the diagonal $r_{n+1} = r_n$ at r^*. Using Graphical iteration method (as shown in Figure

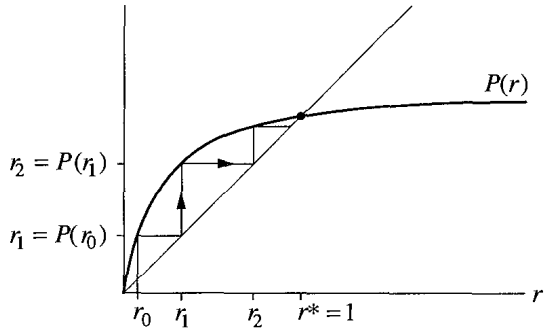


Figure 4.8: The graph of P.

4.8):

- *Start with an initial input r_0, draw a vertical line to the graph of $P(r)$.*
- *From this point, draw a horizontal line to the diagonal $r_{n+1} = r_n$.*
- *Repeat to observe the iterative convergence.*

The Graphical iteration diagram confirms that $r^* = 1$ is stable, as successive iterations converge to this point. This aligns with the fact that the system has a stable limit cycle at $r = 1$.
The system has a unique periodic orbit at $r = 1$, which is stable. The Poincaré map provides a discrete-time approach to analyze this stability.

For planar systems, translating the origin to $x_0 \in \mathcal{E} \cap S$ makes the normal line S pass through the origin. The point $0 \in \mathcal{E} \cap S$ divides S into segments S^+ and S^-, with $s > 0$ for S^+ (which lies in the exterior of $\mathcal{E}$) and $s < 0$ for S^-.
The Poincaré map $P(s)$ is defined for $|s| < \delta$, with $P(0) = 0$. Introducing the displacement function $d(s) = P(s) - s$ (since P is of class C^1 according to Theorem 4.3.1 then d is also of the same class), we have $d(0) = 0$ and, by the Mean Value Theorem, $d(s) = d'(0)s$ for small s.

- If $d'(0) < 0$, the cycle $\mathcal{E}$ is a stable (or ω-limit) cycle.
- If $d'(0) > 0$, the cycle $\mathcal{E}$ is an unstable (or α-limit) cycle.

Thus, the stability of $\mathcal{E}$ is determined by the derivative $P'(0)$. Specifically:

- If $P'(0) < 1$, $\mathcal{E}$ is a stable limit cycle.

- If $P'(0) > 1$, $\mathcal{E}$ is an unstable limit cycle.

The following theorem provides a formula for $P'(0)$ based on $f(x)$, which is useful for stability analysis.

Theorem 4.3.2. *Let the system (4.3.7) be a planar system i.e., $E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, and assume that $f \in C^1(E)$. Let $\gamma(t)$ be a periodic solution of this system with period T. Then the derivative of the Poincaré map $P(s)$ along a straight line S normal to $\mathcal{E} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = \gamma(t), 0 \leq t \leq T\}$ at $\mathbf{x} = 0$, is given by:*

$$P'(0) = \exp\left(\int_0^T \nabla \cdot f(\gamma(t))\, dt\right).$$

Corollary 4.3.1. *Under the hypotheses of Theorem 4.3.2, the periodic solution $\gamma(t)$ is a stable limit cycle if*

$$\int_0^T \nabla \cdot f(\gamma(t))\, dt < 0,$$

and it is an unstable limit cycle if

$$\int_0^T \nabla \cdot f(\gamma(t))\, dt > 0.$$

If this quantity is zero, the cycle may be stable, unstable, semi-stable, or it may belong to a continuous band of cycles.

Example 4.3.2. *Consider the system*

$$\begin{cases} \dot{x} = -y + x(1 - x^2 - y^2), \\ \dot{y} = x + y(1 - x^2 - y^2). \end{cases}$$

This system has a limit cycle represented by the periodic solution

$$\gamma(t) = (\cos t, \sin t)^T.$$

The function $\nabla \cdot f(x, y)$ is given by

$$\nabla \cdot f(x, y) = 2 - 4x^2 - 4y^2.$$

Now, we compute the integral:

$$\int_0^T \nabla \cdot f(\gamma(t))\, dt = \int_0^{2\pi} \left(2 - 4\cos^2 t - 4\sin^2 t\right) dt = -4\pi.$$

Thus, with $s = r - 1$, it follows from Theorem 4.3.2 that

$$P'(0) = e^{-4r},$$

Since $P'(0) < 1$, the cycle $\gamma(t)$ is a stable limit cycle.

4.4 Stability of periodic orbits

Theorem 4.4.1. *Let $\mathcal{U}$ be an open subset of $\mathbb{R}^2$ and suppose that $f \in C^1(\mathcal{U})$. Let $\gamma(t)$ be a T-periodic solution of (4.3.7). Then, $\gamma(t)$ is a stable limit cycle if*

$$\int_0^T \nabla \cdot f(\gamma(t))dt < 0,$$

and it is an unstable limit cycle if

$$\int_0^T \nabla \cdot f(\gamma(t))dt > 0.$$

It may be a stable, unstable, or semi-stable limit cycle if this quantity equals zero.

4.5 Bendixon's and Dulac's criteria

Consider the autonomous planar system

$$\dot{x} = F(x, y)\,, \quad \dot{y} = G(x, y), \tag{4.5.8}$$

with $(x, y) \in D \subset \mathbb{R}^2$.

Theorem 4.5.1. *(Criterion of bendixon) Assume that the domain D is simply connected and the functions, F and G, are continuously differentiable in D. If the divergence $\nabla.(F, G)$ of the vector field (F, G) is not identically zero and does not*

change sign in D, then the system (4.5.8) has no periodic orbit contained entirely in D.

Example 4.5.1. *Consider the nonlinear oscillator system with nonlinear damping*

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= -m(x) - n(x)y, \end{cases}$$

where $m(x)$ and $n(x)$ are smooth functions, with $n(x) > 0$ for all $x \in \mathbb{R}$.
Let $(F(x,y), G(x,y)) = (y, -m(x) - n(x)y)$. The divergence of (F,G) is negative definite, i.e. $\nabla.(F,G) = -n(x)$. From Bendixon's criterion theorem, the previous system has no periodic solutions.

A more general result of the theorem of Bendixon is given by Dulac's Criterion.

Theorem 4.5.2. *(Criterion of Dulac) Assume that the domain D is simply connected and the function, F and G, are continuously differentiable in D. If there exists a continuously differentiable function H in D such that $\nabla.(H(F,G))$ is not identically zero and does not change sign in D, then the following statements hold*

- *The system (4.5.8) has no periodic orbit contained entirely in D.*
- *If S is an annular region contained in D on which $\nabla.(H(F,G))$ does not change sign, then there exists at most one limit cycle of the system (4.5.8).*

4.6 The Poincaré-Bendixon theorem

Consider the equation (4.3.7) $\dot{x} = f(x)$ with $x \in \mathcal{U} \subseteq \mathbb{R}^n$. Previously, we defined an orbit $\mathcal{O}$ (or trajectory) passing through a point x_0 in $\mathcal{U}$ as the set

$$\mathcal{O}(x_0) = \{x \in \mathcal{U} : x = \phi(t, x_0), \text{ for all } t \in \mathbb{R} \text{ such that } \phi(t, x_0) \text{ is defined}\},$$

where ϕ is the flow of the system (4.3.7).
In this section, we also define the positive half-trajectory through the point $x_0 \in \mathcal{U}$ by the curve

$$\mathcal{O}^+(x_0) = \{x \in \mathcal{U} : x = \phi(t, x_0),\ t \geq 0\}.$$

Likewise, we define $\mathcal{O}^-(x_0)$ and we obviously note that $\mathcal{O}(x_0) = \mathcal{O}^+(x_0) \cup \mathcal{O}^-(x_0)$.

Remark 4.6.1. *If the point x_0 is not specifically involved in the motion of the trajectory, we simply denote it by $\mathcal{O}$. Similarly, we note that $\mathcal{O} = \mathcal{O}^+ \cup \mathcal{O}^-$*

Definition 4.6.1. *If there exists a sequence $t_n \longrightarrow \infty$ such that*

$$\lim_{n\to-\infty} \phi(t_n, x) = \ell_1, \quad \textit{where } \ell_1 \in \mathcal{U},$$

then the point ℓ_1 is called an ω-limit point of the trajectory $\mathcal{O}$ of the system (4.3.7). Similarly, If there exists a sequence $t_n \longrightarrow -\infty$ such that

$$\lim_{n\to-\infty} \phi(t_n, x) = \ell_2, \quad \textit{where } \ell_2 \in \mathcal{U},$$

then the point ℓ_2 is called an α-limit point of the trajectory $\mathcal{O}$ of the system (4.3.7).

Definition 4.6.2. *An ω-limit set of a trajectory $\mathcal{O}$ is the set of all ω-limit points of $\mathcal{O}$, and is denoted by $\omega(\mathcal{O})$. Similarly, An α-limit set of a trajectory $\mathcal{O}$ is the set of all α-limit points of $\mathcal{O}$, and is denoted by $\alpha(\mathcal{O})$. The limit set of $\mathcal{O}$, $\omega(\mathcal{O}) \cup \alpha(\mathcal{O})$, is the set of all limit points of $\mathcal{O}$.*

Theorem 4.6.1. *If ℓ is an ω-limit point of a trajectory $\mathcal{O}$ of (4.3.7), i.e. $\ell \in \omega(\mathcal{O})$, then $\mathcal{O}_\ell \subset \omega(\mathcal{O})$. Similarly, if $\ell \in \alpha(\mathcal{O})$ then $\mathcal{O}_\ell \subset \alpha(\mathcal{O})$.*

This theorem shows that given a point $\ell \in \omega(\mathcal{O})$ of a system (4.3.7), then all other points of the trajectory $\phi(., \ell)$ of (4.3.7) through the point ℓ are also ω-limit points of $\mathcal{O}$. Analogously, the same result applies in the case of α-limit points.

Remarks 4.6.1.

- *An equilibrium point x_0 of system (4.3.7) is its own α and ω-limit set.*
- *A stable node (or focus) is the ω-limit set of every trajectory in some neighborhood of the point.*

Definition 4.6.3. *If ℓ is any regular point in $\alpha(\mathcal{O})$ or $\omega(\mathcal{O})$ then the trajectory through ℓ is called a limit orbit of $\mathcal{O}$.*

Example 4.6.1. *We consider the following differential system*

$$\begin{cases} \dot{x} &= -y + x(1 - x^2 - y^2), \\ \dot{y} &= x + y(1 - x^2 - y^2). \end{cases}$$

When we change the system into polar coordinates we get

$$\begin{cases} \dot{r} &= r(1 - r^2), \\ \dot{\theta} &= 1. \end{cases}$$

The origin is an equilibrium point for the system and in the phase plane, the flow:

- *Spirals around the origin in a counter-clockwise direction.*
- *Spirals outward for* $0 < r < 1$ *because* $\dot{r} > 0$.
- *Spirals inward for* $r > 1$ *because* $\dot{r} < 0$.

Since $\dot{r} = 0$ *on* $r = 1$, *the flow approaches the unit circle in the counter-clockwise direction describing a trajectory* $\mathcal{O}_0$. *This trajectory passes through the point* $(\cos(\theta_0), \sin(\theta_0))$ *on the unit circle at* $t = 0$, *see Figure 4.9. The trajectory* $\mathcal{O}_0$ *is called a stable limit*

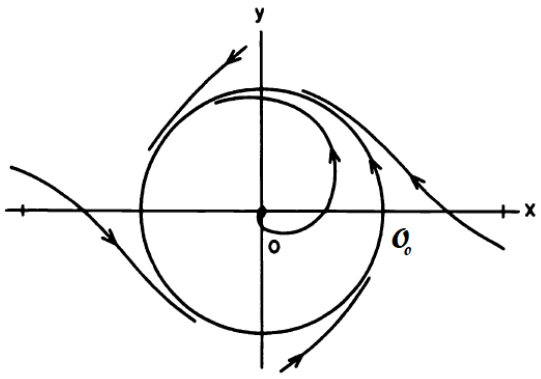


Figure 4.9: A stable limit cycle $\mathcal{O}_0$.

cycle. It is the ω*-limit set of every trajectory of this system except the equilibrium point at the origin.* $\mathcal{O}_0$ *consist of one limit orbit and it is its own* α *and* ω*-limit set.*

Theorem 4.6.2. *(The Poincaré-Bendixon theorem). Consider the system (4.3.7) where* $f \in C^1(\mathcal{U})$ *with* $\mathcal{U} \subset \mathbb{R}^2$. *Assume that (4.3.7) has a trajectory* $\mathcal{O}$ *with* $\mathcal{O}^+$

contained in a compact subset D of $\mathcal{U}$ and has a finite number of equilibrium points in D. Then, ω-limit set $\omega(\mathcal{O})$ is either an equilibrium point of (4.3.7), a periodic orbit of (4.3.7), or that $\omega(\mathcal{O})$ is composed of a finite number of equilibria $\ell_1, \ldots, \ell_m$ of (4.3.7) and a countable number of limit orbits of (4.3.7) whose α and ω limit sets belong to $\{\ell_1, \ldots, \ell_m\}$.

This theorem is proved in [12].
In the following theorem, we present a more specific result that guarantees the existence of closed orbits when no equilibrium points are present.

Theorem 4.6.3. *(Poincaré-Bendixson Theorem (special case)). Let Γ be a nonempty compact limit set of a C^1 planar dynamical system. If Γ contains no equilibrium points, then Γ is a closed orbit.*

For a proof of this theorem, see [8].

Example 4.6.2. *Consider the system*

$$\begin{cases} \dot{x} &= -y + x(x^2 + y^2 - 2x - 3), \\ \dot{y} &= x + y(x^2 + y^2 - 2x - 3). \end{cases} \tag{4.6.9}$$

The system (4.6.9) has one equilibrium point located at the origin $(0, 0)$. First, to determine the possibility of periodic orbits for the system we apply the criterion of Bendixon, so we compute the divergence of the vector function on the righthand side of system (4.6.9) we get

$$4x^2 + 4y^2 - 6x - 6 = 4\left[(x - \frac{3}{4})^2 + y^2 - \frac{33}{16}\right].$$

We deduce that no closed orbits can be contained in the interior of the circle with center $(\frac{3}{4}, 0)$ and radius $\dfrac{\sqrt{33}}{4}$ because inside that circle (Bendixon circle) the divergence is sign definite. Closed orbits are possible only if they enclose or intersect with this Bendixon circle.
We write the system (4.6.9) in polar coordinates we find

$$\begin{cases} \dot{r} &= r(r^2 - 2r\cos(\theta) - 3), \\ \dot{\theta} &= 1. \end{cases} \tag{4.6.10}$$

If $r < 1$ we have $\dot{r} < 0$, and if $r > 3$ we have $\dot{r} > 0$. According to the Poincaré-Bendixon theorem, the annulus $1 < r < 3$ must contain one or more limit cycles.

4.7 Exercises

Exercice 4.7.1. *Consider the Lorenz system:*

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = \rho x - y - xz, \\ \dot{z} = xy - \beta z, \end{cases} \tag{4.7.11}$$

where $\sigma > 0$, $\rho > 0$, and $\beta > 0$.

(a) Show that this system is invariant under the transformation $(x, y, z, t) \mapsto (-x, -y, z, t)$.

(b) Show that the z-axis is invariant under the flow of this system and that it consists of three trajectories.

(c) Show that this system has equilibrium points at the origin and at

$$(\pm\sqrt{\beta(\rho - 1)}, \pm\sqrt{\beta(\rho - 1)}, \rho - 1)$$

for $\rho > 1$. For $\rho > 1$, show that there is a one-dimensional unstable manifold $W^u(0)$ at the origin.

(d) For $\rho \in (0, 1)$, use the Lyapunov function

$$V(x, y, z) = \sigma x^2 + y^2 + \beta z^2$$

to show that the origin is globally stable; i.e., for $\rho \in (0, 1)$, the origin is the ω-limit set of every trajectory of this system.

Exercice 4.7.2. *Show that the system*

$$\begin{cases} \dot{x} = -y + x(1 - x^2 - y^2)^2, \\ \dot{y} = x + y(1 - x^2 - y^2)^2 \end{cases} \tag{4.7.12}$$

has a semi-stable limit cycle Γ. *Sketch the phase portrait for this system.*

Exercice 4.7.3. *Consider the system*

$$\begin{aligned}\dot{x} &= -y + x(1 - x^2 - y^2 - z^2)(4 - x^2 - y^2 - z^2),\\ \dot{y} &= x + y(1 - x^2 - y^2 - z^2)(4 - x^2 - y^2 - z^2),\\ \dot{z} &= 0.\end{aligned}$$

For $z = z_0$, *a constant, write the resulting system in polar coordinates and deduce that this system has two invariant spheres. Sketch the phase portrait for this system and describe all limit sets of trajectories of this system. Does this system have an attracting set?*

Exercice 4.7.4. *Show that* $\gamma(t) = (2\cos(2t), \sin(2t))^T$ *is a periodic solution of the system*

$$\begin{aligned}\dot{x} &= -4y + x(1 - \frac{x^2}{4} - y^2),\\ \dot{y} &= x + y(1 - \frac{x^2}{4} - y^2).\end{aligned}$$

that lies on the ellipse $\dfrac{x^2}{2^2} + y^2 = 1$; *i.e.,* $\gamma(t)$ *represents a cycle* Γ *of this system. Then use the corollary 4.3.1 to show that* Γ *is a stable limit cycle.*

Exercice 4.7.5. *Solve the linear system* $\dot{x} = Ax$ *with*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

and show that at any point $(x_0, 0)$, *on the x-axis, the Poincaré map for the focus at the origin is given by*

$$P(x_0) = x_0 \exp\left(\frac{2\pi a}{|b|}\right).$$

For $d(x) = P(x) - x$, *compute* $d'(0)$ *and show that* $d(-x) = -d(x)$.

Exercice 4.7.6. *(a) Use the Dulac function* $B(x, y) = be^{-2\beta x}$ *to show that the system*

$$\dot{x} = y, \quad \dot{y} = -ax - by + \alpha x^2 + \beta y^2$$

has no limit cycle in $\mathbb{R}^2$.

(b) Show that the system

$$\dot{x} = \frac{y}{1+x^2}, \quad \dot{y} = \frac{-x + y(1+x^2+x^4)}{1+x^2}$$

has no limit cycle in $\mathbb{R}^2$.

Exercice 4.7.7. *Consider the dynamical system*

$$\begin{cases} \dot{x} = -y + x(r^4 - 3r^2 + 1), \\ \dot{y} = x + y(r^4 - 3r^2 + 1), \end{cases}$$

where $r^2 = x^2 + y^2$.

(a) Show that $\dot{r} < 0$ *on the circle* $r = 1$ *and that* $\dot{r} > 0$ *on the circle* $r = 2$. *Use the Poincaré-Bendixson theorem and the fact that the only critical point of this system is at the origin to show that there exists a periodic orbit in the annular region*

$$A_1 = \{x \in \mathbb{R}^2 \mid 1 < |x| < 2\}.$$

(b) Show that the origin is an unstable focus for this system. Use the Poincaré-Bendixson theorem to show that there exists a periodic orbit in the annular region

$$A_2 = \{x \in \mathbb{R}^2 \mid 0 < |x| < 1\}.$$

(c) Determine the stable and unstable limit cycles of this system.

Chapter 5

Introduction to Local Bifurcations

Bifurcation theory is a field in mathematics that studies qualitative changes in the behavior of a system as a parameter is varied. These changes often involve the sudden appearance or disappearance of equilibria, periodic orbits, or more complex structures in dynamical systems. Bifurcations can signal the onset of instability or chaos in systems and are categorized into types, such as saddle-node, transcritical, pitchfork, and Hopf bifurcations. Each type corresponds to a different way in which the solutions of a system change in response to parameter variations.

5.1 Basic concepts of bifurcation

Definition 5.1.1. *A bifurcation occurs when a small change in a system's parameters causes a qualitative or topological change in its long-term behavior (its phase portrait or equilibrium structure). Essentially, a bifurcation marks a point where the nature of the solutions changes.*

Definition 5.1.2. *In bifurcations, the following key points apply*

- *A fixed point is where the system's state does not change over time. Bifurcations involve changes in the number or stability of these fixed points as parameters are varied.*
- *The stability of a fixed point determines whether small disturbances around it die out (stable) or grow (unstable). Stability is often assessed using eigenvalues, which are calculated from the system's linearization near the fixed point.*

A bifurcation occurs when these eigenvalues cross a critical threshold, such as changing from negative to positive.

- *There are different types of bifurcations, each describing how the system's behavior changes:*
  - ***Saddle-node bifurcation**: Two fixed points, one stable and one unstable, merge and annihilate each other.*
  - ***Transcritical bifurcation**: Two fixed points exchange their stability as the parameter changes.*
  - ***Pitchfork bifurcation**: A single fixed point splits into three (two stable, one unstable) or merges into one.*
  - ***Hopf bifurcation**: A stable fixed point becomes unstable, giving rise to a periodic oscillation, known as a limit cycle.*

Definition 5.1.3. *A bifurcation diagram is a plot that shows how the system's fixed points or periodic orbits change as a parameter is varied. It provides a visual representation of the system's transitions and bifurcations. Stable points are typically shown with a solid line, while unstable points are drawn with a dashed line. Often, unstable points are omitted for simplicity.*

In this chapter, we focus on common local bifurcations in one and two dimensions. We limit our discussion to codimension 1 bifurcations, which occur when a single real parameter is varied.

5.2 Bifurcation in one-dimensional systems

Consider the scalar differential equation given by

$$\frac{dx(t)}{dt} = f(x(t), \mu), \tag{5.2.1}$$

where $x(t)$ is a real-valued function of time t, f is a real-valued vector field, and μ is the **bifurcation paramater**. We assume that equation (5.2.1) is well-defined and satisfies the hypotheses of the Cauchy-Lipschitz theorem, ensuring that for each

initial condition, a unique solution exists. Furthermore, we assume that the vector field is of class C^k (where $k \geq 2$) in a neighborhood of the point $(0,0)$, satisfying the following conditions:

$$\begin{array}{ll} (i) & f(0,0) = 0; \\ (ii) & \dfrac{\partial f}{\partial x}(0,0) = 0. \end{array} \tag{5.2.2}$$

The condition (i) indicates that $x = 0$ is an equilibrium point of equation (5.2.1) when $\mu = 0$. Our focus is on the local bifurcations that arise in the neighborhood of this equilibrium as the parameter μ is varied. The condition (ii) is necessary but not sufficient for the emergence of local bifurcations at $\mu = 0$.

5.2.1 Saddle-node bifurcation

A saddle-node bifurcation occurs when a stable and an unstable equilibrium collide and annihilate as a parameter varies. This fundamental bifurcation marks a transition in system dynamics and appears in various physical and biological models.

Theorem 5.2.1. *Consider the differential equation (5.2.1) with f a vector field of class C^k (where $k \geq 2$) in a neighborhood of $(0,0)$, satisfying*

$$\frac{\partial f}{\partial \mu}(0,0) = a, \quad \frac{\partial^2 f}{\partial x^2}(0,0) = 2b \ \text{ with } \ ab \neq 0. \tag{5.2.3}$$

The following properties are valid in the neighborhood of 0 in $\mathbb{R}$ for a sufficiently small μ.

1. *If $ab < 0$ (resp. $ab > 0$) the differential equation (5.2.1) has no equilibria for $\mu < 0$ (resp. for $\mu > 0$).*

2. *If $ab < 0$ (resp. $ab > 0$) the equation (5.2.1) possesses two equilibria $x_{\pm}(\epsilon)$, $\epsilon = \sqrt{|\mu|}$ for $\mu > 0$ (resp. $\mu < 0$), with opposite stabilities.*

Then for equation (5.2.1), a saddle-node bifurcation occurs at $\mu = 0$.

Remark 5.2.1. *The saddle-node bifurcation is also known as "fold bifurcation" and "turning point bifurcation." These names reflect the characteristic shape of the bifurcating equilibria in the (μ, x)-plane.*
The term "saddle node" specifically denotes the emergence of a saddle point and a

node as equilibria in n-dimensional systems, illustrating the different stability properties of the resulting solutions.

Recalling the general expansion of f around $(0,0)$ up to second order, we have

$$\begin{aligned} f(x,\mu) &= f(0,0) + \frac{\partial f}{\partial x}(0,0)x + \frac{\partial f}{\partial \mu}(0,0)\mu + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(0,0)x^2 + \frac{\partial f}{\partial \mu \partial x}(0,0)\mu x \\ &\quad + \frac{1}{2}\frac{\partial f}{\partial \mu^2}(0,0)\mu^2 + o(|(x,\mu)|^2). \end{aligned} \tag{5.2.4}$$

Consequently, from the conditions of the assumptions on the vector field and if we neglect quadratic terms in μ, then near $(x,\mu) = (0,0)$, we obtain

$$f(x,\mu) = a\mu + bx^2 + o(|\mu| + x^2) \quad \text{as} \ \ (x,\mu) \to (0,0). \tag{5.2.5}$$

In fact. Since $a \neq 0$, the implicit function theorem guarantees the existence of a unique solution $\mu = g(x)$, for x near zero, to the equation $f(x,\mu) = 0$. The function g is of class C^k with $k \geq 2$ in a neighborhood of the origin and satisfies $g(0) = 0$. The Taylor expansion of g is given by

$$\mu = -\frac{b}{a}x^2 + o(x^2).$$

As a result, if $ab\mu > 0$, equation (5.2.1) has no equilibria; if $\mu = 0$, there is a single equilibrium at $x = 0$; and if $ab\mu < 0$, there exists a pair of equilibria $x_\pm(\mu) = \pm\sqrt{-a\mu/b} + o(\sqrt{|\mu|})$.
In the case where $ab\mu < 0$, we find

$$\frac{\partial f}{\partial x}(x_\pm(\mu),\mu) = 2bx_\pm(\mu) + o(\sqrt{|\mu|}),$$

implying that the equilibrium $x_-(\mu)$ is attractive (asymptotically stable) when $b > 0$, and repelling (unstable) when $b < 0$. Conversely, the equilibrium $x_+(\mu)$ exhibits the opposite stability behavior.

Example 5.2.1. *Consider the one-dimensional system*

$$\dot{x} = \mu - x^2.$$

For $\mu > 0$, the system has two critical points at $x = \pm\sqrt{\mu}$, where the derivative of the right-hand side is

$$Df(x, \mu) = -2x, \quad Df(\pm\sqrt{\mu}, \mu) = \mp 2\sqrt{\mu}.$$

The critical point at $x = \sqrt{\mu}$ is stable, while the critical point at $x = -\sqrt{\mu}$ is unstable. (Here, $Df(x, \mu)$ denotes the derivative of the function $f(x, \mu)$ with respect to x.) For $\mu = 0$, there is a single critical point at $x = 0$, which is nonhyperbolic, since $Df(0, 0) = 0$; the vector field $f(x) = -x^2$ is unstable at this point, and $\mu = 0$ represents a bifurcation value. For $\mu < 0$, there are no critical points.
The phase portraits for this differential equation are shown in Figure 5.1. The bifurcation diagram in Figure 5.2 illustrates the bifurcation occurring at $\mu = 0$. The curve $\mu - x^2 = 0$ determines the position of the critical points, where a solid curve indicates stable critical points, and a dashed curve indicates unstable ones. This bifurcation is called a saddle-node bifurcation.

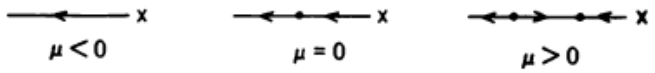


Figure 5.1: Phase portraits for the differential equation of Example 5.2.1.

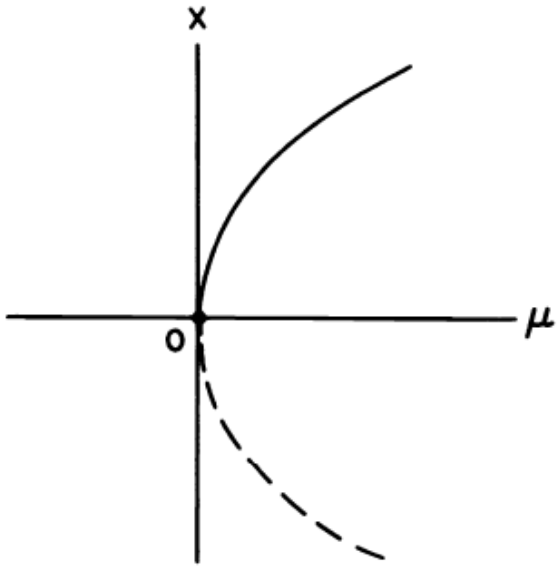


Figure 5.2: Bifurcation diagram for the saddle-node bifurcation in Example 5.2.1.

5.2.2 Pitchfork bifurcation

A pitchfork bifurcation occurs when a single equilibrium splits into three as a parameter varies, typically under symmetry constraints. This bifurcation plays a key role in dynamical systems, with distinct supercritical and subcritical cases. In the supercritical case, a stable equilibrium gives rise to two new stable equilibria with the original one becoming unstable, while in the subcritical case, two unstable equilibria merge into a single unstable equilibrium. This bifurcation appears in various physical and biological models.

Theorem 5.2.2. *Consider the differential equation (5.2.1), where f is a vector field of class C^k (with $k \geq 3$) defined in a neighborhood of $(0,0)$. The function f satisfies the conditions in (5.2.2) and is odd with respect to x, that is,*

$$f(-x,\mu) = -f(x,\mu). \tag{5.2.6}$$

Moreover, assume that

$$\frac{\partial^2 f}{\partial \mu \partial x}(0,0) = a, \quad \frac{\partial^3 f}{\partial x^3}(0,0) = 6b \ \text{ with } \ ab \neq 0. \tag{5.2.7}$$

The following properties are valid in the neighborhood of 0 in $\mathbb{R}$ for a sufficiently small μ.

1. *If $ab < 0$ (resp. $ab > 0$) the differential equation (5.2.1) has one equilibrium $x = 0$ for $\mu < 0$ (resp. for $\mu > 0$). This equilibrium is stable when $b < 0$ and unstable when $b > 0$.*

2. *If $ab < 0$ (resp. $ab > 0$) the equation (5.2.1) possesses the trivial equilibrium $x = 0$ and two nontrivial equilibria $x_\pm(\epsilon)$, $\epsilon = \sqrt{|\mu|}$ for $\mu > 0$ (resp. $\mu < 0$), which are symmetric, $x_+(\epsilon) = -x_-(\epsilon)$. Moreover, the map $\epsilon \to x_\pm(\epsilon)$ is of class C^{k-3} in a neighborhood of 0 in $\mathbb{R}$, and $x_\pm(\epsilon) = O(\epsilon)$. The nontrivial equilibria are stable when $b < 0$ and unstable when $b > 0$, whereas the trivial equilibrium has opposite stability.*

Then for equation (5.2.1), a pitchfork bifurcation occurs at $\mu = 0$.

As a result of conditions (5.2.2), (5.2.6), and (5.2.7), the function f admits the following Taylor expansion

$$f(x,\mu) = xh(x^2,\mu) \quad \text{with} \quad h(x^2,\mu) = a\mu + bx^2 + o(|\mu| + x^2) \quad \text{as} \quad (x,\mu) \to (0,0),$$

where h is of class $C^{(k-1)/2}$ in a neighborhood of $(0,0)$.

Example 5.2.2. *Consider the one-dimensional system*

$$\dot{x} = \mu x - x^3.$$

For $\mu > 0$, the system has critical points at $x = 0$ and $x = \pm\sqrt{\mu}$. For $\mu \leq 0$, the only critical point is at $x = 0$.
At $\mu = 0$, there is a nonhyperbolic critical point at $x = 0$, since $Df(0,0) = 0$. The vector field $f(x) = -x^3$ is unstable, and $\mu = 0$ is a bifurcation value.
The phase portraits for this differential equation are shown in Figure 5.3. The bifurcation diagram is shown in Figure 5.4. This type of bifurcation is called a pitchfork bifurcation.

Figure 5.3: Phase portraits for the differential equation in Example 5.2.2 .

Exercice 5.2.1. *Consider the truncated equation*

$$\frac{dx}{dt} = a\mu x + bx^3.$$

1. *Create bifurcation diagrams in the (u,μ)-plane for the truncated equation, considering different values of a and b.*

2. *Prove the theorem.*

5.2.3 Transcritical bifurcation

A transcritical bifurcation occurs when two equilibria exchange their stability as a parameter varies. Unlike the saddle-node bifurcation, the equilibria do not disappear

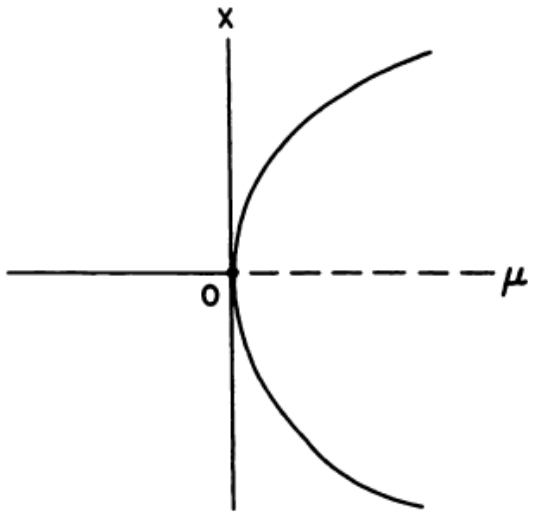


Figure 5.4: Bifurcation diagram for the pitchfork bifurcation in Example 5.2.2 .

but instead intersect and exchange their roles. This bifurcation often arises in models with conserved quantities or population dynamics.

Theorem 5.2.3. *Consider the differential equation (5.2.1) with f a vector field of class C^k (where $k \geq 2$) in a neighborhood of $(0,0)$, satisfying*

$$\frac{\partial f}{\partial \mu}(0,0) = a, \quad \frac{\partial^2 f}{\partial x^2}(0,0) = 2b \quad \textit{with} \quad ab \neq 0. \tag{5.2.8}$$

Assume that the following properties are valid in the neighborhood of 0 in $\mathbb{R}$ for a sufficiently small μ:

1. *If $ab < 0$ (resp. $ab > 0$) the differential equation (5.2.1) has no equilibria for $\mu < 0$ (resp. for $\mu > 0$).*

2. *If $ab < 0$ (resp. $ab > 0$) the equation (5.2.1) possesses two equilibria $x_\pm(\epsilon)$, $\epsilon = \sqrt{|\mu|}$ for $\mu > 0$ (resp. $\mu < 0$), with opposite stabilities. Moreover, the map $\epsilon \to x_\pm(\epsilon)$ is of class C^{k-2} in a neighborhood of 0 in $\mathbb{R}$, and $x_\pm(\epsilon) = O(\epsilon)$.*

Then for equation (5.2.1), a saddle-node bifurcation occurs at $\mu = 0$.

Example 5.2.3. *Consider the one-dimensional system*

$$\dot{x} = \mu x - x^2.$$

The critical points are at $x = 0$ and $x = \mu$. For $\mu = 0$, there is a single critical point

at $x = 0$, which is nonhyperbolic since $Df(0,0) = 0$. The vector field $f(x) = -x^2$ is unstable at this point, and $\mu = 0$ is a bifurcation value.

The phase portraits for this differential equation are shown in Figure 5.5. The bifurcation diagram is shown in Figure 5.6. At the bifurcation value $\mu = 0$, there is an exchange of stability at the critical points, marking the occurrence of a transcritical bifurcation. Details can be found in [5].

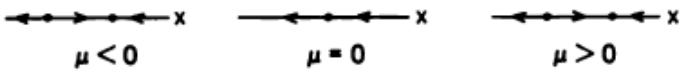


Figure 5.5: Phase portraits for the differential system in Example 5.2.3.

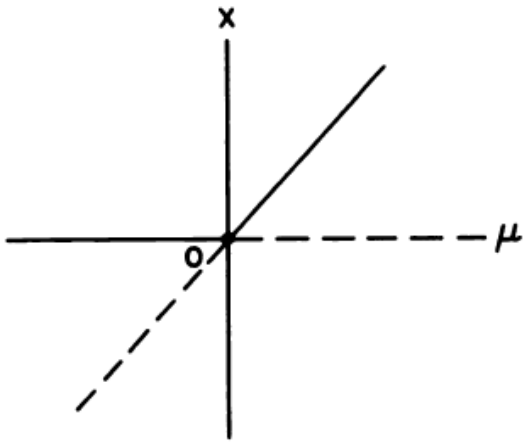


Figure 5.6: Diagram bifurcation for the transcritical bifurcation in Example 5.2.3.

5.3 Bifurcation in two-dimensional systems

This section extends our earlier study of bifurcations to two-dimensional systems. Here, both fixed points and closed orbits can be created, destroyed, or destabilized as parameters vary, enabling us to understand how oscillations begin or end. A bifurcation is defined as a change in the topological structure of the phase portrait due to parameter variation, affecting fixed points, closed orbits, or saddle connections.

5.3.1 Saddle-node bifurcation

A saddle-node bifurcation is the basic mechanism for creating or destroying fixed points. Consider the following two-dimensional prototype system

$$\begin{cases} \dot{x} &= \mu - x^2 \\ \dot{y} &= -y \end{cases} \tag{5.3.9}$$

For $\mu > 0$, there is a stable fixed point at $(x^*, y^*) = (\sqrt{\mu}, 0)$ and an unstable fixed point at $(x^*, y^*) = (-\sqrt{\mu}, 0)$. At $\mu = 0$, the phase portraits shown in the following figure illustrate the collision and annihilation of these fixed points.

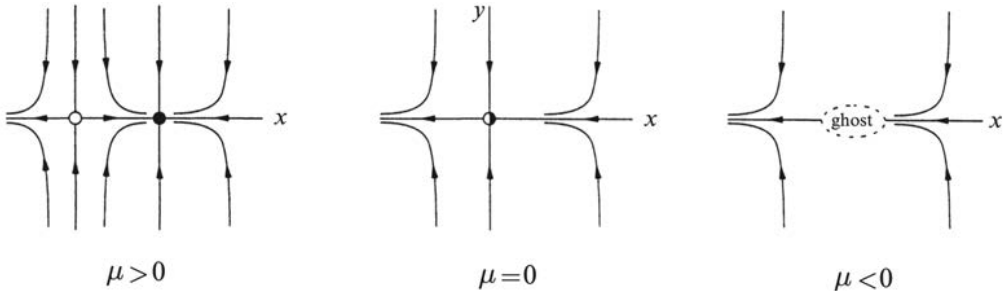


Figure 5.7: collision and annihilation of fixed points of system (5.3.9).

Figure 5.7 shows two fixed points: a stable node at $(x^*, y^*) = (\sqrt{\mu}, 0)$ and a saddle at $(x^*, y^*) = (-\sqrt{\mu}, 0)$. As μ decreases, the saddle and node approach each other, collide at $\mu = 0$, and disappear when $\mu < 0$. Even after the fixed points annihilate, they continue to influence the flow by leaving behind a bottleneck region. This region traps trajectories, delaying their exit. The time spent in the bottleneck increases as $\mu - \mu_c \sim (\mu_c - \mu)^{-1/2}$, where μ_c is the value at which the saddle-node bifurcation occurs.

Now, consider a more general scenario where we have the two-dimensional system

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y),$$

that depends on a parameter μ. Suppose that for some value of μ, the nullclines intersect as shown in Figure 5.8 Each intersection corresponds to a fixed point since $\dot{x} = 0$ and $\dot{y} = 0$ simultaneously. To track the movement of fixed points as μ changes, we simply observe these intersections. Now, suppose the nullclines move apart as μ

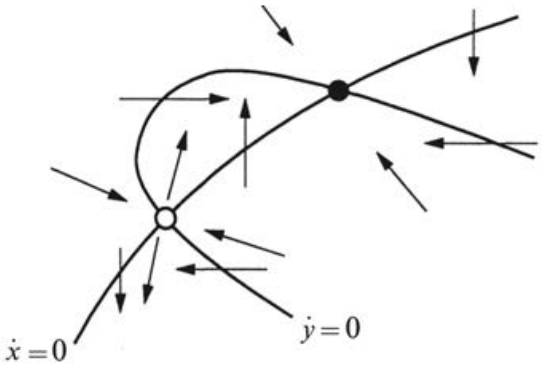


Figure 5.8: Intersection of nullclines.

varies, becoming tangent at $\mu = \mu_c$. At this point, the fixed points approach each other, collide, and vanish. Once the nullclines separate, there are no intersections, and the fixed points disappear abruptly. This behavior characterizes all saddle-node bifurcations locally. For further details, see [14].

5.3.2 Transcritical and pitchfork bifurcations

Based on our earlier work in one-dimensional bifurcations and analogous to the saddle-node case, we can construct the prototypical cases of transcritical and pitchfork bifurcations at a stable fixed point which are given by

1. **Transcritical Bifurcation:**

$$\begin{cases} \dot{x} = \mu x - x^2, \\ \dot{y} = -y. \end{cases}$$

2. **Supercritical Pitchfork Bifurcation:**

$$\begin{cases} \dot{x} = \mu x - x^3, \\ \dot{y} = -y. \end{cases}$$

3. **Subcritical Pitchfork Bifurcation:**

$$\begin{cases} \dot{x} &= \mu x + x^3, \\ \dot{y} &= -y. \end{cases} \tag{5.3.10}$$

The analysis in each case follows a similar pattern. Here, we will focus on the supercritical pitchfork bifurcation.

For $\mu < 0$, the only fixed point is a stable node at the origin. For $\mu = 0$, the origin remains stable, but the decay along the x-direction becomes very slow (algebraic) rather than exponential (this is the phenomenon of "critical slowing down"). For $\mu > 0$, the origin loses stability and gives rise to two new stable fixed points symmetrically located at $(x^*, y^*) = (\pm\sqrt{\mu}, 0)$. By computing the Jacobian at each fixed point, one

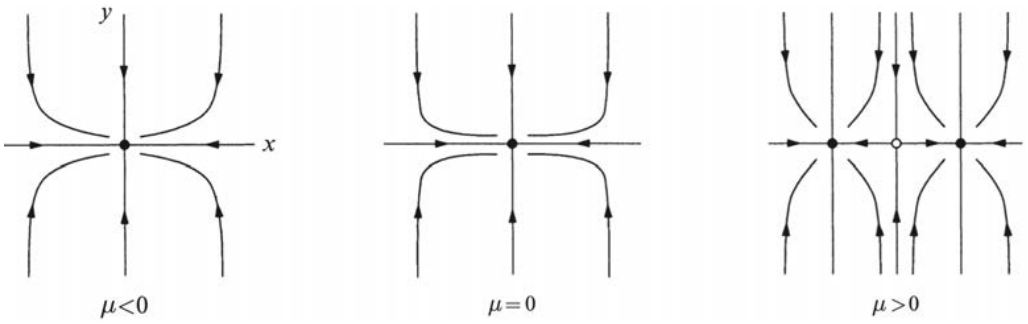


Figure 5.9: Phase portraits of system (5.3.10).

can verify that the origin becomes a saddle, while the two new fixed points are stable nodes. The phase portraits are shown in Figure 5.9.

Pitchfork bifurcations frequently occur in systems exhibiting symmetry. Consider the following example.

Example 5.3.1.

$$\begin{cases} \dot{x} &= \mu x + y + \sin x, \\ \dot{y} &= x - y. \end{cases} \tag{5.3.11}$$

The system (5.3.11) is invariant under the change of variables $x \to -x$, $y \to -y$, so the phase portrait must be symmetric under reflection through the origin. The

origin is a fixed point for all μ, and its Jacobian is

$$J = \begin{pmatrix} \mu + 1 & 1 \\ 1 & -1 \end{pmatrix},$$

which has $\tau = \mu$ and $\Delta = -(\mu + 2)$. Hence, the origin is a stable fixed point if $\mu < -2$ and a saddle if $\mu > -2$. This suggests that a pitchfork bifurcation occurs at $\mu_c = -2$.

To confirm this, we seek a symmetric pair of fixed points near the origin for μ close to μ_c. (At this stage, we do not know whether the bifurcation is sub- or supercritical). The fixed points satisfy $y = x$, and thus

$$(\mu + 1)x + \sin x = 0.$$

One solution is $x = 0$, which we have already identified. Now, assume x is small and nonzero, and expand the sine function as a power series:

$$(\mu + 1)x + x - \frac{x^3}{3!} + O(x^5) = 0.$$

After dividing through by x and neglecting higher-order terms, we obtain

$$\mu + 2 - \frac{x^2}{6} \approx 0.$$

Thus, there is a pair of fixed points with $x^* \approx \pm\sqrt{6(\mu + 2)}$ for μ slightly greater than -2. Therefore, a supercritical pitchfork bifurcation occurs at $\mu_c = -2$. (If the bifurcation had been subcritical, the pair of fixed points would have existed when the origin was stable, not after it has become a saddle.) Since the bifurcation is supercritical, we know that the new fixed points are stable without needing further analysis.

To draw the phase portrait near $(0, 0)$ for μ slightly greater than -2, it is helpful to find the eigenvectors of the Jacobian at the origin. This can be done exactly, but a simple approximation is to use the Jacobian close to that at the bifurcation. Thus,

for the Jacobian

$$J \approx \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

the eigenvectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ *and* $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$*, with eigenvalues* $\lambda = 0$ *and* $\lambda = -2$*, respectively. For* μ *slightly greater than* -2*, the origin becomes a saddle, and the zero eigenvalue becomes slightly positive. This information implies the phase portrait shown in Figure 5.10.*

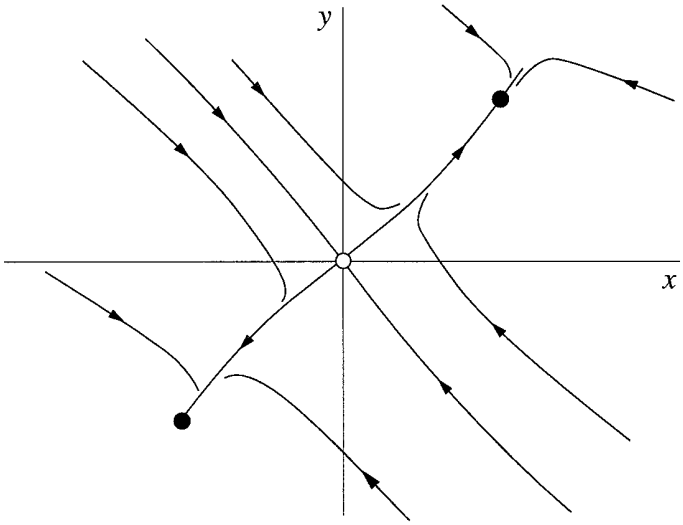


Figure 5.10: Phase portrait of system (5.3.11).

In the examples discussed above, bifurcations occur when the eigenvalue $\Delta = 0$, or equivalently, when one of the eigenvalues equals zero. More broadly, saddle-node, transcritical, and pitchfork bifurcations are all categorized as zero-eigenvalue bifurcations. While other examples of zero-eigenvalue bifurcations exist, these are among the most prevalent. Such bifurcations invariably involve the collision and subsequent annihilation of two or more fixed points.

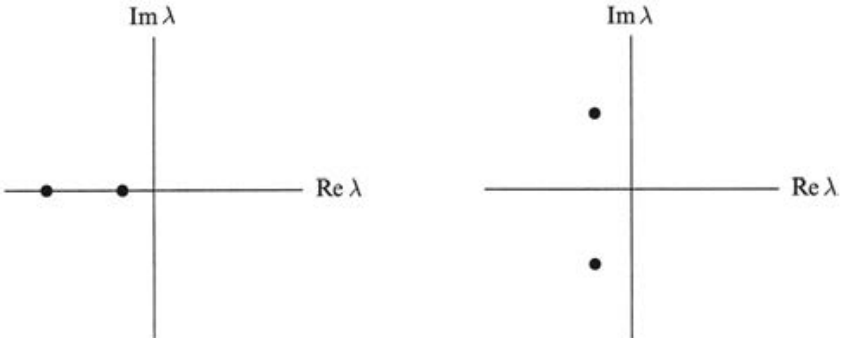
In the next subsection, we will explore a fundamentally distinct type of bifurcation, one that lacks a counterpart in one-dimensional systems. This bifurcation allows for a fixed point to lose stability without interacting with or colliding with any other fixed points.

5.3.3 Hopf bifurcations

a stable fixed point may become unstable only if the real part of one of the eigenvalues λ of the Jacobian becomes greater than zero as the control parameter μ varies. Recall that the eigenvalues λ solve

$$\det(J) - \lambda I = 0,$$

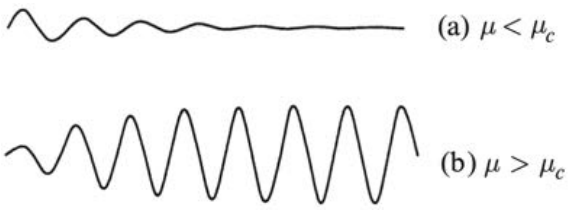
where J is the Jacobian and I is the identity matrix.
In two-dimension, the resulting quadratic equation either has two real roots or two complex conjugate roots. So when a fixed point is stable, either both roots are real and negative, or the roots are complex conjugates with negative real parts:



In the saddle-node, transcritical, and pitchfork bifurcations, one of the purely real eigenvalues passes through $\lambda = 0$ when the fixed point becomes unstable. In contrast, a Hopf bifurcation occurs when a pair of complex conjugate eigenvalues cross the imaginary axis. Similar to pitchfork bifurcations, Hopf bifurcations can be classified into two cases: supercritical and subcritical.

(i) Supercritical Hopf bifurcation

The complex eigenvalues produce oscillatory solutions. One possibility is that oscillations are damped for $\mu < \mu_c$ and growing for $\mu > \mu_c$.



This scenario corresponds to a supercritical Hopf bifurcation.

We can represent a simple model of the supercritical Hopf bifurcation using polar coordinates, expressed in terms of the radius r and angle θ in the 2-D phase space:

$$\begin{cases} \dot{r} &= \mu r - r^3, \\ \dot{\theta} &= \omega + br^2. \end{cases} \tag{5.3.12}$$

Here, μ governs the stability of the fixed point, ω represents the frequency of oscillations when r is infinitesimal, and b determines how the frequency depends on the amplitude of large oscillations.

The r-equation resembles the form of the supercritical pitchfork bifurcation, while the θ-equation provides the rotational component, effectively acting as a driving mechanism.

When $\mu \leq 0$, the origin ($r = 0$) is a stable spiral. Conversely, for $\mu > 0$, the origin becomes an unstable spiral.

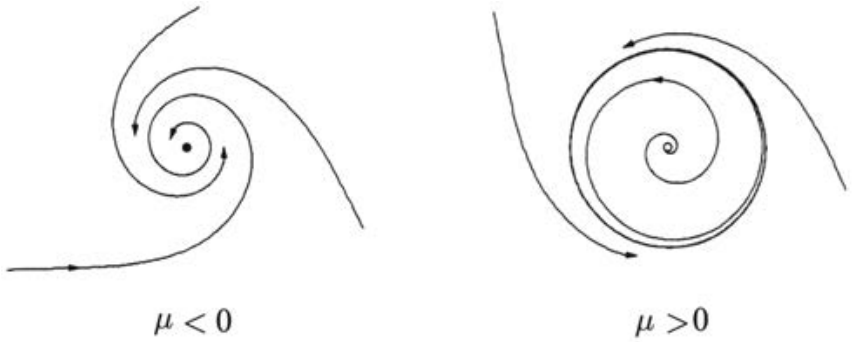


Figure 5.11: Phase portraits for μ above and below the bifurcation.

Note that for $\mu > 0$ there is a stable limit cycle at $r = \sqrt{\mu}$. When $\mu = 0$ the origin is still stable, however the amplitude of oscillations decays slower than exponentially (since $\dot{r} = -r^2$), and this is another case of critical slowing down.

To analyze how the eigenvalues behave during the bifurcation, we rewrite the system in Cartesian coordinates, as this simplifies finding the Jacobian.
Let $x = r\cos\theta$ and $y = r\sin\theta$. Then

$$\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta,$$

substituting $\dot{r} = \mu r - r^3$ and $\dot{\theta} = \omega + br^2$:

$$\dot{x} = (\mu r - r^3)\cos\theta - r(\omega + br^2)\sin\theta.$$

Expanding this in terms of x and y:

$$\dot{x} = (\mu - (x^2 + y^2))x - (\omega + b(x^2 + y^2))y.$$

Similarly, for $\dot{y}$:

$$\dot{y} = (\omega + b(x^2 + y^2))x + (\mu - (x^2 + y^2))y.$$

Ignoring cubic terms near the origin $(x, y \approx 0)$, the system reduces to:

$$\begin{cases} \dot{x} &= \mu x - \omega y, \\ \dot{y} &= \omega x + \mu y. \end{cases}$$

The Jacobian at the origin is thus

$$J = \begin{bmatrix} \mu & -\omega \\ \omega & \mu \end{bmatrix},$$

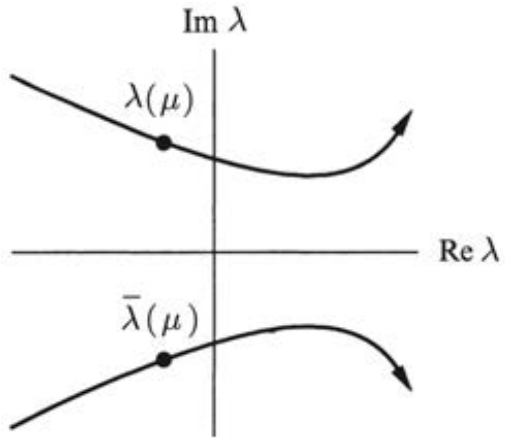
which has the eigenvalues $\lambda = \mu \pm i\omega$. As expected, the eigenvalues cross the imaginary axis from left to right as μ increases from negative to positive values. We can extract two important characteristics of supercritical Hopf bifurcations from the prototypical form (5.3.12):

- The size of the limit cycle grows continuously from zero and increases proportionally to $\sqrt{\mu - \mu_c}$, for μ close to μ_c.
- The frequency of the limit cycle is given approximately by $\omega = \Im(\lambda)$, evaluated at $\mu = \mu_c$. This formula is exact at the birth of the limit cycle and accurate within $O(\mu - \mu_c)$ for μ close to μ_c. Therefore, the period is

$$T = \frac{2\pi}{\Im(\lambda)} + O(\mu - \mu_c).$$

In this system, the eigenvalues cross the imaginary axis as a straight line parallel to

the real axis. However, the paths are usually curved: Moreover, the limit cycle is



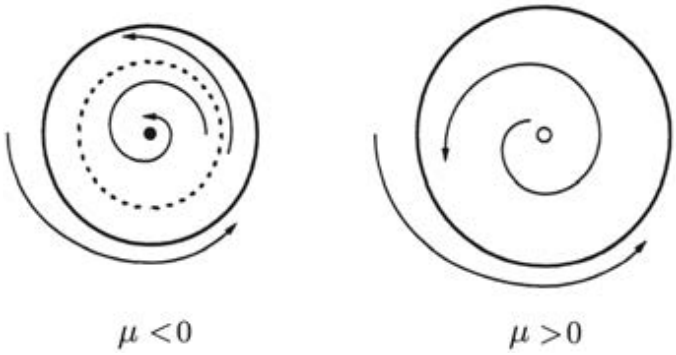
usually elliptical and not circular near μ_c.

(ii) Subcritical Hopf bifurcation

As with pitchfork bifurcations, there is also a subcritical variety of Hopf bifurcations in which the stability of the fixed point and the limit cycle is reversed. This type exhibits some very interesting properties. Here's the prototype:

$$\begin{cases} \dot{r} &= \mu r + r^3 - r^5, \\ \dot{\theta} &= \omega + br^2. \end{cases} \tag{5.3.13}$$

The important difference from the earlier supercritical case is that the cubic term r^3 is now destabilizing, as it drives trajectories away from the origin. The phase portraits look like this:



When $\mu < 0$, there are two attractors: a stable limit cycle and a stable fixed point. These two attractors are separated by an unstable limit cycle. As $\mu \to 0^-$, the unstable limit cycle shrinks to the origin.

The subcritical Hopf bifurcation occurs at $\mu = 0$. At this point, the origin becomes unstable, and the large amplitude limit cycle becomes the only attractor. Consequently, any solution near the origin is forced to grow immediately into a large amplitude oscillation.

Here's a pictorial summary of the supercritical and subcritical cases:

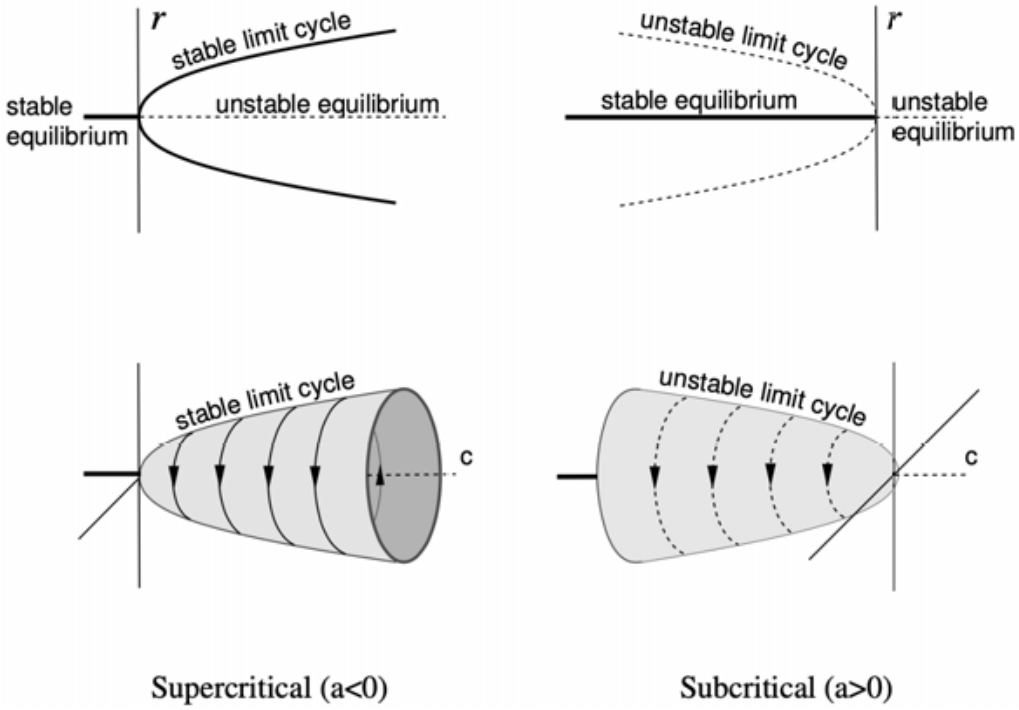


Figure 5.12: Summary of the supercritical and subcritical cases.

5.4 Exercises

Exercice 5.4.1. *Show that the first-order system*

$$\dot{x} = r - x - e^{-x},$$

undergoes a saddle-node bifurcation as r is varied, and find the value of r at the bifurcation point.

Exercice 5.4.2. *Show that the first-order system*

$$\dot{x} = x(1 - x^2) - a(1 - e^{-bx}),$$

undergoes a transcritical bifurcation at $x = 0$ when the parameters a, b satisfy a certain equation, to be determined. (This equation defines a bifurcation curve in the (a, b) parameter space.) Then find an approximate formula for the fixed point that bifurcates from $x = 0$, assuming that the parameters are close to the bifurcation curve.

Exercice 5.4.3. *Equations similar to*

$$\dot{x} = -x + \beta \tanh(x),$$

arise in statistical mechanical models of magnets and neural networks. Show that this equation undergoes a supercritical pitchfork bifurcation as β is varied. Then give a numerically accurate plot of the fixed points for each β.

Exercice 5.4.4. *Define the function*

$$f(x) = \begin{cases} x^3 \sin\left(\dfrac{1}{x}\right), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0. \end{cases}$$

Show that $f \in C^1(\mathbb{R})$. Consider the one-dimensional system

$$\dot{x} = f(x) - \mu,$$

with f defined above.

(a) Show that for $\mu = 0$ there are an infinite number of critical points in any neighborhood of the origin, that the nonzero critical points are hyperbolic and alternate in stability, and that the origin is a nonhyperbolic critical point.

(b) Show that $\mu = 0$ is a bifurcation value.

(c) Draw a bifurcation diagram and show that there are an infinite number of bifurcation values that accumulate at $\mu = 0$. What type of bifurcations occur at the nonzero bifurcation values?

Exercice 5.4.5. *Consider the system*

$$\dot{x} = y - 2x, \quad \dot{y} = \mu + x^2 - y.$$

(a) Sketch the nullclines.

(b) Find and classify the bifurcations that occur as μ varies.

(c) Sketch the phase portrait as a function of μ.

Exercice 5.4.6. *Consider the system*

$$\dot{x} = \mu x - y + xy^2, \quad \dot{y} = x + \mu y + y^3.$$

Show that a Hopf bifurcation occurs at the origin as μ varies. Is the bifurcation subcritical or supercritical?

List of Figures

2.1 Situations that would lead to a scalar autonomous differential equation having nonmonotone solutions. 45
2.2 Case $t_0, x_0 > 0$, subcase 2 , in the resolution of (2.2.13). 69
2.3 The extension $\tilde{\phi}$ on the interval $\tilde{\mathcal{I}}$ of the solution ϕ (defined on the interval $\mathcal{I}$). 71
2.4 ϕ_1 is a global and maximal solution on $\mathcal{I}$; ϕ_2 is a maximal solution on $\mathcal{I}$, but it is not global on $\mathcal{I}$. 74

3.1 Phase portrait with a saddle, centers, closed orbits, and homoclinic trajectories 98
3.2 The motion of an undamped particle in a double-well potential 99

4.1 A stable limit cycle γ_0. 106
4.2 A dynamical system with S^2 as part of its attracting set. 107
4.3 A dynamical system with a cylinder as its attracting set. 108
4.4 A dynamical system with an invariant torus T^2 as its attracting set. . 108
4.5 The Lorenz attractor. 110
4.6 Periodic solutions of the pendulum equation (4.2.6). 111
4.7 The Poincaré map associated with a cycle. 114
4.8 The graph of P. 116
4.9 A stable limit cycle $\mathcal{O}_0$. 121

5.1 Phase portraits for the differential equation of Example 5.2.1. 130
5.2 Bifurcation diagram for the saddle-node bifurcation in Example 5.2.1. 130
5.3 Phase portraits for the differential equation in Example 5.2.2 132
5.4 Bifurcation diagram for the pitchfork bifurcation in Example 5.2.2 . . 133

5.5 Phase portraits for the differential system in Example 5.2.3. 134
5.6 Diagram bifurcation for the transcritical bifurcation in Example 5.2.3. 134
5.7 collision and annihilation of fixed points of system (5.3.9). 135
5.8 Intersection of nullclines. 136
5.9 Phase portraits of system (5.3.10). 137
5.10 Phase portrait of system (5.3.11). 139
5.11 Phase portraits for μ above and below the bifurcation. 141
5.12 Summary of the supercritical and subcritical cases. 144

Bibliography

[1] Julien Arino. Fundamental theory of ordinary differential equations. *Lecture Notes. University of Manitoba*, 2006.

[2] Vladimir I Arnold. *Ordinary differential equations*. Springer Science & Business Media, 1992.

[3] Earl A Coddington, Norman Levinson, and T Teichmann. Theory of ordinary differential equations, 1956.

[4] J Dieudonńe. *Foundations of Modern Analysis*. Academic Press, NY, 1969.

[5] Grégory Faye. An introduction to bifurcation theory. *NeuroMathComp Laboratory, INRIA, Sophia Antipolis, CNRS, ENS Paris, France*, 2011.

[6] J Guckenheimer and P Holmes. Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. *Applied mathematical sciences*, 42, 1983.

[7] Philip Hartman. *Ordinary differential equations*. SIAM, 2002.

[8] Morris W Hirsch and Stephen Smale. *Differential equations, dynamical systems, and linear algebra*. Academic Press. INC., 1974.

[9] Po-Fang Hsieh and Yasutaka Sibuya. *Basic theory of ordinary differential equations*. Springer Science & Business Media, 2012.

[10] Yuri A Kuznetsov, Iu A Kuznetsov, and Y Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer, 1998.

[11] Horacio J Marquez. *Nonlinear control systems: analysis and design*, volume 1. Wiley-Interscience Hoboken, NJ, USA, 2003.

[12] Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.

[13] Colin Sparrow. The lorenz equations: Bifurcations, chaos, and strange attractors. *Applied mathematical sciences*, 1982.

[14] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

[15] Ferdinand Verhulst. *Nonlinear differential equations and dynamical systems*. Springer Science & Business Media, 2012.

[16] Kôsaku Yosida et al. Lectures on differential and integral equations. 1960.