Ordinary Differential Equations

Simon Brendle

Contents

Preface

These notes grew out of courses taught by the author at Stanford University during the period of 2006 – 2009. The material is all classical. The author is grateful to Messrs. Chad Groft, Michael Eichmair, and Jesse Gell-Redman, who served as course assistants during that time.

Chapter 1

Introduction

1.1. Linear ordinary differential equations and the method of integrating factors

A differential equation is an equation which relates the derivatives of an unknown function to the unknown function itself and known quantities. We distinguish two basic types of differential equations: An ordinary differential equation is a differential equation for an unknown function which depends on a single variable (usually denoted by t and referred to as time). By contrast, if the unknown function depends on two or more variables, the equation is a partial differential equation. In this text, we will restrict ourselves to ordinary differential equations, as the theory of partial differential equations is considerably more difficult.

Perhaps the simplest example of an ordinary differential equation is the equation

$$x'(t) = a\,x(t), \tag{1}$$

where $x(t)$ is a real-valued function and a is a constant. This is an example of a linear differential equation of first order. Its general solution is described in the following proposition:

Proposition 1.1. *A function $x(t)$ is a solution of (1) if and only if $x(t) = c\,e^{at}$ for some constant c.*

Proof. Let $x(t)$ be an arbitrary solution of (1). Then

$$\frac{d}{dt}(e^{-at}\,x(t)) = e^{-at}\,(x'(t) - a\,x(t)) = 0.$$

Therefore, the function $e^{-at}\,x(t)$ is constant. Consequently, $x(t) = c\,e^{at}$ for some constant c.

Conversely, suppose that $x(t)$ is a function of the form $x(t) = c\,e^{at}$ for some constant c. Then $x'(t) = ca\,e^{at} = a\,x(t)$. Therefore, any function of the form $x(t) = c\,e^{at}$ is a solution of (1). □

We now consider a more general situation. Specifically, we consider the differential equation

$$x'(t) = a(t)\,x(t) + f(t). \tag{2}$$

Here, $a(t)$ and $f(t)$ are given continuous functions which are defined on some interval $J \subset \mathbb{R}$. Like (1), the equation (2) is a linear differential equation of first order. However, while the equation (1) has constant coefficients, coefficients in the equation (2) are allowed to depend on t. In the following proposition, we describe the general solution of (2):

Proposition 1.2. *Fix a time $t_0 \in J$, and let $\varphi(t) = \int_{t_0}^t a(s)\,ds$. Then a function $x(t)$ is a solution of (2) if and only if*

$$x(t) = e^{\varphi(t)} \left(\int_{t_0}^t e^{-\varphi(s)}\,f(s)\,ds + c \right)$$

for some constant c.

Proof. Let $x(t)$ be an arbitrary solution of (2). Then

$$\begin{aligned} \frac{d}{dt}(e^{-\varphi(t)}\,x(t)) &= e^{-\varphi(t)}\,(x'(t) - \varphi'(t)\,x(t)) \\ &= e^{-\varphi(t)}\,(x'(t) - a(t)\,x(t)) \\ &= e^{-\varphi(t)}\,f(t). \end{aligned}$$

Integrating this identity, we obtain

$$e^{-\varphi(t)}\,x(t) = \int_{t_0}^t e^{-\varphi(s)}\,f(s)\,ds + c$$

for some constant c. This implies

$$x(t) = e^{\varphi(t)} \left(\int_{t_0}^t e^{-\varphi(s)}\,f(s)\,ds + c \right)$$

for some constant c.

Conversely, suppose that $x(t)$ is of the form

$$x(t) = e^{\varphi(t)} \left(\int_{t_0}^t e^{-\varphi(s)}\,f(s)\,ds + c \right)$$

for some constant c. Then

$$\begin{aligned} x'(t) &= \varphi'(t)\, e^{\varphi(t)} \left(\int_{t_0}^t e^{-\varphi(s)}\, f(s)\, ds + c \right) \\ &\quad + e^{\varphi(t)}\, \frac{d}{dt} \left(\int_{t_0}^t e^{-\varphi(s)}\, f(s)\, ds \right) \\ &= a(t)\, x(t) + f(t). \end{aligned}$$

Therefore, $x(t)$ solves the differential equation (2). This completes the proof. $\square$

1.2. The method of separation of variables

We next describe another class of differential equations of first order that can be solved in closed form. We say that a differential equation is separable if it can be written in the form

$$x'(t) = f(x(t))\, g(t). \tag{3}$$

Here, $f(x)$ and $g(t)$ are continuous functions which we assume to be given. Moreover, we assume that $U \subset \mathbb{R}$ is an open set such that $f(x) \neq 0$ for all $x \in U$.

In order to solve (3), we first choose a function $\varphi : U \to \mathbb{R}$ such that $\varphi'(x) = \frac{1}{f(x)}$. Suppose now that $x(t)$ is a solution of (3) which takes values in U. Then we obtain

$$\frac{d}{dt}\varphi(x(t)) = \varphi'(x(t))\, x'(t) = \frac{1}{f(x(t))}\, x'(t) = g(t).$$

Integrating both sides of this equation gives

$$\varphi(x(t)) = \int g(t)\, dt + c$$

for some constant c. Thus, the solution $x(t)$ can be written in the form

$$x(t) = \varphi^{-1}\Big(\int g(t)\, dt + c \Big),$$

where φ^{-1} denotes the inverse of φ. Note that the general solution of the differential equation involves an arbitrary constant c. If we prescribe the value of the function $x(t)$ at some time t_0, then this uniquely determines the integration constant c, and we obtain a unique solution of the given differential equation with that initial condition.

As an example, let us consider the differential equation

$$x'(t) = t\, (1 + x(t)^2).$$

To solve this equation, we observe that $\int \frac{1}{1+x^2}\,dx = \arctan(x)$. Hence, if $x(t)$ is a solution of the given differential equation, then

$$\frac{d}{dt}\arctan(x(t)) = \frac{1}{1+x(t)^2}\,x'(t) = t.$$

Integrating this equation, we obtain

$$\arctan(x(t)) = \frac{t^2}{2} + c$$

for some constant c. Thus, we conclude that

$$x(t) = \tan\Big(\frac{t^2}{2} + c\Big).$$

1.3. Problems

Problem 1.1. Find the solution of the differential equation $x'(t) = -\frac{2t}{1+t^2}\,x(t) + 1$ with initial condition $x(0) = 1$.

Problem 1.2. Find the solution of the differential equation $x'(t) = \frac{t}{t+1}\,y(t) + 1$ with initial condition $x(0) = 0$.

Problem 1.3. Find the general solution of the differential equation $x'(t) = x(t)\,(1 - x(t))$. This differential is related to the logistic growth model.

Problem 1.4. Find the general solution of the differential equation $x'(t) = x(t)\,\log\frac{1}{x(t)}$. This equation describes the Gompertz growth model.

Problem 1.5. Let $x(t)$ be the solution of the differential equation $x'(t) = \cos x(t)$ with initial condition $x(0) = 0$.
(i) Using separation of variables, show that

$$\log(1 + \sin x(t)) - \log(1 - \sin x(t)) = 2t.$$

(Hint: Write $\frac{2}{\cos x} = \frac{\cos x}{1+\sin x} + \frac{\cos x}{1-\sin x}$.)
(ii) Show that

$$x(t) = \arcsin\Big(\frac{e^t - e^{-t}}{e^t + e^{-t}}\Big) = \arctan(e^t) - \arctan(e^{-t}).$$

Chapter 2

Systems of linear differential equations

2.1. The exponential of a matrix

Let $A \in \mathbb{C}^{n \times n}$ be an $n \times n$ matrix. The operator norm of A is defined by

$$\|A\|_{\mathrm{op}} = \sup_{x \in \mathbb{C}^n,\, \|x\| \leq 1} \|Ax\|.$$

It is straightforward to verify that the operator norm is submultiplicative; that is,

$$\|AB\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} \, \|B\|_{\mathrm{op}}.$$

Iterating this estimate, we obtain

$$\|A^k\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}}^k$$

for every nonnegative integer k. This implies that the sequence

$$\sum_{k=0}^{m} \frac{1}{k!} \, A^k$$

is a Cauchy sequence in $\mathbb{C}^{n \times n}$. Its limit

$$\exp(A) := \lim_{m \to \infty} \sum_{k=0}^{m} \frac{1}{k!} \, A^k$$

is referred to as the matrix exponential of A.

Proposition 2.1. *Let $A, B \in \mathbb{C}^{n \times n}$ be two $n \times n$ matrices satisfying $AB = BA$. Then*

$$\exp(A + B) = \exp(A) \, \exp(B).$$

Proof. Since A and B commute, we obtain

$$(A+B)^l = \sum_{j=0}^{l} \binom{l}{j} A^j\, B^{l-j},$$

hence

$$\frac{1}{l!}\,(A+B)^l = \sum_{j=0}^{l} \frac{1}{j!\,(l-j)!}\, A^j\, B^{l-j}.$$

Summation over l gives

$$\sum_{l=0}^{2m} \frac{1}{l!}\,(A+B)^l = \sum_{j,k\geq 0,\, j+k\leq 2m} \frac{1}{j!\,k!}\, A^j\, B^k.$$

From this, we deduce that

$$\begin{aligned} &\sum_{k=0}^{2m} \frac{1}{k!}\,(A+B)^k - \Big(\sum_{j=0}^{m} \frac{1}{j!}\, A^j\Big) \Big(\sum_{k=0}^{m} \frac{1}{k!}\, B^k\Big) \\ &= \sum_{j,k\geq 0,\, j+k\leq 2m,\, \max\{j,k\}>m} \frac{1}{j!\,k!}\, A^j\, B^k. \end{aligned}$$

This gives

$$\begin{aligned} &\Big\|\sum_{l=0}^{2m} \frac{1}{l!}\,(A+B)^l - \Big(\sum_{j=0}^{m} \frac{1}{j!}\, A^j\Big) \Big(\sum_{k=0}^{m} \frac{1}{k!}\, B^k\Big)\Big\|_{\text{op}} \\ &\leq \sum_{j,k\geq 0,\, j+k\leq 2m,\, \max\{j,k\}>m} \frac{1}{j!\,k!}\, \|A\|_{\text{op}}^j\, \|B\|_{\text{op}}^k \\ &\leq \Big(\sum_{j=m+1}^{\infty} \frac{1}{j!}\, \|A\|_{\text{op}}^j\Big) \Big(\sum_{k=m+1}^{\infty} \frac{1}{k!}\, \|B\|_{\text{op}}^k\Big), \end{aligned}$$

and the right hand side converges to 0 as $m \to \infty$. From this, the assertion follows. ☐

Corollary 2.2. *For any matrix $A \in \mathbb{C}^{n\times n}$, the matrix* $\exp(A)$ *is invertible, and its inverse is given by* $\exp(-A)$.

Corollary 2.3. *We have* $\exp((s+t)A) = \exp(sA)\,\exp(tA)$ *for every matrix $A \in \mathbb{C}^{n\times n}$ and all $s,t \in \mathbb{R}$.*

In the remainder of this section, we derive an alternative formula for the exponential of a matrix. This formula is inspired by Euler's formula for the exponential of a number:

Proposition 2.4. *For every matrix $A \in \mathbb{C}^{n\times n}$, we have*

$$\exp(A) = \lim_{m\to\infty} \Big(I + \frac{1}{m}\, A\Big)^m.$$

Proof. We compute

$$\begin{aligned}
&\exp(A) - \Big(I + \frac{1}{m}\,A\Big)^m \\
&= \Big[\exp\Big(\frac{1}{m}\,A\Big)\Big]^m - \Big(I + \frac{1}{m}\,A\Big)^m \\
&= \sum_{l=0}^{m-1} \Big[\exp\Big(\frac{1}{m}\,A\Big)\Big]^{m-l-1} \Big[\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big] \Big(I + \frac{1}{m}\,A\Big)^l.
\end{aligned}$$

This gives

$$\begin{aligned}
&\Big\|\exp(A) - \Big(I + \frac{1}{m}\,A\Big)^m\Big\|_{\text{op}} \\
&\leq \sum_{l=0}^{m-1} \Big\|\exp\Big(\frac{1}{m}\,A\Big)\Big\|_{\text{op}}^{m-l-1} \Big\|\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big\|_{\text{op}} \Big\|I + \frac{1}{m}\,A\Big\|_{\text{op}}^l \\
&\leq \sum_{l=0}^{m-1} e^{\frac{m-l-1}{m}\,\|A\|_{\text{op}}} \Big\|\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big\|_{\text{op}} \Big(1 + \frac{1}{m}\,\|A\|_{\text{op}}\Big)^l \\
&\leq \sum_{l=0}^{m-1} e^{\frac{m-l-1}{m}\,\|A\|_{\text{op}}} \Big\|\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big\|_{\text{op}}\, e^{\frac{l}{m}\,\|A\|_{\text{op}}} \\
&= m\, e^{\frac{m-1}{m}\,\|A\|_{\text{op}}} \Big\|\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big\|_{\text{op}}.
\end{aligned}$$

On the other hand,

$$\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A = \sum_{k=2}^{\infty} \frac{1}{k!}\,\frac{1}{m^k}\,A^k,$$

hence

$$\Big\|\exp\Big(\frac{1}{m}\,A\Big) - I - \frac{1}{m}\,A\Big\|_{\text{op}} \leq \sum_{k=2}^{\infty} \frac{1}{k!}\,\frac{1}{m^k}\,\|A\|_{\text{op}}^k \leq \frac{1}{m^2}\,\|A\|_{\text{op}}^2\, e^{\frac{1}{m}\,\|A\|_{\text{op}}}.$$

Putting these facts together, we conclude that

$$\Big\|\exp(A) - \Big(I + \frac{1}{m}\,A\Big)^m\Big\|_{\text{op}} \leq \frac{1}{m}\,\|A\|_{\text{op}}^2\, e^{\|A\|_{\text{op}}}.$$

From this, the assertion follows easily. □

2.2. Calculating the matrix exponential of a diagonalizable matrix

In this section, we consider a matrix $A \in \mathbb{C}^{n\times n}$ which is diagonalizable. In other words, there exists an invertible matrix $S \in \mathbb{C}^{n\times n}$ and a diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{bmatrix}$$

such that $A = SDS^{-1}$. Equivalently, a matrix A is diagonalizable if there exists a basis of $\mathbb{C}^n$ which consists of eigenvectors of A.

In order to compute the exponential of such a matrix we need two auxiliary results. The first one relates the matrix exponentials of two matrices that are similar to each other.

Proposition 2.5. *Suppose that $A, B \in \mathbb{C}^{n\times n}$ are similar, so that $A = SBS^{-1}$ for some invertible matrix $S \in \mathbb{C}^{n\times n}$. Then $\exp(tA) = S\exp(tB)S^{-1}$ for all $t \in \mathbb{R}$.*

Proof. Using induction on k, it is easy to show that $A^k = SB^kS^{-1}$ for all integers $k \geq 0$. Consequently,

$$\exp(tA) = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} SB^kS^{-1} = S\exp(tB)S^{-1}.$$

This completes the proof. □

The second result gives a formula for the exponential of a diagonal matrix:

Proposition 2.6. *Suppose that*

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{bmatrix}$$

is a diagonal matrix. Then

$$\exp(tD) = \begin{bmatrix} e^{t\lambda_1} & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & e^{t\lambda_n} \end{bmatrix}.$$

Proof. Using induction on k, we can show that

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n^k \end{bmatrix}$$

for every integer $k \geq 0$. Summation over k gives

$$\exp(tD) = \sum_{k=0}^{\infty} \frac{t^k}{k!} D^k = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_1^k & 0 & \dots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_n^k \end{bmatrix}.$$

From this, the assertion follows. □

To summarize, if $A = SDS^{-1}$ is a diagonalizable matrix, then its matrix exponential is given by

$$\exp(tA) = S \exp(tD) S^{-1} = S \begin{bmatrix} e^{t\lambda_1} & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & e^{t\lambda_n} \end{bmatrix} S^{-1}.$$

As an example, let us consider the matrix

$$A = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix},$$

where $\alpha, \beta \in \mathbb{R}$. The matrix A has eigenvalues $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$. Moreover, the vectors

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

and

$$v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

are eigenvectors of A with eigenvalues λ_1 and λ_2. Consequently, A is diagonalizable and

$$A = S \begin{bmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{bmatrix} S^{-1},$$

where

$$S = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

Therefore, the matrix exponential of A is given by

$$\begin{aligned}
\exp(tA) &= S \begin{bmatrix} e^{t(\alpha+i\beta)} & 0 \\ 0 & e^{t(\alpha-i\beta)} \end{bmatrix} S^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} e^{t(\alpha+i\beta)} & 0 \\ 0 & e^{t(\alpha-i\beta)} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & \frac{i}{2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2}\left(e^{t(\alpha+i\beta)} + e^{t(\alpha-i\beta)}\right) & -\frac{i}{2}\left(e^{t(\alpha+i\beta)} + e^{t(\alpha-i\beta)}\right) \\ \frac{i}{2}\left(e^{t(\alpha+i\beta)} + e^{t(\alpha-i\beta)}\right) & \frac{1}{2}\left(e^{t(\alpha+i\beta)} + e^{t(\alpha-i\beta)}\right) \end{bmatrix} \\
&= \begin{bmatrix} e^{\alpha t} \cos(\beta t) & e^{\alpha t} \sin(\beta t) \\ -e^{\alpha t} \sin(\beta t) & e^{\alpha t} \cos(\beta t) \end{bmatrix}.
\end{aligned}$$

2.3. Generalized eigenspaces and the $L + N$ decomposition

In order to compute the exponential of a matrix that is not diagonalizable, it will be necessary to consider decompositions of $\mathbb{C}^n$ into generalized eigenspaces. We will need the following theorem due to Cayley and Hamilton:

Theorem 2.7. *Let A be a $n \times n$ matrix, and let $p_A(\lambda) = \det(\lambda I - A)$ denote the characteristic polynomial of A. Then $p_A(A) = 0$.*

Proof. The proof involves several steps.

Step 1: Suppose first that A is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, i.e.

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \lambda_n \end{bmatrix}.$$

Then

$$p(A) = \begin{bmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & p(\lambda_n) \end{bmatrix}$$

for every polynomial p. In particular, if $p = p_A$ is the characteristic polynomial of A, then $p_A(\lambda_j) = 0$ for all j, hence $p_A(A) = 0$.

Step 2: Suppose next that A is an upper triangular matrix whose diagonal entries are pairwise distinct. In this case, A has n distinct eigenvalues. In particular, A is diagonalizable. Hence, we can find a diagonal matrix B and an invertible matrix S such that $A = SBS^{-1}$. Clearly, A and B have the same characteristic polynomial, so $p_A(A) = p_B(A) = Sp_B(B)S^{-1} = 0$ by Step 1.

Step 3: Suppose now that A is a arbitrary upper triangular matrix. We can find a sequence of matrices A_k such that $\lim_{k\to\infty} A_k = A$ and each matrix A_k is upper triangular with n distinct diagonal entries. This implies $p_A(A) = \lim_{k\to\infty} p_{A_k}(A_k) = 0$.

Step 4: Finally, if A is a general $n \times n$ matrix, we can find an upper triangular matrix B such that $A = SBS^{-1}$. Again, A and B have the same characteristic polynomial, so we obtain $p_A(A) = p_B(A) = S p_B(B) S^{-1} = 0$ by Step 3. □

We will also need the following tool from algebra:

Proposition 2.8. *Suppose that $f(\lambda)$ and $g(\lambda)$ are two polynomials that are relatively prime. (This means that any polynomial that divides both $f(\lambda)$ and $g(\lambda)$ must be constant, i.e. of degree 0.) Then we can find polynomials $p(\lambda)$ and $q(\lambda)$ such that $p(\lambda)\, f(\lambda) + q(\lambda)\, g(\lambda) = 1$.*

This is standard result in algebra. The polynomials $p(\lambda)$ and $q(\lambda)$ can be found using the Euclidean algorithm. A proof can be found in most algebra textbooks.

Proposition 2.9. *Let A be an $n \times n$ matrix, and let $f(\lambda)$ and $g(\lambda)$ be two polynomials that are relatively prime. Moreovr, let x be a vector satisfying $f(A)\, g(A)\, x = 0$. Then there exists a unique pair of vectors y, z such that $f(A)\, y = 0$, $g(A)\, z = 0$, and $y + z = x$. In other words,* $\ker(f(A)\, g(A)) = \ker f(A) \oplus \ker g(A)$.

Proof. Since the polynomials $f(\lambda)$ and $g(\lambda)$ are relatively prime, we can find polynomials $p(\lambda)$ and $q(\lambda)$ such that

$$p(\lambda)\, f(\lambda) + q(\lambda)\, g(\lambda) = 1.$$

This implies

$$p(A)\, f(A) + q(A)\, g(A) = I.$$

In order to prove the existence part, we define vectors y, z by $y = q(A)\, g(A)\, x$ and $z = p(A)\, f(A)\, x$. Then

$$f(A)\, y = f(A)\, q(A)\, g(A)\, x = q(A)\, f(A)\, g(A)\, x = 0,$$
$$g(A)\, z = g(A)\, p(A)\, f(A)\, x = p(A)\, f(A)\, g(A)\, x = 0,$$

and

$$y + z = (p(A)\, f(A) + q(A)\, g(A))\, x = x.$$

Therefore, the vectors y, z have all the required properties.

In order to prove the uniqueness part, it suffices to show that $\ker f(A) \cap \ker g(A) = \{0\}$. Assume that x lies in the intersection of $\ker f(A)$ and $\ker g(A)$, so that $f(A)\, x = 0$ and $g(A)\, x = 0$. This implies $p(A)\, f(A)\, x = 0$ and $q(A)\, g(A)\, x = 0$. Adding both equations, we obtain $x = (p(A)\, f(A) +$

$q(A)\,g(A))\,x = 0$. This shows that show that $\ker f(A) \cap \ker g(A) = \{0\}$, as claimed. □

Corollary 2.10. *Let A be a $n \times n$ matrix, and denote by $p_A(\lambda) = \det(\lambda I - A)$ the characteristic polynomial of A. Let us write the polynomial $p_A(\lambda)$ in the form*

$$p_A(\lambda) = (\lambda - \lambda_1)^{\nu_1} \cdots (\lambda - \lambda_m)^{\nu_m},$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of A and $\nu_1, \ldots, \nu_m$ denote their respective algebraic multiplicities. Then we have the direct sum decomposition

$$\mathbb{C}^n = \ker(A - \lambda_1 I)^{\nu_1} \oplus \ldots \oplus (A - \lambda_m I)^{\nu_m}.$$

The spaces $\ker(A - \lambda_j I)^{\nu_j}$ are referred to as the generalized eigenspaces of A.

Proof. For abbreviation, let $g_j(\lambda) = (\lambda - \lambda_j)^{\nu_j}$, so that

$$p_A(\lambda) = g_1(\lambda) \cdots g_m(\lambda).$$

For each $k \in \{1, \ldots, m\}$, the polynomials $g_1(\lambda) \cdots g_{k-1}(\lambda)$ and $g_k(\lambda)$ are relatively prime, as they have no roots in common. Consequently,

$$\ker(g_1(A) \cdots g_{k-1}(A) g_k(A)) = \ker(g_1(A) \cdots g_{k-1}(A)) \oplus \ker g_k(A).$$

Repeated application of this result yields the direct sum decomposition

$$\ker p_A(A) = \ker g_1(A) \oplus \ldots \oplus \ker g_m(A).$$

On the other hand, $p_A(A) = 0$ by the Cayley-Hamilton theorem, so that $\ker p_A(A) = \mathbb{C}^n$. As a result, we obtain the following decomposition of $\mathbb{C}^n$ into generalized eigenspaces:

$$\mathbb{C}^n = \ker g_1(A) \oplus \ldots \oplus \ker g_m(A).$$

□

Theorem 2.11. *Let $A \in \mathbb{C}^{n \times n}$ be given. Then we can find matrices $L, N \in \mathbb{C}^n$ with the following properties:*

(i) *$L + N = A$.*

(ii) *$LN = NL$.*

(iii) *L is diagonalizable.*

(iv) *N is nilpotent, i.e. $N^n = 0$.*

Moreover, the matrices L and N are unique.

Proof. We first prove the existence statement. Consider the decomposition of $\mathbb{C}^n$ into generalized eigenspaces:

$$\mathbb{C}^n = \ker(A - \lambda_1 I)^{\nu_1} \oplus \ldots \oplus (A - \lambda_m I)^{\nu_m}.$$

Consider the linear transformation from $\mathbb{C}^n$ into itself that sends a vector $x \in \ker(A - \lambda_j I)^{\nu_j}$ to $\lambda_j x$ ($j = 1, \ldots, m$). Let L be the $n \times n$ matrix associated with this linear transformation. This implies $Lx = \lambda_j x$ for all $x \in \ker(A - \lambda_j I)^{\nu_j}$. Clearly, $\ker(L - \lambda_j I) = \ker(A - \lambda_j I)^{\nu_j}$ for $j = 1, \ldots, m$. Therefore, there exists a basis of $\mathbb{C}^n$ that consists of eigenvectors of L. Consequently, L is diagonalizable.

We claim that A and L commute, i.e. $LA = AL$. It suffices to show that $LAx = ALx$ for all vectors $x \in \ker(A - \lambda_j I)^{\nu_j}$ and all $j = 1, \ldots, m$. Indeed, if x belongs to the generalized eigenspace $\ker(A - \lambda_j I)^{\nu_j}$, then Ax lies in the same generalized eigenspace. Therefore, $Lx = \lambda_j x$ and $LAx = \lambda_j Ax$. Putting these facts together, we obtain $LAx = \lambda_j Ax = ALx$, as claimed. Therefore, $LA = AL$.

We now put $N = A - L$. Clearly, $L + N = A$ and $LN = LA - L^2 = AL - L^2 = NL$. Hence, it remains to show that $N^n = 0$. As above, it is enough to show that $N^n x = 0$ for all vectors $x \in \ker(A - \lambda_j I)^{\nu_j}$ and all $j = 1, \ldots, m$. By definition of L and N, we have $Nx = Ax - Lx = (A - \lambda_j I)x$ for all $x \in \ker(A - \lambda_j I)^{\nu_j}$. From this it is easy to see that $N^n x = (A - \lambda_j I)^n x$. However, $(A - \lambda_j I)^n x = 0$ since $x \in \ker(A - \lambda_j I)^{\nu_j}$ and $\nu_j \leq n$. Thus, we conclude that $N^n x = 0$ for all $x \in \ker(A - \lambda_j I)^{\nu_j}$. This completes the proof of the existence part.

We next turn to the proof of the uniqueness statement. Suppose that $L, N \in \mathbb{C}^{n \times n}$ satsify (i) – (iv). We claim that $Lx = \lambda_j x$ for all vectors $x \in \ker(A - \lambda_j I)^{\nu_j}$ and all $j = 1, \ldots, m$. To this end, we use the formula $L - \lambda_j I = (A - \lambda_j I) - N$. Since N commutes with $A - \lambda_j I$, it follows that

$$(L - \lambda_j I)^{2n} = \sum_{l=0}^{2n} \binom{2n}{l} (-N)^l \, (A - \lambda_j I)^{2n-l}.$$

Using the identity $N^n = 0$, we obtain

$$(L - \lambda_j I)^{2n} = \sum_{l=0}^{n-1} \binom{2n}{l} (-N)^l \, (A - \lambda_j I)^{2n-l}.$$

Suppose that $x \in \ker(A - \lambda_j I)^{\nu_j}$. Since $\nu_j \leq n$, we have $(A - \lambda_j I)^{2n-l} x = 0$ for all $l = 0, \ldots, n-1$. This implies $(L - \lambda_j I)^{2n} x = 0$. Since L is diagonalizable, we it follows that $(L - \lambda_j I)x = 0$. Thus, we conclude that $Lx = \lambda_j x$ for all vectors $x \in \ker(A - \lambda_j I)^{\nu_j}$ and all $j = 1, \ldots, m$.

Since

$$\mathbb{C}^n = \ker(A - \lambda_1 I)^{\nu_1} \oplus \ldots \oplus (A - \lambda_m I)^{\nu_m},$$

there is exactly one matrix L such that $Lx = \lambda_j x$ for $x \in \ker(A - \lambda_j I)^{\nu_j}$ and $j = 1, \ldots, m$. This completes the proof of the uniqueness statement. $\square$

As an example, let us compute the $L+N$ decomposition of the matrix

$$A = \begin{bmatrix} 0 & 2 & 3 & -4 \\ 0 & 1 & 2 & -4 \\ -1 & -2 & -4 & 6 \\ 0 & -1 & -2 & 3 \end{bmatrix}.$$

We begin by computing the eigenvalues and eigenvectors of A. The characteristic polynomial of A is given by

$$\begin{aligned} \det(\lambda I - A) &= \det \begin{bmatrix} \lambda & -2 & -3 & 4 \\ 0 & \lambda-1 & -2 & 4 \\ 1 & 2 & \lambda+4 & -6 \\ 0 & 1 & 2 & \lambda-3 \end{bmatrix} \\ &= \lambda \det \begin{bmatrix} \lambda-1 & -2 & 4 \\ 2 & \lambda+4 & -6 \\ 1 & 2 & \lambda-3 \end{bmatrix} + \det \begin{bmatrix} -2 & -3 & 4 \\ \lambda-1 & -2 & 4 \\ 1 & 2 & \lambda-3 \end{bmatrix} \\ &= \lambda(\lambda^3-\lambda) + (3\lambda^2+1) \\ &= \lambda^4 + 2\lambda^2 + 1 \\ &= (\lambda-i)^2\,(\lambda+i)^2. \end{aligned}$$

Thus, the eigenvalues of A are i and $-i$, and they both have algebraic multiplicity 2. A straightforward calculation shows that the generalized eigenspaces are given by

$$\ker(A-iI)^2 = \ker \begin{bmatrix} -4 & -4i & -6i & -2+8i \\ -2 & -2i & 2-4i & -4+8i \\ 4+2i & -2+4i & -4+8i & 6-12i \\ 2 & 2i & 4i & -6i \end{bmatrix} = \operatorname{span}\{v_1, v_2\}$$

and

$$\ker(A+iI)^2 = \ker \begin{bmatrix} -4 & 4i & 6i & -2-8i \\ -2 & 2i & 2+4i & -4-8i \\ 4-2i & -2-4i & -4-8i & 6+12i \\ 2 & -2i & -4i & 6i \end{bmatrix} = \operatorname{span}\{v_3, v_4\},$$

where v_1, v_2, v_3, v_4 are defined by

$$v_1 = \begin{bmatrix} 1 \\ i \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 2+i \\ 0 \\ -2+i \\ -1 \end{bmatrix},$$

$$v_3 = \begin{bmatrix} 1 \\ -i \\ 0 \\ 0 \end{bmatrix}, \qquad v_4 = \begin{bmatrix} 2-i \\ 0 \\ -2-i \\ -1 \end{bmatrix}.$$

Thus, we conclude that $Lv_1 = iv_1$, $Lv_2 = iv_2$, $Lv_3 = -iv_3$, $Lv_4 = -iv_4$.

Let S be the 4×4-matrix with column vectors v_1, v_2, v_3, v_4:

$$S = \begin{bmatrix} 1 & 2+i & 1 & 2-i \\ i & 0 & -i & 0 \\ 0 & -2+i & 0 & -2-i \\ 0 & -1 & 0 & -1 \end{bmatrix}.$$

The inverse of S is given by

$$S^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{i}{2} & -\frac{1}{2} & 2 \\ 0 & 0 & -\frac{i}{2} & -\frac{1-2i}{2} \\ \frac{1}{2} & \frac{i}{2} & -\frac{1}{2} & 2 \\ 0 & 0 & \frac{i}{2} & -\frac{1+2i}{2} \end{bmatrix}.$$

Using the identities $Lv_1 = iv_1$, $Lv_2 = iv_2$, $Lv_3 = -iv_3$, $Lv_4 = -iv_4$, we obtain

$$L = S \begin{bmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} S^{-1} = \begin{bmatrix} 0 & 1 & 2 & -3 \\ -1 & 0 & 1 & -4 \\ 0 & 0 & -2 & 5 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Consequently,

$$N = A - L = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & -2 & -2 & 1 \\ 0 & -1 & -1 & 1 \end{bmatrix}.$$

It is easy to check that $LN = NL$ and $N^2 = 0$, as expected.

2.4. Calculating the exponential of a general $n \times n$ matrix

Let now A be an arbitrary $n \times n$ matrix. By Theorem 2.11, we may write $A = L + N$, where L is diagonalizable, N is nilpotent, and $LN = NL$. This gives

$$\exp(tA) = \exp(tL)\, \exp(tN).$$

Since L is diagonalizable, we may write $L = SDS^{-1}$, where D is a diagonal matrix. This gives

$$\exp(tL) = S\exp(tD)S^{-1}.$$

On the other hand,

$$\exp(tN) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\,N^k = \sum_{k=0}^{n-1} \frac{t^k}{k!}\,N^k$$

since $N^n = 0$. Putting these facts together, we obtain

$$\exp(tA) = S\exp(tD)S^{-1} \sum_{k=0}^{n-1} \frac{t^k}{k!}\,N^k.$$

Since the exponential of a diagonal matrix is trivial to compute, we thus obtain an explicit formula for $\exp(tA)$.

As an example, let us compute the matrix exponential of the matrix

$$A = \begin{bmatrix} 0 & 2 & 3 & -4 \\ 0 & 1 & 2 & -4 \\ -1 & -2 & -4 & 6 \\ 0 & -1 & -2 & 3 \end{bmatrix}.$$

We have seen above that

$$L = \begin{bmatrix} 0 & 1 & 2 & -3 \\ -1 & 0 & 1 & -4 \\ 0 & 0 & -2 & 5 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

and

$$N = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & -2 & -2 & 1 \\ 0 & -1 & -1 & 1 \end{bmatrix}.$$

Moreover,

$$L = S \begin{bmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} S^{-1},$$

where

$$S = \begin{bmatrix} 1 & 2+i & 1 & 2-i \\ i & 0 & -i & 0 \\ 0 & -2+i & 0 & -2-i \\ 0 & -1 & 0 & -1 \end{bmatrix}.$$

This gives

$$\exp(tL) = S \begin{bmatrix} e^{it} & 0 & 0 & 0 \\ 0 & e^{it} & 0 & 0 \\ 0 & 0 & e^{-it} & 0 \\ 0 & 0 & 0 & e^{-it} \end{bmatrix} S^{-1}.$$

Furthermore,

$$\exp(tN) = I + tN = \begin{bmatrix} 1 & t & t & -t \\ t & 1+t & t & 0 \\ -t & -2t & 1-2t & t \\ 0 & -t & -t & 1+t \end{bmatrix}$$

since $N^2 = 0$. Putting these facts together, we obtain

$$\exp(tA) = S \begin{bmatrix} e^{it} & 0 & 0 & 0 \\ 0 & e^{it} & 0 & 0 \\ 0 & 0 & e^{-it} & 0 \\ 0 & 0 & 0 & e^{-it} \end{bmatrix} S^{-1} \begin{bmatrix} 1 & t & t & -t \\ t & 1+t & t & 0 \\ -t & -2t & 1-2t & t \\ 0 & -t & -t & 1+t \end{bmatrix}.$$

2.5. Solving systems of linear differential equations using matrix exponentials

In this section, we consider a system of linear differential equations of the form

$$x_i'(t) = \sum_{j=1}^n a_{ij}\, x_j(t),$$

where $x_1(t), \ldots, x_n(t)$ are the unkown functions and a_{ij} is a given set of real numbers. It is convenient to rewrite this system in matrix form. To that end, we write

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

and

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}.$$

With this understood, the system can be rewritten as

$$x'(t) = Ax(t).$$

Proposition 2.12. *The function $t \mapsto \exp(tA)$ is differentiable, and*

$$\frac{d}{dt}\exp(tA) = A\exp(tA).$$

Proof. Using the identity

$$\frac{1}{h}\,(\exp(hA) - I) = \sum_{k=1}^{\infty} \frac{h^{k-1}}{k!}\,A^k,$$

we obtain

$$\lim_{h\to 0}\frac{1}{h}\,(\exp(hA) - I) = A.$$

This implies

$$\lim_{h\to 0}\frac{1}{h}\,(\exp((t+h)A) - \exp(tA)) = \lim_{h\to 0}\frac{1}{h}\,(\exp(hA) - I)\,\exp(tA) = A\exp(tA).$$

This proves the claim. □

Theorem 2.13. *Given any vector $x_0 \in \mathbb{R}^n$ and any matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique solution of the differential equation $x'(t) = Ax(t)$ with initial condition $x(0) = x_0$. Moreover, the solution can be expressed as $x(t) = \exp(tA)x_0$.*

Proof. Let $x(t)$ be a solution of the differential equation $x'(t) = Ax(t)$ with initial condition $x(0) = x_0$. Then

$$\frac{d}{dt}(\exp(-tA)\,x(t)) = \exp(-tA)\,x'(t) - \exp(-tA)\,Ax(t) = 0.$$

Consequently, the function $t \mapsto \exp(-tA)\,x(t)$ is constant. From this, we deduce that $\exp(-tA)\,x(t) = x_0$, hence $x(t) = \exp(tA)x_0$. This proves the uniqueness statement.

Conversely, if we define $x(t)$ by the formula $x(t) = \exp(tA)x_0$, then $x(0) = x_0$ and

$$x'(t) = \frac{d}{dt}\exp(tA)x_0 = A\exp(tA)x_0 = Ax(t).$$

Thus, $x(t)$ solves the given initial value problem. □

We next consider an inhomogeneous system of the form

$$x'(t) = Ax(t) + f(t),$$

where $f(t)$ is a given function which takes values in $\mathbb{R}^n$.

Theorem 2.14. *Let A be an $n \times n$ matrix, and let $f : I \to \mathbb{R}^n$ be a continuous function which is defined on an open interval I containing 0. Then* *there exists a unique solution of the differential equation $x'(t) = Ax(t)+f(t)$ with initial condition $x(0) = x_0$. Moreover, the solution can be expressed as*

$$x(t) = \exp(tA)\left(x_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right).$$

Proof. Let $x(t)$ be a solution of the differential equation $x'(t) = Ax(t)+f(t)$ with initial condition $x(0) = x_0$. Then

$$\frac{d}{dt}(\exp(-tA)\, x(t)) = \exp(-tA)\, x'(t) - \exp(-tA)\, Ax(t) = \exp(-tA)\, f(t).$$

Integrating this identity gives

$$\exp(-tA)\, x(t) = x_0 + \int_0^t \exp(-sA)\, f(s)\, ds,$$

hence

$$x(t) = \exp(tA)\left(x_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right).$$

This proves the uniqueness statement.

Conversely, if we define $x(t)$ by the formula

$$x(t) = \exp(tA)\left(x_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right),$$

then $x(0) = x_0$ and

$$\begin{aligned} x'(t) &= A\exp(tA)\left(x_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right) \\ &\quad + \exp(tA)\,\frac{d}{dt}\left(x_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right) \\ &= Ax(t) + f(t). \end{aligned}$$

Therefore, $x(t)$ is a solution of the given initial value problem. □

As an example, let us find the solution of the inhomogeneous system

$$x'(t) = \begin{bmatrix} 1 & -2 \\ 3 & 3 \end{bmatrix} x(t) + e^{2t}\begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

with initial condition

$$x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The coefficient matrix

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

has eigenvalues $\lambda_1 = 2+\sqrt{5}\,i$ and $\lambda_2 = 2-\sqrt{5}\,i$. The associated eigenvectors are given by

$$v_1 = \begin{bmatrix} -2 \\ 1+\sqrt{5}\,i \end{bmatrix}, \quad v_2 = \begin{bmatrix} -2 \\ 1-\sqrt{5}\,i \end{bmatrix}.$$

Hence, if we define

$$S = \begin{bmatrix} -2 & -2 \\ 1+\sqrt{5}\,i & 1-\sqrt{5}\,i \end{bmatrix},$$

then

$$A = S \begin{bmatrix} 2+\sqrt{5}\,i & 0 \\ 0 & 2-\sqrt{5}\,i \end{bmatrix} S^{-1}$$

and

$$\exp(tA) = S \begin{bmatrix} e^{(2+\sqrt{5}\,i)t} & 0 \\ 0 & e^{(2-\sqrt{5}\,i)t} \end{bmatrix} S^{-1}.$$

Moreover, we compute

$$\begin{aligned}
\exp(-tA)\,f(t) &= S \begin{bmatrix} e^{-\sqrt{5}\,i\,t} & 0 \\ 0 & e^{\sqrt{5}\,i\,t} \end{bmatrix} S^{-1} \begin{bmatrix} -2 \\ 3 \end{bmatrix} \\
&= \frac{1}{2\sqrt{5}\,i}\, S \begin{bmatrix} e^{-\sqrt{5}\,i\,t} & 0 \\ 0 & e^{\sqrt{5}\,i\,t} \end{bmatrix} \begin{bmatrix} 2+\sqrt{5}\,i \\ -2+\sqrt{5}\,i \end{bmatrix} \\
&= \frac{1}{2\sqrt{5}\,i}\, S \begin{bmatrix} e^{-\sqrt{5}\,i\,t}\,(2+\sqrt{5}\,i) \\ e^{\sqrt{5}\,i\,t}\,(-2+\sqrt{5}\,i) \end{bmatrix}.
\end{aligned}$$

This implies

$$\int_0^t \exp(-sA)\,f(s)\,ds = \frac{1}{10}\, S \begin{bmatrix} (e^{-\sqrt{5}\,i\,t}-1)\,(2+\sqrt{5}\,i) \\ (e^{\sqrt{5}\,i\,t}-1)\,(2-\sqrt{5}\,i) \end{bmatrix}.$$

Thus, the solution of the initial value problem is given by

$$\begin{aligned}
x(t) &= \exp(tA) \int_0^t \exp(-sA)\,f(s)\,ds \\
&= \frac{1}{10}\, S \begin{bmatrix} e^{(2+\sqrt{5}\,i)t} & 0 \\ 0 & e^{(2-\sqrt{5}\,i)t} \end{bmatrix} \begin{bmatrix} (e^{-\sqrt{5}\,i\,t}-1)\,(2+\sqrt{5}\,i) \\ (e^{\sqrt{5}\,i\,t}-1)\,(2-\sqrt{5}\,i) \end{bmatrix} \\
&= \frac{1}{10}\, S \begin{bmatrix} (e^{2t}-e^{(2+\sqrt{5}\,i)t})\,(2+\sqrt{5}\,i) \\ (e^{2t}-e^{(2-\sqrt{5}\,i)t})\,(2-\sqrt{5}\,i) \end{bmatrix} \\
&= \frac{1}{10} \begin{bmatrix} -(e^{2t}-e^{(2+\sqrt{5}\,i)t})\,(4+2\sqrt{5}\,i) - (e^{2t}-e^{(2-\sqrt{5}\,i)t})\,(4-2\sqrt{5}\,i) \\ -(e^{2t}-e^{(2+\sqrt{5}\,i)t})\,(3-3\sqrt{5}\,i) - (e^{2t}-e^{(2-\sqrt{5}\,i)t})\,(3+3\sqrt{5}\,i) \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} -4e^{2t} + 4e^{2t}\cos(\sqrt{5}\,t) - 2\sqrt{5}\,e^{2t}\sin(\sqrt{5}\,t) \\ -3e^{2t} + 3e^{2t}\cos(\sqrt{5}\,t) + 3\sqrt{5}\,e^{2t}\sin(\sqrt{5}\,t) \end{bmatrix}.
\end{aligned}$$

2.6. Asymptotic behavior of solutions

Proposition 2.15. *Assume that all eigenvalues of A have negative real part. Then $\exp(tA) \to 0$ as $t \to \infty$. In other words, every solution of the differential equation $x'(t) = Ax(t)$ converges to 0 as $t \to \infty$.*

Proof. Let $\lambda_1, \ldots, \lambda_m$ denote the eigenvalues of A, and let $\nu_1, \ldots, \nu_m$ denote their algebraic multiplicities. Suppose that $x \in \ker(A - \lambda_j I)^{\nu_j}$, so that x is a generalized eigenvector of A. Then

$$\exp(tA)x = e^{t\lambda_j} \exp(t(A - \lambda_j I))x = e^{t\lambda_j} \sum_{k=0}^{n-1} \frac{t^k}{k!} (A - \lambda_j I)^k x.$$

Since $\mathrm{Re}(\lambda_j) < 0$, it follows that

$$\|\exp(tA)x\| = e^{t\,\mathrm{Re}(\lambda_j)} \left\| \sum_{k=0}^{n-1} \frac{t^k}{k!} (A - \lambda_j I)^k x \right\| \to 0$$

as $t \to \infty$. Thus, $\lim_{t\to\infty} \exp(tA)x = 0$ whenever x is a generalized eigenvector of A. Since every vector in $\mathbb{C}^n$ can be written as a sum of generalized eigenvectors, we conclude that $\lim_{t\to\infty} \exp(tA)x = 0$ for all $x \in \mathbb{C}^n$. From this, the assertion follows easily. □

Proposition 2.16. *Assume that all eigenvalues of A have negative real part. Moreover, suppose that $f : \mathbb{R} \to \mathbb{R}^n$ is a continuous function which is periodic with period τ. Then there exists a unique solution $\bar{x}(t)$ of the differential equation $x'(t) = Ax(t) + f(t)$ which is periodic with period τ. Moreover, if $x(t)$ is another solution of the differential equation $x'(t) = Ax(t) + f(t)$, then $x(t) - \bar{x}(t) \to 0$ as $t \to \infty$.*

Proof. We first observe that the matrix $\exp(-\tau A) - I$ is invertible. Indeed, if 1 is an eigenvalue of the matrix $\exp(-\tau A)$, then 1 is an eigenvalue of the matrix $\exp(-m\tau A)$ for every integer m. This is impossible since $\exp(-m\tau A) \to 0$ as $m \to \infty$. Consequently, the matrix $\exp(-\tau A) - I$ is invertible.

We now define a vector $\bar{x}_0$ by

$$(\exp(-\tau A) - I)\,\bar{x}_0 = \int_0^\tau \exp(-sA)\, f(s)\, ds.$$

Moreover, we define

$$\bar{x}(t) = \exp(tA)\left(\bar{x}_0 + \int_0^t \exp(-sA)\, f(s)\, ds\right).$$

Since the function $f(t)$ is periodic with period τ, we have

$$\begin{aligned}
\bar{x}(t+\tau) &= \exp((t+\tau)A)\left(\bar{x}_0 + \int_0^{t+\tau} \exp(-sA)\, f(s)\, ds\right) \\
&= \exp((t+\tau)A)\left(\exp(-\tau A)\,\bar{x}_0 + \int_\tau^{t+\tau} \exp(-sA)\, f(s)\, ds\right) \\
&= \exp((t+\tau)A)\left(\exp(-\tau A)\,\bar{x}_0 + \int_0^{t} \exp(-(s+\tau)A)\, f(s+\tau)\, ds\right) \\
&= \exp(tA)\left(\bar{x}_0 + \int_0^{t} \exp(-sA)\, f(s+\tau)\, ds\right) \\
&= \bar{x}(t)
\end{aligned}$$

for all t. Therefore, the function $\bar{x}(t)$ is periodic with period τ. Moreover, by Theorem 2.14, the function $\bar{x}(t)$ satisfies $\bar{x}'(t) = A\bar{x}(t) + f(t)$ for all t. This proves that there is at least one periodic solution.

We next assume that $x(t)$ is an arbitrary solution of the differential equation $x'(t) = Ax(t) + f(t)$. Then the difference $x(t) - \bar{x}(t)$ satisfies the differential equation

$$\frac{d}{dt}(x(t) - \bar{x}(t)) = A(x(t) - \bar{x}(t)).$$

This implies

$$x(t) - \bar{x}(t) = \exp(tA)\,(x(0) - \bar{x}(0)),$$

and the right hand side converges to 0 as $t \to \infty$.

Finally, we show that $\bar{x}(t)$ is the only solution which is periodic with period τ. To prove this, suppose that $x(t)$ is another solution which is periodic with period τ. Then the function $x(t) - \bar{x}(t)$ is periodic with period τ. Since $x(t) - \bar{x}(t) \to 0$ as $t \to \infty$, we conclude that $x(t) - \bar{x}(t) = 0$ for all t. This completes the proof. □

In the remainder of this section, we analyze what happens when A has some eigenvalues with negative real part and some eigenvalues with positive real part. As usual, we denote by $\lambda_1, \ldots, \lambda_m$ the eigenvalues of A and by $\nu_1, \ldots, \nu_m$ their algebraic multiplicities. We assume that $\mathrm{Re}(\lambda_j) \neq 0$ for $j = 1, \ldots, m$. After rearranging the eigenvalues, we may assume that $\mathrm{Re}(\lambda_j) < 0$ for $j = 1, \ldots, l$ and $\mathrm{Re}(\lambda_j) > 0$ for $j = l+1, \ldots, m$. Let us define

$$p_-(\lambda) = \prod_{j=1}^{l} (\lambda - \lambda_j)^{\nu_j}$$

and

$$p_+(\lambda) = \prod_{j=l+1}^{m} (\lambda - \lambda_j)^{\nu_j}.$$

We claim that $p_-(\lambda)$ and $p_+(\lambda)$ are relatively prime. Suppose by contradiction that there is a polynomial $q(\lambda)$ of degree at least 1 which divides both $p_-(\lambda)$ and $p_+(\lambda)$. By the fundamental theorem of algebra, $q(\lambda)$ has at least one root in the complex plane, and this number must be a common root of $p_-(\lambda)$ and $p_+(\lambda)$. But this impossible since the roots of $p_-(\lambda)$ all have positive real part and the roots of $p_+(\lambda)$ all have negative real part. Thus, we conclude that $p_-(\lambda)$ and $p_+(\lambda)$ are relatively prime. Moreover, the product $p_-(\lambda)\, p_+(\lambda)$ is the characteristic polynomial of A. Using the Cayley-Hamilton, we conclude that

$$\mathbb{C}^n = \ker(p_-(A)\, p_+(A)) = \ker p_-(A) \oplus p_+(A).$$

Let P_- and P_+ denote the canonical projections associated with this direct sum decomposition. The subspaces $\ker p_-(A)$ and $\ker p_+(A)$ are referred to as the stable and unstable subspaces, and the projections P_- and P_+ are referred to as the spectral projections.

We now assume that A has real entries. Since the eigenvalues of a real matrix occur in pairs of complex conjugate numbers, the polynomials $p_-(\lambda)$ and $p_+(\lambda)$ have real coefficients. Thus, the subspaces $\ker p_-(A), \ker p_+(A) \subset \mathbb{C}^n$ are invariant under complex conjugation, and the projections P_- and P_+ are matrices with real entries.

Proposition 2.17. *Assume that A has no eigenvalues on the imaginary axis. Let us choose $\alpha > 0$ small enough such that $|\mathrm{Re}(\lambda_j)| > \alpha$ for all eigenvalues of A. If $x \in \ker p_-(A)$, then $e^{\alpha t} \exp(tA)x \to 0$ as $t \to \infty$. Similarly, if $x \in \ker p_+(A)$, then $e^{-\alpha t} \exp(tA)x \to 0$ as $t \to -\infty$.*

Proof. Suppose first that $x \in \ker(A - \lambda_j I)^{\nu_j}$ for some $j \in \{1, \ldots, l\}$. Then

$$\exp(tA)x = e^{t\lambda_j}\, \exp(t(A - \lambda_j I))x = e^{t\lambda_j} \sum_{k=0}^{n-1} \frac{t^k}{k!}\, (A - \lambda_j I)^k x.$$

Since $\mathrm{Re}(\lambda_j) < -\alpha$, it follows that

$$e^{\alpha t}\, \|\exp(tA)x\| = e^{t\,(\mathrm{Re}(\lambda_j)+\alpha)} \left\| \sum_{k=0}^{n-1} \frac{t^k}{k!}\, (A - \lambda_j I)^k x \right\| \to 0$$

as $t \to \infty$. Thus, $\lim_{t\to\infty} e^{\alpha t} \exp(tA)x = 0$ whenever $x \in \ker(A - \lambda_j I)^{\nu_j}$ and $j \in \{1, \ldots, l\}$. Since

$$\ker p_-(A) = \ker(A - \lambda_1 I)^{\nu_1} \oplus \ldots \oplus (A - \lambda_l I)^{\nu_l},$$

we conclude that $\lim_{t\to\infty} e^{\alpha t} \exp(tA)x = 0$ for all $x \in \ker p_-(A)$. An analogous argument shows that $\lim_{t\to-\infty} e^{-\alpha t} \exp(tA)x = 0$ for all $x \in \ker p_+(A)$. $\square$

Corollary 2.18. *Assume that A has no eigenvalues on the imaginary axis. Let us choose $\alpha > 0$ small enough such that $|\mathrm{Re}(\lambda_j)| > \alpha$ for all eigenvalues of A. Then there exist positive constants N and α such that*

$$e^{\alpha t} \exp(tA)\, P_- \to 0$$

as $t \to \infty$ and

$$e^{-\alpha t} \exp(tA)\, P_+ \to 0$$

as $t \to -\infty$.

2.7. Problems

Problem 2.1. Consider the following 3×3 matrix:

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}.$$

Find a diagonalizable matrix L such that $AL = LA$ and $A - L$ is nilpotent.

Problem 2.2. Consider the following 4×4 matrix:

$$A = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 2 & -1 & 1 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Find a diagonalizable matrix L such that $AL = LA$ and $A - L$ is nilpotent.

Problem 2.3. Find the general solution of the inhomogeneous differential equation

$$x'(t) = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 5t \end{bmatrix}.$$

Problem 2.4. Consider a matrix $A \in \mathbb{C}^{n \times n}$. Show that

$$\{\mu : \mu \text{ is an eigenvalue of } \exp(tA)\} = \{e^{t\lambda} : \lambda \text{ is an eigenvalue of } A\}.$$

This result is called the spectral mapping theorem. (Hint: It suffices to prove this in the special case A is an upper triangular matrix.)

Problem 2.5. Consider a matrix $A \in \mathbb{C}^{n \times n}$. As usual, let $\lambda_1, \ldots, \lambda_m$ denote the eigenvalues of A and by $\nu_1, \ldots, \nu_m$ their algebraic multiplicities. We assume that $\mathrm{Re}(\lambda_j) < 0$ for $j = 1, \ldots, k$; $\mathrm{Re}(\lambda_j) = 0$ for $j = k+1, \ldots, l$; and $\mathrm{Re}(\lambda_j) > 0$ for $j = l+1, \ldots, m$. Show that

$$\{x \in \mathbb{C}^n : \limsup_{t \to \infty} \|\exp(tA)x\| < \infty\} = \bigoplus_{j=1}^{k} \ker(A - \lambda_j I)^{\nu_j} \oplus \bigoplus_{j=k+1}^{l} \ker(A - \lambda_j I).$$

Problem 2.6. Consider a matrix $A \in \mathbb{R}^{n \times n}$. Suppose that all eigenvalues of A have negative real part. Show that the limit

$$\lim_{\tau \to \infty} \int_0^\tau \big[\exp(tA)\big]^T \exp(tA)\, dt =: S$$

exists. Show that the matrix S is positive definite, and

$$\langle Sx, Ax \rangle = -\frac{1}{2}\, \|x\|^2$$

for all $x \in \mathbb{R}^n$.

Problem 2.7. Consider a matrix $A \in \mathbb{C}^{n \times n}$. Suppose that every eigenvalue of A satisfies $|\lambda| < 1$. Show that $\lim_{k \to \infty} A^k = 0$.

Chapter 3

Nonlinear systems

3.1. Peano's existence theorem

The proof of Peano's theorem relies on a compactness theorem due to Arzelà and Ascoli:

Arzelà-Ascoli Theorem. *Let J be a compact interval, and let $x_k : J \to \mathbb{R}^n$ be a sequence of continuous functions. Assume that the functions $x_k(t)$ are uniformly bounded, so that*

$$\sup_{k \in \mathbb{N}} \sup_{t \in J} \|x_k(t)\| < \infty.$$

Suppose further that the sequence x_k is equicontinuous. This means that, given any $\varepsilon > 0$, there exists a real number $\delta > 0$ so that

$$\sup_{k \in \mathbb{N}} \sup_{s,t \in J,\, |s-t| \leq \delta} |x_k(s) - x_k(t)| \leq \varepsilon.$$

Then a subsequence of the original sequence $x_k(t)$ converges uniformly to some limit function $x(t)$.

Proof. Let $A = \{t_1, t_2, \ldots\}$ be a countable dense subset of J. Since the sequence $x_k(t_1)$ is bounded, the Bolzano-Weierstrass theorem implies that we can find a sequence of integers $\{k_{l,1} : l = 1, 2, \ldots\}$ going to infinity such that the limit $\lim_{l \to \infty} x_{k_{l,1}}(t_1)$ exists. Similarly, since the sequence $x_k(t_2)$ is bounded, there exists a subsequence $\{k_{l,2} : l = 1, 2, \ldots\}$ of the sequence $\{k_{l,1} : l = 1, 2, \ldots\}$ with the property that the limit $\lim_{l \to \infty} x_{k_{l,1}}(t_1)$ exists. Proceeding inductively, we obtain for each $m \in \mathbb{N}$ a sequence of integers $\{k_{l,m} : l = 1, 2, \ldots\}$ with the property that the limit $\lim_{l \to \infty} x_{k_{l,m}}(t_m)$ exists. Moreover, the sequence $\{k_{l,m} : l = 1, 2, \ldots\}$ is a subsequence of the previous sequence $\{k_{l,m-1} : l = 1, 2, \ldots\}$.

We now consider the diagonal sequence $\{k_{l,l} : l = 1, 2, \dots\}$. Given any integer $m \in \mathbb{N}$, we have $\{k_{l,l} : l = m, m+1, \dots\} \subset \{k_{l,m} : l = 1, 2, \dots\}$. Since the limit $\lim_{l\to\infty} x_{k_{l,m}}(t_m)$ exists, we conclude that the limit $\lim_{l\to\infty} x_{k_{l,l}}(t_m)$ exists as well. To summarize, the limit $\lim_{l\to\infty} x_{k_{l,l}}(t)$ exists for each $t \in A$.

We want to show that the functions $x_{k_{l,l}}(t)$ converge uniformly. Suppose that ε is a given positive real number. Since the functions $x_k(t)$ are equicontinuous, we can find a real number $\delta > 0$ so that

$$\sup_{k\in\mathbb{N}} \sup_{s,t\in J,\,|s-t|\leq\delta} |x_k(s) - x_k(t)| \leq \varepsilon.$$

Moreover, since the set A is dense, we can find a finite subset $E \subset A$ such that $J \subset \bigcup_{s\in E}^{m} (s-\delta, s+\delta)$. Finally, we can find a positive integer l_0 such that

$$\max_{s\in E} |x_{k_{\tilde{l},\tilde{l}}}(s) - x_{k_{l,l}}(s)| \leq \varepsilon$$

for all $\tilde{l} \geq l \geq l_0$. We claim that

$$|x_{k_{\tilde{l},\tilde{l}}}(t) - x_{k_{l,l}}(t)| \leq 3\varepsilon$$

for all $t \in J$ and $\tilde{l} \geq l \geq l_0$. Indeed, for each $t \in J$ there exists a number $s \in E$ such that $|s - t| < \delta$. This gives

$$\begin{aligned} &|x_{k_{\tilde{l},\tilde{l}}}(t) - x_{k_{l,l}}(t)| \\ &\leq |x_{k_{\tilde{l},\tilde{l}}}(s) - x_{k_{l,l}}(s)| + |x_{k_{l,l}}(s) - x_{k_{l,l}}(t)| + |x_{k_{\tilde{l},\tilde{l}}}(s) - x_{k_{\tilde{l},\tilde{l}}}(t)| \leq 3\varepsilon, \end{aligned}$$

as claimed.

In particular, for each $t \in J$, the sequence $x_{k_{l,l}}(t)$, $l = 1, 2, \dots$, is a Cauchy sequence. Consequently, the limit $x(t) = \lim_{l\to\infty} x_{k_{l,l}}(t)$ exists for each $t \in J$. Moreover,

$$|x(t) - x_{k_{l,l}}(t)| = \lim_{\tilde{l}\to\infty} |x_{k_{\tilde{l},\tilde{l}}}(t) - x_{k_{l,l}}(t)| \leq 3\varepsilon$$

for all $t \in J$ and all $l \geq l_0$. Since $\varepsilon > 0$ is arbitrary, the sequence $x_k(t)$ converges uniformly to $x(t)$. This completes the proof. □

Theorem 3.1. *Let $U \subset \mathbb{R}^n$ be an open set, and let $F : U \to \mathbb{R}^n$ be a continuous mapping. Fix a point $x_0 \in U$. Then the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$ admits a solution $x(t)$, which is defined on an interval $(-\delta, \delta)$. Here, δ is a positive real number which depends on x_0.*

Proof. Since U is open, we can find a positive real number $r > 0$ such that $B_{2r}(x_0) \subset U$. Let

$$M = \sup_{x\in B_r(x_0)} \|F(x)\|,$$

and suppose that δ is chosen such that $0 < \delta < \frac{r}{M}$.

For every positive integer k, we define a continuous function $x_k : [-\delta, \delta] \to B_r(x_0)$ as follows. For $t \in [-\frac{\delta}{k}, \frac{\delta}{k}]$, we define

$$x_k(t) = x_0 + t\,F(x_0).$$

Suppose now that $j \in \{1, \ldots, k-1\}$, and we have defined $x_k(t)$ for $t \in [-\frac{j\delta}{k}, \frac{j\delta}{k}]$. We then define

$$x_k(t) = x_k\Big(\frac{j\delta}{k}\Big) + \Big(t - \frac{j\delta}{k}\Big)\,F\Big(x_k\Big(\frac{j\delta}{k}\Big)\Big)$$

for $t \in [\frac{j\delta}{k}, \frac{(j+1)\delta}{k}]$ and

$$x_k(t) = x_k\Big(-\frac{j\delta}{k}\Big) + \Big(t + \frac{j\delta}{k}\Big)\,F\Big(x_k\Big(-\frac{j\delta}{k}\Big)\Big)$$

for $t \in [-\frac{(j+1)\delta}{k}, -\frac{j\delta}{k}]$. Using induction on j, one can show that

$$|x_k(t)| \leq M\,|t| < r$$

for all $t \in [-\frac{j\delta}{k}, \frac{j\delta}{k}]$. Thus, the function x_k maps the interval $[-\delta, \delta]$ into the ball $B_r(x_0)$, as claimed.

It follows immediately from the definition of x_k that

$$|x_k(s) - x_k(t)| \leq M\,|s - t|$$

for all $s, t \in [-\delta, \delta]$. Consequently, the functions $x_k(t)$ are equicontinuous. By the Arzelà-Ascoli theorem, there exists a sequence of positive integers $k_l \to \infty$ such that the functions $x_{k_l}(t)$ converge uniformly to some limit function $x(t)$ as $l \to \infty$. It is clear that x is a continuous function which maps the interval $[-\delta, \delta]$ into the closed ball of radius r centered at x_0.

We claim that $x(t)$ is a solution of the given initial value problem. To see this, we write

$$x_k(t) = x_0 + \int_0^t F(y_k(s))\,ds,$$

where the function $y_k(t)$ is defined by

$$y_k(t) = \begin{cases} x_0 & \text{if } t \in [-\frac{\delta}{k}, \frac{\delta}{k}] \\ x_k\big(\frac{j\delta}{k}\big) & \text{if } t \in (\frac{j\delta}{k}, \frac{(j+1)\delta}{k}] \text{ for some } j \in \{1, \ldots, k-1\} \\ x_k\big(-\frac{j\delta}{k}\big) & \text{if } t \in [-\frac{(j+1)\delta}{k}, -\frac{j\delta}{k}) \text{ for some } j \in \{1, \ldots, k-1\}. \end{cases}$$

It is easy to see that

$$\sup_{t \in [-\delta, \delta]} \|x_k(t) - y_k(t)\| \leq \frac{M}{k}.$$

This implies

$$\sup_{t \in [-\delta, \delta]} \|x(t) - y_{k_l}(t)\| \leq \sup_{t \in [-\delta, \delta]} \|x(t) - x_{k_l}(t)\| + \frac{M}{k_l} \to 0$$

as $l \to \infty$. In other words, the functions $y_{k_l}(t)$ converge uniformly to $x(t)$ as $l \to \infty$. Since continuous functions on compact sets are uniformly continuous, we conclude that the functions $F(y_{k_l}(t))$ converge uniformly to $F(x(t))$ as $l \to \infty$. In particular,

$$\begin{aligned} x(t) &= \lim_{l\to\infty} x_{k_l}(t) \\ &= \lim_{l\to\infty} \Big(x_0 + \int_0^t F(y_{k_l}(s))\, ds \Big) \\ &= x_0 + \int_0^t F(x(s))\, ds. \end{aligned}$$

Putting $t = 0$ gives $x(0) = x_0$. Moreover, since $x(t)$ is continuous, the function $F(x(t))$ is continuous as well. Thus, $x(t)$ is continuously differentiable with derivative $x'(t) = F(x(t))$. This completes the proof. □

It is instructive to consider the special case when $x'(t) = Ax(t)$ for some matrix $A \in \mathbb{R}^{n\times n}$. In this case, the approximating functions $x_k(t)$ satisfy

$$x_k\Big(\frac{j\delta}{k}\Big) = \Big(I + \frac{\delta}{k}\,A\Big)^j\, x_0$$

and

$$x_k\Big(-\frac{j\delta}{k}\Big) = \Big(I - \frac{\delta}{k}\,A\Big)^j\, x_0$$

for $j \in \{1, \ldots, k\}$. Using Proposition 2.4, it is not difficult to show that $x_k(t) \to \exp(tA)x_0$ as $k \to \infty$.

3.2. Existence theory via the method of Picard iterates

Theorem 3.2. *Let $U \subset \mathbb{R}^n$ be an open set, and let $F : U \to \mathbb{R}^n$ be a continuously differentiable mapping. Fix a point $x_0 \in U$. Then the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$ admits a solution $x(t)$, which is defined on an interval $(-\delta, \delta)$. Here, δ is a positive real number which depends on x_0.*

Since U is open, we can find a positive real number $r > 0$ such that $B_{3r}(x_0) \subset U$. Moreover, we define

$$M = \sup_{x\in B_{2r}(x_0)} \|F(x)\|$$

and

$$L = \sup_{x\in B_{2r}(x_0)} \|DF(x)\|_{\mathrm{op}}.$$

We now choose δ small enough so that

$$0 < \delta < \min\{r, \frac{r}{M}, \frac{1}{2L}\}.$$

The strategy is to construct a family of functions $x_k(t)$ defined on the interval $(-\delta, \delta)$ such that $x_0(t) = x_0$ and

$$x_{k+1}(t) = x_0 + \int_0^t F(x_k(s))\, ds$$

for all $t \in (-\delta, \delta)$. The functions $x_k(t)$ are called the Picard iterates.

Lemma 3.3. *The function $x_k(t)$ is well-defined, and*

$$\|x_k(t) - x_{k-1}(t)\| \leq 2^{-k+1}\, r$$

for all $t \in (-\delta, \delta)$.

Proof. We argue by induction on k. For $k = 1$, the assertion is trivial. We now assume that the assertion has been established for all integers less than or equal to k. In other words, $x_j(t)$ is defined for all $j \leq k$, and we have

$$\|x_j(t) - x_{j-1}(t)\| \leq 2^{-j+1}\, r$$

for all $t \in (-\delta, \delta)$ and all $j \leq k$. Summation over j gives

$$\max\{\|x_k(t) - x_0\|, \|x_{k-1}(t) - x_0\|\} \leq \sum_{j=1}^{k} \|x_j(t) - x_{j-1}(t)\| \leq \sum_{j=1}^{k} 2^{-j+1}\, r < 2r$$

for all $t \in (-\delta, \delta)$. Consequently, $x_k(t), x_{k-1}(t) \in B_{2r}(x_0) \subset U$. Therefore, $x_{k+1}(t)$ can be defined.

By definition of L, we have $\|F(\xi) - F(\eta)\| \leq L\, \|\xi - \eta\|$ whenever $\xi, \eta \in B_{2r}(x_0)$. Since $x_k(t), x_{k-1}(t) \in B_{2r}(x_0)$ for all $t \in (-\delta, \delta)$, we conclude that

$$\|F(x_k(t)) - F(x_{k-1}(t))\| \leq L\, \|x_k(t) - x_{k-1}(t)\| \leq 2^{-k+1}\, r\, L$$

for all $t \in (-\delta, \delta)$. Using the identities

$$x_{k+1}(t) = x_0 + \int_0^t F(x_k(s))\, ds$$

and

$$x_k(t) = x_0 + \int_0^t F(x_{k-1}(s))\, ds,$$

we obtain

$$x_{k+1}(t) - x_k(t) = \int_0^t (F(x_k(s)) - F(x_{k-1}(s)))\, ds.$$

This implies

$$\|x_{k+1}(t) - x_k(t)\| \leq 2^{-k+1}\, r\, L\, |t| \leq 2^{-k+1}\, r\, L\, \delta \leq 2^{-k}\, r$$

for all $t \in (-\delta, \delta)$. $\square$

Using Lemma 3.3, we obtain

$$\|x_{\tilde{k}}(t) - x_k(t)\| \leq \sum_{j=k+1}^{\tilde{k}} \|x_j(t) - x_{j-1}(t)\| \leq \sum_{j=k+1}^{\tilde{k}} 2^{-j+1}\, r \leq 2^{-k+1}\, r$$

for all $\tilde{k} \geq k \geq 1$ and all $t \in (-\delta, \delta)$. Thus, for each $t \in (-\delta, \delta)$, the sequence $x_k(t)$ is a Cauchy sequence. Let us define

$$x(t) := \lim_{k \to \infty} x_k(t).$$

Then

$$\|x(t) - x_k(t)\| = \lim_{\tilde{k} \to \infty} \|x_{\tilde{k}}(t) - x_k(t)\| \leq 2^{-k+1}\, r$$

for all $t \in (-\delta, \delta)$. Thus, the sequence $x_k(t)$ converges uniformly to $x(t)$. In particular, the limit function $x(t)$ is continuous. Moreover,

$$\begin{aligned} x(t) &= \lim_{k \to \infty} x_{k+1}(t) \\ &= \lim_{k \to \infty} \left(x_0 + \int_0^t F(x_k(s))\, ds \right) \\ &= x_0 + \int_0^t F(x(s))\, ds \end{aligned}$$

for all $t \in (-\delta, \delta)$. Putting $t = 0$ gives $x(0) = x_0$. Moreover, since $x(t)$ is continuous, the function $F(x(t))$ is continuous as well. Thus, $x(t)$ is continuously differentiable with derivative $x'(t) = F(x(t))$. This completes the proof of Theorem 3.2.

As an example, let us consider the differential equation $x'(t) = Ax(t)$ where A is an $n \times n$ matrix. In this case, the Picard iterates satisfy

$$x_{k+1}(t) = x_0 + \int_0^t Ax_k(s)\, ds.$$

Using induction on k, one can show that

$$x_k(t) = \sum_{l=0}^{k} \frac{t^l}{l!}\, A^l x_0.$$

Thus, $x_k(t) \to \exp(tA)x_0$ as expected.

3.3. Uniqueness and the maximal time interval of existence

Theorem 3.4. *Let $U \subset \mathbb{R}^n$ be an open set, and let $F : U \to \mathbb{R}^n$ be a continuously differentiable mapping. Suppose that $x(t)$ and $y(t)$ are two solutions of the differential equation $x'(t) = F(x(t))$. Moreover, suppose that $x(t)$ is defined on some open interval I and $y(t)$ is defined on some* *open interval J. If $0 \in I \cap J$ and $x(0) = y(0)$, then $x(t) = y(t)$ for all $t \in I \cap J$.*

Proof. We first show that $x(t) = y(t)$ for all $t \in I \cap J \cap (0, \infty)$. Suppose that this false. Let $\tau = \inf\{t \in I \cap J \cap (0, \infty) : x(t) \neq y(t)\}$. Clearly, $x(\tau) = y(\tau)$. For abbreviation, let $\bar{x} := x(\tau) = y(\tau)$. Let us fix a real number $r > 0$ such that $B_{2r}(\bar{x}) \subset U$. We can find a real number $\delta > 0$ such that $[\tau, \tau + \delta] \subset I \cap J$ and $x(t), y(t) \in B_r(\bar{x})$ for all $t \in [\tau, \tau + \delta]$. Hence, if we put $L = \sup_{x \in B_r(\bar{x})} \|DF(x)\|_{\mathrm{op}}$, then we obtain

$$\|F(x(t)) - F(y(t))\| \leq L \, \|x(t) - y(t)\|$$

for all $t \in [\tau, \tau + \delta]$. This implies

$$\begin{aligned} \frac{d}{dt}\|x(t) - y(t)\|^2 &= 2 \, \langle x(t) - y(t), x'(t) - y'(t) \rangle \\ &= 2 \, \langle x(t) - y(t), F(x(t)) - F(y(t)) \rangle \\ &\leq 2L \, \|x(t) - y(t)\|^2 \end{aligned}$$

for $t \in [\tau, \tau + \delta]$. Consequently, the function $t \mapsto e^{-2Lt} \, \|x(t) - y(t)\|^2$ is monotone decreasing on the interval $[\tau, \tau + \delta]$. Since $x(\tau) = y(\tau)$, it follows that $x(t) = y(t)$ for all $t \in [\tau, \tau + \delta]$. This contradicts the definition of τ. Thus, $x(t) = y(t)$ for all $t \in I \cap J \cap (0, \infty)$. An analogous argument shows that $x(t) = y(t)$ for all $t \in I \cap J \cap (-\infty, 0)$. □

In the following, we assume that $U \subset \mathbb{R}^n$ is an open set, and $F : U \to \mathbb{R}^n$ is continuously differentiable. Theorem 3.4 guarantees that, given any point $x_0 \in U$, there exists a unique maximal solution of the initial value problem

$$\begin{cases} x'(t) = F(x(t)) \\ x(0) = x_0. \end{cases} \tag{4}$$

To see this, we fix a point $x_0 \in U$ and define

$$\begin{aligned} J_{x_0}^\star = \{\tau \in \mathbb{R} : &\text{ there exists a solution of (4) which is} \\ &\text{defined on some open interval containing } \tau\}. \end{aligned}$$

In other words, $J_{x_0}^\star$ is the union of the domains of definition of all solutions of the initial value problem (4). Clearly, $J_{x_0}^\star$ is an open interval. Using Theorem 3.1 or Theorem 3.2, we conclude that $J_{x_0}^\star$ is non-empty. Moreover, by Theorem 3.4, two solutions of (4) agree on the intersection of their domains of definition. Hence, there exists a unique function $x : J_{x_0}^\star \to U$ which solves (4).

We next characterize the maximal interval $J_{x_0}^\star$ on which the solution is defined.

Theorem 3.5. *Assume that $U \subset \mathbb{R}^n$ is an open set, and $F : U \to \mathbb{R}^n$ is continuously differentiable. Fix a point $x_0 \in U$, and let $x(t)$ be the unique maximal solution of the initial value problem (4), and let $J^\star_{x_0} = (\alpha, \beta)$ denote its domain of definition. If $\beta < \infty$, then*

$$\limsup_{t \nearrow \beta} \min\Big\{\operatorname{dist}(x(t), \mathbb{R}^n \setminus U), \frac{1}{\|x(t)\|}\Big\} = 0.$$

Proof. We argue by contradiction. Suppose that $\beta < \infty$ and

$$\inf_{k \in \mathbb{N}} \min\Big\{\operatorname{dist}(x(t_k), \mathbb{R}^n \setminus U), \frac{1}{\|x(t_k)\|}\Big\} > 0$$

for some sequence of times $t_k \nearrow \beta$. Then the sequence $\{x(t_k) : k \in \mathbb{N}\}$ is contained in a compact subset of U.

Using either Theorem 3.1 or Theorem 3.2, we can find a uniform constant $\delta > 0$ with the property that $(-\delta, \delta) \subset J^\star_{x(t_k)}$ for all k. It is important that the constant δ does not depend on k; this is possible since the sequence $\{x(t_k) : k \in \mathbb{N}\}$ is contained in a compact subset of U. Consequently, $[0, t_k + \delta) \subset J^\star_{x_0}$ for each k. This implies $t_k + \delta \leq \beta$ for all k. This contradicts the fact that $\lim_{k \to \infty} t_k = \beta$. This completes the proof of Theorem 3.5. $\square$

We now define

$$\Omega = \{(x_0, t) \in U \times \mathbb{R}^n : t \in J^\star_{x_0}\}.$$

Moreover, for each point $(x_0, t) \in \Omega$, we define

$$\Phi(x_0, t) = \varphi_t(x_0) = x(t),$$

where $x : J^\star_{x_0} \to U$ denotes the unique maximal solution of (4). The mapping $\Phi : \Omega \to U$ is referred as the flow associated with the differential equation $x'(t) = F(x(t))$. The map Φ enjoys the following semigroup property: If $(x_0, s) \in \Omega$ and $(\Phi(x_0, s), t) \in \Omega$, then $(x_0, s + t) \in \Omega$ and $\Phi(x_0, s + t) = \Phi(\Phi(x_0, s), t)$. This is an easy consequence of the uniqueness theorem above.

3.4. Continuous dependence on the initial data

As above let $U \subset \mathbb{R}^n \times \mathbb{R}$, and let $F : U \to \mathbb{R}^n$ be a continuous mapping which is continuously differentiable in the spatial variables. We denote by Φ the flow associated with this differential equation, and by $\Omega \subset U \times \mathbb{R}$ the domain of definition of Φ.

Theorem 3.6. *The set Ω is open, and $\Phi : \Omega \to U$ is continuous.*

In order to prove this theorem, we consider an arbitrary point $(x_0, t_0) \in \Omega$. Our goal is to show that a Ω contains a ball of positive radius centered at (x_0, t_0) and that Φ is continuous at (x_0, t_0).

It suffices to consider the case $t_0 \geq 0$. (The case $t_0 < 0$ can be handled in an analogous fashion.) Let

$$C = \{\Phi(x_0, t) : 0 \leq t \leq t_0\}.$$

The set C is closed and bounded subset of U. Since U is open, the set C has positive distance from the boundary of U. Hence, we can find a positive real number r such that

$$0 < 4r < \mathrm{dist}(C, \mathbb{R}^n \setminus U).$$

Let

$$M = \sup\{\|F(x)\| : \mathrm{dist}(x, C) \leq 4r\}$$

and

$$L = \sup\{\|DF(x)\|_{\mathrm{op}} : \mathrm{dist}(x, C) \leq 4r\}.$$

Lemma 3.7. *Fix a point y_0 such that $\|x_0 - y_0\| \leq e^{-Lt_0}\, r$. Then*

$$\|\Phi(x_0, t) - \Phi(y_0, t)\| \leq 2e^{Lt}\, \|x_0 - y_0\|$$

whenever $0 \leq t \leq t_0$ and $(y_0, t) \in \Omega$.

Proof. For abbreviation, let $x(t) = \Phi(x_0, t)$ and $y(t) = \Phi(y_0, t)$. Suppose that there exists a real number $t \in [0, t_0]$ such that $(y_0, t) \in \Omega$ and

$$\|x(t) - y(t)\| > 2e^{Lt}\, \|x_0 - y_0\|.$$

Observe first that $x_0 \neq y_0$. (If $x_0 = y_0$, then the uniqueness theorem implies that $x(t) = y(t)$ for all t.) Let

$$\tau = \inf\{t \in [0, t_0] : (y_0, t) \in \Omega \text{ and } \|x(t) - y(t)\| > 2e^{Lt}\, \|x_0 - y\|\}.$$

Then $0 < \tau \leq t_0$, $(y, \tau) \in \Omega$, and

$$\|x(\tau) - y(\tau)\| = 2e^{L\tau}\, \|x_0 - y_0\|.$$

Moreover,

$$\|x(t) - y(t)\| \leq 2e^{Lt}\, \|x_0 - y_0\|$$

for all $t \in [0, \tau]$. This implies

$$\|x(t) - y(t)\| \leq 2e^{Lt_0}\, \|x_0 - y_0\| \leq 2r$$

for all $t \in [0, \tau]$. Hence, for every $t \in [0, \tau]$, the line segment joining $x(t)$ and $y(t)$ is contained in the set $\{x \in \mathbb{R}^n : \mathrm{dist}(x, C) \leq 2r\}$. By definition of L, we have

$$\|F(x(t)) - F(y(t))\| \leq L\, \|x(t) - y(t)\|$$

for all $t \in [0, \tau]$. This implies

$$\begin{aligned}
\frac{d}{dt}\|x(t) - y(t)\|^2 &= 2\, \langle x(t) - y(t), x'(t) - y'(t) \rangle \\
&= 2\, \langle x(t) - y(t), F(x(t)) - F(y(t)) \rangle \\
&\leq 2L\, \|x(t) - y(t)\|^2
\end{aligned}$$

for all $t \in [0, \tau]$. Consequently, the function $t \mapsto e^{-2Lt} \|x(t) - y(t)\|^2$ is also monotone decreasing on $[0, \tau]$. Consequently,

$$2 \|x_0 - y_0\| \leq e^{-L\tau} \|x(\tau) - y(\tau)\| \leq \|x(0) - y(0)\| = \|x_0 - y_0\|.$$

This is a contradiction. □

Lemma 3.8. *Fix a point y_0 such that $\|x_0 - y_0\| \leq e^{-Lt_0} r$. If $0 \leq s < t_0 + \frac{r}{M}$ and $(y_0, t) \in \Omega$, then* $\mathrm{dist}(\Phi(y_0, t), C) \leq 4r$.

Proof. Suppose that there exists a number t such that $0 \leq t < t_0 + \frac{r}{M}$, $(y_0, s) \in \Omega$, and $\mathrm{dist}(\Phi(y_0, t), C) > 4r$. In particular, $(y_0, t_0) \in \Omega$. We define

$$\tau = \inf\{t \in [0, t_0 + \frac{r}{M}) : (y_0, t) \in \Omega \text{ and } \mathrm{dist}(\Phi(y_0, t), C) > 4r\}.$$

It follows from Lemma 3.7 that

$$\mathrm{dist}(\Phi(y_0, t), C) \leq \|\Phi(x_0, t) - \Phi(y_0, t)\| \leq 2e^{Lt} \|x_0 - y_0\| \leq 2r$$

for $0 \leq t \leq t_0$. Consequently, $t_0 < \tau < t_0 + \frac{r}{M}$, $(y_0, \tau) \in \Omega$, and $\mathrm{dist}(\Phi(y_0, \tau), C) = 4r$. Moreover, $\mathrm{dist}(\Phi(y_0, t), C) \leq 4r$ for all $t \in [t_0, \tau]$. This implies

$$\|\frac{d}{dt}\Phi(y_0, t)\| = \|F(\Phi(y_0, t))\| \leq M$$

for all $t \in [t_0, \tau]$ by definition of M. Integrating this inequality over the interval $[t_0, \tau]$ gives $\|\Phi(y_0, t_0) - \Phi(y_0, \tau)\| \leq r$. On the other hand, $\|\Phi(x_0, t_0) - \Phi(y_0, t_0)\| \leq 2r$ by Lemma 3.7. Putting these facts together, we conclude that $\|\Phi(x_0, t_0) - \Phi(y_0, \tau)\| \leq 3r$. This contradicts the fact that $\|\Phi(x_0, t_0) - \Phi(y_0, \tau)\| \geq \mathrm{dist}(\Phi(y_0, \tau), C) = 4r$. This proves the assertion. □

Lemma 3.9. *Suppose that $\|x_0 - y_0\| \leq e^{-Lt_0} r$ and $0 \leq t < t_0 + \frac{r}{M}$. Then $(y_0, t) \in \Omega$, and $\|\Phi(x_0, t_0) - \Phi(y_0, t)\| \leq 2e^{Lt_0} \|x_0 - y_0\| + M |t_0 - t|$.*

Proof. We first show that $(y_0, t) \in \Omega$ whenever $0 \leq t < t_0 + \frac{r}{M}$. Suppose by contradiction that the solution $t \mapsto \Phi(y_0, t)$ ceases to exist at some time $\tau < t_0 + \frac{r}{M}$. It follows from Lemma 3.8 that $\mathrm{dist}(\Phi(y_0, t), C) \leq 4r$ for all $t \in [0, \tau)$. This implies that the solution $t \mapsto \Phi(y_0, t)$ is contained in a compact subset of U, contradicting the global existence and uniqueness theorem. Since the set C is compact and $\mathrm{dist}(C, \mathbb{R}^n \setminus U) > 4r$, these statements are in contradiction. Thus, we conclude that $(y_0, t) \in \Omega$ whenever $0 \leq t < t_0 + \frac{r}{M}$.

Using Lemma 3.8, we obtain $\mathrm{dist}(\Phi(y_0, t), C) \leq 4r$ for $0 \leq t < t_0 + \frac{r}{M}$. Hence, by definition of M, we have

$$\|\frac{d}{dt}\Phi(y_0, t)\| = \|F(\Phi(y_0, t))\| \leq M$$

for $0 \leq t < t_0 + \frac{r}{M}$. This implies

$$\|\Phi(y_0, t_0) - \Phi(y_0, t)\| \leq M |t_0 - t|$$

for all $0 \leq t < t_0 + \frac{r}{M}$. Therefore,

$$\begin{aligned}\|\Phi(x_0, t_0) - \Phi(y_0, t)\| &\leq \|\Phi(x_0, t_0) - \Phi(y_0, t_0)\| + \|\Phi(y_0, t_0) - \Phi(y_0, t)\| \\ &\leq 2e^{Lt_0} \|x_0 - y_0\| + M\,|t_0 - t|.\end{aligned}$$

This proves the claim. □

To summarize, we have shown that Ω contains a neighborhood of the point (x_0, t_0), and Φ is continuous at the point (x_0, t_0).

3.5. Differentiability of flows and the linearized equation

As usual, we consider an autonomous system of ordinary differential equations of the form $x'(t) = F(x(t))$, where $F : U \to \mathbb{R}^n$ is a continuously differentiable function defined on some open subset U of $\mathbb{R}^n$. We denote by Φ the flow associated with this differential equation, and by $\Omega \subset U \times \mathbb{R}$ the domain of definition of Φ. We have shown earlier that the set Ω is open and the map $\Phi : \Omega \to U$ is continuous.

We next establish differentiability of Φ. To that end, we need the following auxiliary result:

Proposition 3.10. *Let $A : [0, T] \to \mathbb{R}^{n\times n}$ be a continuous function which takes values values in the space of $n \times n$ matrices. Then there exists a continuously differentiable function $M : [0, T] \to \mathbb{R}^{n\times n}$ such that $M'(t) = A(t)M(t)$ and $M(0) = I$.*

Proof. For abbreviation, let $N = \sup_{t\in[0,T]} \|A(t)\|_{\mathrm{op}}$. For each $k \in \mathbb{N}$, we define a continuous function $M_k : [0, T] \to \mathbb{R}^{n\times n}$ by

$$M_k(t) = I + t\,A(0)$$

for $t \in [0, \frac{1}{k}]$ and

$$M_k(t) = \Big[I + \Big(t - \frac{jT}{k}\Big)\, A\Big(\frac{jT}{k}\Big)\Big]\, M_k\Big(\frac{jT}{k}\Big)$$

for $t \in [\frac{jT}{k}, \frac{(j+1)T}{k}]$. Using induction on j, we can show that

$$\|M_k(t)\|_{\mathrm{op}} \leq e^{Nt}$$

for all $t \in [0, \frac{jT}{k}]$. By the Arzelà-Ascoli theorem, a subsequence of the sequence $M_k(t)$ converges uniformly to some function $M(t)$. The function $M : [0, T] \to \mathbb{R}^{n\times n}$ is clearly continuous. Moreover, by following the arguments in the proof of Theorem 3.1, one readily verifies that

$$M(t) = I + \int_0^t A(s)M(s)\,ds$$

for all $t \in [0, T]$. Thus, $M(t)$ is continuously differentiable and $M'(t) = A(t)M(t)$. $\square$

Theorem 3.11. *Fix a point $(x_0, t_0) \in \Omega$. Let $x(t) = \Phi(x_0, t)$ denote the unique maximal solution with initial vector x_0, and let $A(t) = DF(x(t))$. Moreover, suppose that $M : [0, t_0] \to \mathbb{R}^{n\times n}$ is a solution of the ODE $M'(t) = A(t)M(t)$ with initial condition $M(0) = I$. Then the map φ_{t_0} is differentiable at x_0, and its differential is given by $D\varphi_{t_0} = M(t_0)$.*

In the following, we describe the proof of Theorem 3.11. We need to show that

$$\lim_{y\to 0} \frac{1}{\|x_0 - y\|} \|\varphi_{t_0}(x_0) - \varphi_{t_0}(y) - M(t_0)(x_0 - y)\| = 0.$$

To verify this, let $\varepsilon > 0$ be given. Since F is continuously differentiable, we can find a real number $r > 0$ such that for each $t \in [0, t_0]$ we have $B_{2r}(x(t)) \subset U$ and $\sup_{y\in B_r(x(t))} \|DF(y) - A(t)\|_{\mathrm{op}} \leq \varepsilon$. Integrating this inequality over a line segment, we obtain

$$\|F(x(t)) - F(y) - A(t)\,(x(t) - y)\| \leq \varepsilon \,\|y - x(t)\|$$

for all $t \in [0, t_0]$ and all points $y \in B_r(x(t))$.

Since Φ is continuous, we can find a real number $0 < \delta < r$ such that

$$\sup_{t\in[0,t_0]} \|\Phi(x_0, t) - \Phi(y_0, t)\| \leq r$$

for all points y_0 satisfying $\|x_0 - y_0\| \leq \delta$.

We now consider a point y_0 with $0 < \|x_0 - y_0\| \leq \delta$. For abbreviation, let $K := \sup_{t\in[0,t_0]} \|A(t)\|$ and $L = \sup_{t\in[0,t_0]} \|M(t)\|$. The function $y(t) = \Phi(y_0, t)$ satisfies $y'(t) = F(y(t))$. Hence, the function $u(t) := x(t) - y(t) - M(t)(x_0 - y_0)$ satisfies

$$\begin{aligned}
u'(t) &= x'(t) - y'(t) - M'(t)(x_0 - y_0)\\
&= F(x(t)) - F(y(t)) - A(t)M(t)(x_0 - y_0)\\
&= F(x(t)) - F(y(t)) - A(t)\,(x(t) - y(t)) + A(t)u(t).
\end{aligned}$$

This implies

$$\begin{aligned}
\|u'(t)\| &\leq \|F(x(t)) - F(y(t)) - A(t)\,(x(t) - y(t))\| + \|A(t)\|\,\|u(t)\|\\
&\leq \varepsilon\,\|x(t) - y(t)\| + K\,\|u(t)\|\\
&\leq \varepsilon\,\|M(t)(x_0 - y_0)\| + (K + \varepsilon)\,\|u(t)\|\\
&\leq \varepsilon L\,\|x_0 - y_0\| + (K + \varepsilon)\,\|u(t)\|
\end{aligned}$$

for all $t \in [0, t_0]$.

Lemma 3.12. *We have*
$$e^{-(K+\varepsilon)t_0}\,\|u(t_0)\| \leq t_0\varepsilon L\,\|x_0-y_0\|.$$

Proof. Note that $u(0)=0$ by definition of $u(t)$. If $u(t_0)=0$, the assertion is trivial. Hence, we may assume that $u(t_0)\neq 0$. Let $\tau=\sup\{t\in[0,t_0] : u(t)=0\}$. Clearly, $\tau\in[0,t_0)$, $u(\tau)=0$, and $u(\tau)\neq 0$ for all $t\in(\tau,t_0]$. Note that
$$\frac{d}{dt}\|u(t)\| \leq \|u'(t)\| \leq \varepsilon L\,\|x_0-y_0\| + (K+\varepsilon)\,\|u(t)\|$$
for all $t\in(\tau,t_0]$. This implies
$$\frac{d}{dt}(e^{-(K+\varepsilon)t}\,\|u(t)\|) \leq \varepsilon L\,e^{-(K+\varepsilon)t}\,\|x_0-y_0\| \leq \varepsilon L\,\|x_0-y_0\|$$
for all $t\in(\tau,t_0]$. Consequently, the function $t\mapsto e^{-(K+\varepsilon)t}\,\|u(t)\| - t\varepsilon L\,\|x_0-y_0\|$ is monotone decreasing on the interval $(\tau,t_0]$. Since $u(\tau)=0$, it follows that
$$e^{-(K+\varepsilon)t_0}\,\|u(t_0)\| - t_0\varepsilon L\,\|x_0-y_0\| \leq e^{-(K+\varepsilon)t_0}\,\|u(\tau)\| - \tau\varepsilon L\,\|x_0-y_0\| \leq 0,$$
which establishes the claim. $\square$

We now complete the proof of Theorem 3.11. Using Lemma 3.12, we obtain
$$\begin{aligned}\frac{1}{\|x_0-y_0\|}\,\|\varphi_{t_0}(x_0)-\varphi_{t_0}(y_0)-M(t_0)(x_0-y_0)\| &= \frac{1}{\|x_0-y_0\|}\,\|u(t_0)\| \\ &\leq t_0\varepsilon L\,e^{(K+\varepsilon)t_0}\end{aligned}$$
for all points y_0 satisfying $0<\|x_0-y_0\|\leq\delta$. Since $\varepsilon>0$ is arbitrary, we conclude that
$$\lim_{y\to 0}\frac{1}{\|x_0-y\|}\,\|\varphi_{t_0}(x_0)-\varphi_{t_0}(y)-M(t_0)(x_0-y)\| = 0.$$
This completes the proof.

3.6. Liouville's theorem

Proposition 3.13. *Let $M(t)$ be a continuously differentiable function taking values in $\mathbb{R}^{n\times n}$. If $M'(t)=A(t)M(t)$, then*
$$\det M(t) = \det M(0)\,e^{\int_0^t \operatorname{tr}(A(s))\,ds}.$$

Proof. It suffices to show that
$$\frac{d}{dt}\det M(t) = \operatorname{tr}(A(t))\,\det M(t)$$
for all t. In order to verify this, we fix a time t_0. Let us write $A(t_0)=SBS^{-1}$, where $S\in\mathbb{C}^{n\times n}$ is invertible and $B\in\mathbb{C}^{n\times n}$ is an upper triangular matrix.

Moreover, let $\tilde{M}(t) = (I + (t - t_0)\, A(t_0))M(t_0)$. Then $\tilde{M}(t_0) = M(t_0)$ and $\tilde{M}'(t_0) = M'(t_0)$. This implies

$$\frac{d}{dt} \det M(t)\Big|_{t=t_0} = \frac{d}{dt} \det \tilde{M}(t)\Big|_{t=t_0}.$$

Moreover,

$$\begin{aligned}\det \tilde{M}(t) &= \det(I + (t - t_0)\, A(t_0)) \det M(t_0) \\ &= \det(I + (t - t_0)\, B) \det M(t_0) \\ &= \prod_{k=1}^{n} (1 + (t - t_0)\, b_{kk}) \det M(t_0).\end{aligned}$$

This implies

$$\begin{aligned}\frac{d}{dt} \det \tilde{M}(t)\Big|_{t=t_0} &= \sum_{k=1}^{n} b_{kk} \det M(t_0) \\ &= \operatorname{tr}(B) \det M(t_0) \\ &= \operatorname{tr}(A(t_0)) \det M(t_0).\end{aligned}$$

Putting these facts together, we conclude that

$$\frac{d}{dt} \det M(t)\Big|_{t=t_0} = \operatorname{tr}(A(t_0)) \det M(t_0),$$

as claimed. □

Theorem 3.14. *Suppose that $DF(x)$ is trace-free for all $x \in U$. Then φ_t is volume-preserving for each t; that is,* $\det D\varphi_t(x) = 1$ *for all $(x, t) \in \Omega$.*

Proof. Fix a point $(x_0, t_0) \in \Omega$. For abbreviation, let $x(t) = \Phi(x_0, t)$ and $A(t) = DF(x(t))$. Then $D\varphi_{t_0}(x_0) = M(t_0)$, where $M(t)$ is a solution of the differential equation $M'(t) = A(t)M(t)$ with initial condition $M(0) = I$. Since $A(t)$ is tracefree, we conclude that $\det M(t)$ is constant in t. Thus, $\det D\varphi_{t_0}(x_0) = \det M(t_0) = \det M(0) = 1$. □

3.7. Problems

Problem 3.1. Let F be a continuously differentiable vector field on $\mathbb{R}^n$. Moreover, suppose that $\|F(x)\| \leq \psi(\|x\|)$, where $\psi : [0, \infty) \to (0, \infty)$ is a continuous function satisfying $\int_0^\infty \frac{1}{\psi(r)}\, dr = \infty$. Fix a point $x_0 \in \mathbb{R}^n$, and let $x(t)$ denote the unique solution of the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$. Show that $x(t)$ is defined for all $t \in \mathbb{R}$.

Problem 3.2. Let us define a continuous function $F : \mathbb{R} \to \mathbb{R}$ by $F(x) = \sqrt{|x|}$. Show that the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = 0$ has infinitely many solutions.

Problem 3.3. Let $U \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set, and let $F : U \to \mathbb{R}^n$ be a continuously differentiable mapping. Given any point $(x_0, \lambda_0) \in U$, we define $\Phi(x_0, \lambda_0, t) = x(t)$, where $x(t)$ is the unique solution of the differential equation $x'(t) = F(x(t), \lambda_0)$ with initial condition $x(0) = x_0$. Moreover, let Ω be the set of all triplets $(x_0, \lambda_0, t) \in U \times \mathbb{R}$ for which $\Phi(x_0, \lambda_0, t)$ is defined. Show that the set Ω is open and the map $\Phi : \Omega \to \mathbb{R}^n$ is continuous. (Hint: Apply Theorem 3.6 to the system $x'(t) = F(x(t), \lambda(t))$, $\lambda'(t) = 0$.)

Chapter 4

Analysis of equilibrium points

4.1. Stability of equilibrium points

Definition 4.1. Suppose that $\bar{x}$ is an equilibrium point of the autonomous system $x'(t) = F(x(t))$, so that $F(\bar{x}) = 0$. We say that $\bar{x}$ is stable if, given any real number $\varepsilon > 0$ we can find a real number $\delta > 0$ such that $\varphi_t(x) \in B_\varepsilon(\bar{x})$ for all $x \in B_\delta(\bar{x})$ and all $t \geq 0$. We say that $\bar{x}$ is asymptotically stable if $\bar{x}$ is stable and $\lim_{t\to\infty} \varphi_t(x) = \bar{x}$ for all points x in some open neighorhood of $\bar{x}$.

We note that the condition that $\lim_{t\to\infty} \varphi_t(x) = \bar{x}$ for all points x in some open neighorhood of $\bar{x}$ does not imply stability (cf. Problem 4.3 below).

The following result gives a sufficient condition for an equilibrium point to be asymptotically stable:

Theorem 4.2. *Suppose that* 0 *is an equilibrium point of the system* $x'(t) = F(x(t))$*. Moreover, suppose that the eigenvalues of the matrix* $A = DF(0)$ *all have negative real part. Then* 0 *is asymptotically stable.*

Proof. Suppose that $\varepsilon > 0$ is given. Since the eigenvalues of A have negative real part, we have $\lim_{t\to\infty} \exp(tA) = 0$. Let us fix a real number $T > 0$ such that $\|\exp(TA)\|_{\mathrm{op}} < \frac{1}{2}$. By Theorem 3.6, we can find a real number $\eta > 0$ such that $\sup_{t\in[0,T]} \|\varphi_t(y)\| < \varepsilon$ for all points $y \in B_\eta(0)$. Moreover, by Theorem 3.11, we have $D\varphi_T(0) = \exp(TA)$. Since $\|\exp(TA)\|_{\mathrm{op}} < \frac{1}{2}$, we can find a real number $\delta \in (0, \eta)$ such that $\|\varphi_T(y)\| \leq \frac{1}{2} \|y\|$ for all points $y \in B_\delta(0)$.

We now consider an initial point $y_0 \in B_\delta(0)$. Using induction on k, we can show that the solution $\varphi_t(y_0)$ is defined on $[0, kT]$, and $\|\varphi_{kT}(y_0)\| \leq 2^{-k}\,\delta$. Since $\|\varphi_{kT}(y_0)\| < \eta$, we conclude that

$$\sup_{t \in [kT,(k+1)T]} \|\varphi_t(y_0)\| = \sup_{t \in [0,T]} \|\varphi_t(\varphi_{kT}(y_0))\| < \varepsilon$$

for all k. Therefore, $\varphi_t(y_0) \in B_\varepsilon(0)$ for all $t \geq 0$. This shows that 0 is a stable equilibrium point. Finally, since $\varphi_{kT}(y_0) \to 0$ as $k \to \infty$, Theorem 3.6 implies that

$$\sup_{t \in [kT,(k+1)T]} \|\varphi_t(y_0)\| = \sup_{t \in [0,T]} \|\varphi_t(\varphi_{kT}(y_0))\| \to 0$$

as $k \to \infty$. This completes the proof. □

4.2. The stable manifold theorem

In this section, we give a precise description of the qualitative behavior of a dynamical system near an equilibrium point. Our discussion loosely follows the one in [**6**]. We consider an open set $U \subset \mathbb{R}^n$ and a map $F : U \to \mathbb{R}^n$ of class C^1. Let us assume that 0 is a hyperbolic equilibrium point for the differential equation $x'(t) = F(x(t))$. That is, we have $0 \in U$, $F(0) = 0$, and all eigenvalues of the matrix $A = DF(0)$ have non-zero real part. For abbreviation, let $G(x) = F(x) - Ax$. Clearly, $DG(0) = 0$.

Let $U_- = \ker p_-(A)$ and $U_+ = \ker p_+(A)$ denote the stable and unstable subspaces of A, and let P_- and P_+ denote the associated projections. Since U_- and U_+ are invariant under A, the projections P_- and P_+ commute with A. We can find positive constants α and Λ such that

$$\|\exp(tA)\,P_-\|_{\text{op}} \leq \Lambda\,e^{-\alpha t}$$

for all $t \geq 0$ and

$$\|\exp(tA)\,P_+\|_{\text{op}} \leq \Lambda\,e^{\alpha t}$$

for all $t \leq 0$. Suppose that $r > 0$ is sufficiently small such that $B_{2Kr}(0) \subset U$ and

$$\sup_{\|x\| \leq 2\Lambda r} \|DF(x) - DF(0)\|_{\text{op}} \leq \frac{3\alpha}{16\Lambda}.$$

Since $DG(x) = DF(x) - A = DF(x) - DF(0)$, it follows that

$$\sup_{\|x\| \leq 2\Lambda r} \|DG(x)\|_{\text{op}} \leq \frac{3\alpha}{16\Lambda}.$$

Lemma 4.3. *Given any vector $x_- \in U_-$ with $\|x_-\| \leq r$, there exists a function $x(t)$ such that $\|x(t)\| \leq 2\Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$ and*

$$\begin{aligned} x(t) = \exp(tA)x_- &+ \int_0^t \exp((t-s)A)\, P_- G(x(s))\, ds \\ &- \int_t^\infty \exp(-(s-t)A)\, P_+ G(x(s))\, ds \end{aligned} \tag{5}$$

for all $t \geq 0$.

Proof. The strategy is to use an iterative method. We define a sequence of functions $x_k(t)$ by

$$x_0(t) = 0$$

and

$$\begin{aligned} x_k(t) = \exp(tA)x_- &+ \int_0^t \exp((t-s)A)\, P_- G(x_{k-1}(s))\, ds \\ &- \int_t^\infty \exp(-(s-t)A)\, P_+ G(x_{k-1}(s))\, ds. \end{aligned}$$

We claim that the function $x_k(t)$ is well-defined, and

$$\|x_k(t) - x_{k-1}(t)\| \leq 2^{-k+1}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$$

for all $t \geq 0$.

Since $x_1(t) = \exp(tA)\, x_-$ and $x_0(t) = 0$, we have $\|x_1(t) - x_0(t)\| = \|\exp(tA)\, x_-\| = \|\exp(tA)\, P_- x_-\| \leq \Lambda\, e^{-\alpha t}\, \|x_-\|$ for all $t \geq 0$. Therefore, the assertion holds for $k = 1$. We next assume that $k \geq 1$ and the assertion holds for all integers less than or equal to k. In other words, the function $x_j(t)$ is well-defined and

$$\|x_j(t) - x_{j-1}(t)\| \leq 2^{-j+1}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$$

for all $j \leq k$. Summation over j gives

$$\begin{aligned} \max\{\|x_k(t)\|, \|x_{k-1}(t)\|\} &\leq \sum_{j=1}^k \|x_j(t) - x_{j-1}(t)\| \\ &\leq \sum_{j=1}^k 2^{-j+1}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\| \\ &\leq 2\Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\| \leq 2\Lambda r \end{aligned}$$

for all $t \geq 0$. In particular, the function $x_{k+1}(t)$ can be defined. Moreover, $\|G(\xi) - G(\eta)\| \leq \frac{3\alpha}{16\Lambda}\, \|\xi - \eta\|$ whenever $\xi, \eta \in B_{2\Lambda r}(0)$. Since $x_k(t), x_{k-1}(t) \in B_{2\Lambda r}(0)$, we obtain

$$\|G(x_k(t)) - G(x_{k-1}(t))\| \leq \frac{3\alpha}{16\Lambda}\, \|x_k(t) - x_{k-1}(t)\| \leq \frac{3\alpha}{8}\, 2^{-k}\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$$

for all $t \geq 0$. By definition of $x_{k+1}(t)$ and $x_k(t)$, we have

$$\begin{aligned} x_{k+1}(t) - x_k(t) &= \int_0^t \exp((t-s)A)\, P_-\, (G(x_k(s)) - G(x_{k-1}(s)))\, ds \\ &\quad - \int_t^\infty \exp((t-s)A)\, P_+\, (G(x_k(s)) - G(x_{k-1}(s)))\, ds, \end{aligned}$$

hence

$$\begin{aligned} \|x_{k+1}(t) - x_k(t)\| &\leq \int_0^t \Lambda\, e^{-\alpha(t-s)}\, \|G(x_k(s)) - G(x_{k-1}(s))\|\, ds \\ &\quad + \int_t^\infty \Lambda\, e^{-\alpha(s-t)}\, \|G(x_k(s)) - G(x_{k-1}(s))\|\, ds \end{aligned}$$

for all $t \geq 0$. Using our estimate for $\|G(x_k(s)) - G(x_{k-1}(s))\|$, we obtain

$$\begin{aligned} \|x_{k+1}(t) - x_k(t)\| &\leq \int_0^t \frac{3\alpha}{8}\, 2^{-k}\, \Lambda\, e^{-\alpha(t-s)}\, e^{-\frac{\alpha s}{2}}\, \|x_-\|\, ds \\ &\quad + \int_t^\infty \frac{3\alpha}{8}\, 2^{-k}\, \Lambda\, e^{-\alpha(s-t)}\, e^{-\frac{\alpha s}{2}}\, \|x_-\|\, ds \\ &\leq \int_0^t \frac{3\alpha}{8}\, 2^{-k}\, \Lambda\, e^{-\frac{\alpha(t-s)}{2}}\, e^{-\frac{\alpha t}{2}}\, \|x_-\|\, ds \\ &\quad + \int_t^\infty \frac{3\alpha}{8}\, 2^{-k}\, \Lambda\, e^{-\frac{3\alpha(s-t)}{2}}\, e^{-\frac{\alpha t}{2}}\, \|x_-\|\, ds \\ &\leq 2^{-k}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\| \end{aligned}$$

for all $t \geq 0$. This completes the proof of the claim.

We now continue with the proof of Lemma 4.3. The functions $x_k(t)$ satisfy

$$\|x_{k+1}(t) - x_k(t)\| \leq 2^{-k}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$$

for all $t \geq 0$ and all $k \geq 0$. In particular, for every $t \geq 0$ the sequence $\{x_k(t) : k \in \mathbb{N}\}$ is a Cauchy sequence. Hence, we may define a function $x(t)$ by

$$x(t) = \lim_{k \to \infty} x_k(t).$$

It is easy to see that

$$\|x(t) - x_k(t)\| \leq \sum_{j=k}^\infty \|x_{j+1}(t) - x_j(t)\| \leq 2^{-k+1}\, \Lambda\, e^{-\frac{\alpha t}{2}}\, \|x_-\|$$

for all $t \geq 0$ and $k \geq 0$. Hence, the functions $x_k(t)$ converge uniformly to $x(t)$. As a consequence, the function $x(t)$ is continuous and satisfies

$$\begin{aligned} x(t) = \exp(tA)x_- &+ \int_0^t \exp((t-s)A)\, P_- G(x(s))\, ds \\ &- \int_t^\infty \exp(-(s-t)A)\, P_+ G(x(s))\, ds \end{aligned}$$

for all $t \geq 0$. This completes the proof of Lemma 4.3. □

In the next step, we show that the function $x(t)$ is unique. In fact, we prove a stronger statement:

Lemma 4.4. *Consider two vectors $x_-, y_- \in U_-$ with $\|x_-\|, \|y_-\| \leq r$. Moreover, suppose that $x(t)$ and $y(t)$ are two functions satisfying $\|x(t)\|, \|y(t)\| \leq 2\Lambda r$ and*

$$\begin{aligned} x(t) = \exp(tA)x_- &+ \int_0^t \exp((t-s)A)\, P_- G(x(s))\, ds \\ &- \int_t^\infty \exp(-(s-t)A)\, P_+ G(x(s))\, ds \end{aligned}$$

and

$$\begin{aligned} y(t) = \exp(tA)x_- &+ \int_0^t \exp((t-s)A)\, P_- G(x(s))\, ds \\ &- \int_t^\infty \exp(-(s-t)A)\, P_+ G(x(s))\, ds \end{aligned}$$

for all $t \geq 0$. Then

$$\|x(t) - y(t)\| \leq 2\Lambda\, \|x_- - y_-\|.$$

Proof. We compute

$$\begin{aligned} x(t) - y(t) = \exp(tA)\,(x_- - y_-) &+ \int_0^t \exp((t-s)A)\, P_-\,(G(x(s)) - G(y(s)))\, ds \\ &- \int_t^\infty \exp((t-s)A)\, P_+\,(G(x(s)) - G(y(s)))\, ds \end{aligned}$$

for all $t \geq 0$. This implies

$$\begin{aligned}\|x(t) - y(t)\| &\leq \Lambda\, e^{-\alpha t}\, \|x_- - y_-\| + \int_0^t \Lambda\, e^{-\alpha(t-s)}\, \|G(x(s)) - G(y(s))\|\, ds \\ &\quad + \int_t^\infty \frac{3\alpha}{16}\, e^{-\alpha(s-t)}\, \|G(x(s)) - G(y(s))\|\, ds \\ &\leq \Lambda\, e^{-\alpha t}\, \|x_- - y_-\| + \int_0^t \frac{3\alpha}{16}\, e^{-\alpha(t-s)}\, \|x(s) - y(s)\|\, ds \\ &\quad + \int_t^\infty \frac{3\alpha}{16}\, e^{-\alpha(s-t)}\, \|x(s) - y(s)\|\, ds \\ &\leq \Lambda\, \|x_- - y_-\| + \frac{3}{8} \sup_{s \geq 0} \|x(s) - y(s)\|\end{aligned}$$

for all $t \geq 0$. Thus, we conclude that

$$\sup_{t \geq 0} \|x(t) - y(t)\| \leq \Lambda\, \|x_- - y_-\| + \frac{3}{8} \sup_{t \geq 0} \|x(t) - y(t)\|,$$

hence

$$\sup_{t \geq 0} \|x(t) - y(t)\| \leq 2\Lambda\, \|x_- - y_-\|.$$

This proves the assertion. □

We define a function ψ from U_- to U_+ as follows: for every vector $x_- \in U_-$ with $\|x_-\| \leq r$, we define $\psi(x_-) = P_+ x(0) \in U_+$, where $x(t)$ is the unique solution of the integral equation (5). It follows from Lemma 4.4 that $\psi(x_-)$ is well-defined and continuous. We next show that ψ is differentiable:

Lemma 4.5. *Fix a point $x_- \in U_-$ with $\|x_-\| \leq r$, and let $x(t)$ denote the solution of the integral equation (5). Then there exists a bounded continuous function $M : [0, \infty) \to \mathbb{R}^{n \times n}$ such that*

$$\begin{aligned}M(t) &= \exp(tA) P_- + \int_0^t \exp((t-s)A)\, P_-\, DG(x(s))\, M(s)\, ds \\ &\quad - \int_t^\infty \exp(-(s-t)A)\, P_+\, DG(x(s))\, M(s)\, ds.\end{aligned} \tag{6}$$

Moreover, the function ψ is differentiable at x_-, and the differential $D\psi(x_-) : U_- \to U_+$ is given by $D\psi_-(x_-) = P_+ M(0)|_{U_-}$.

Proof. The existence of a solution to the integral equation (6) follows from an iteration procedure similar to the one used in the proof of Lemma 4.3. We omit the details.

We claim that ψ is differentiable at x_- with differential $D\psi_-(x_-) = P_+ M(0)|_{U_-}$. To prove this, let $\varepsilon > 0$ be given. We can find a real number $\delta > 0$ such that $\|DG(\xi) - DG(\eta)\|_{\text{op}} \leq \varepsilon$ for all points $\xi, \eta \in B_{2\Lambda r}(0)$ satisfying $\|\xi - \eta\| \leq 2\Lambda\delta$. Integrating this inequality over a line segment, we conclude that $\|G(\xi) - G(\eta) - DG(\xi)\,(\xi - \eta)\| \leq \varepsilon\,\|\xi - \eta\|$ whenever $\|\xi - \eta\| \leq 2\Lambda\delta$.

We now consider a point $y_- \in U_-$ with $\|y_-\| \leq r$ and $\|x_- - y_-\| \leq \delta$. We define $u(t) := x(t) - y(t) - M(t)\,(x_- - y_-)$, where $x(t)$ and $y(t)$ denote the solutions of our integral equation associated with x_- and y_-. Then

$$\begin{aligned} x(t) - y(t) = \exp(tA)\,(x_- - y_-) &+ \int_0^t \exp((t-s)A)\,P_-\,(G(x(s)) - G(y(s)))\,ds \\ &- \int_t^\infty \exp(-(s-t)A)\,P_+\,(G(x(s)) - G(y(s)))\,ds, \end{aligned}$$

hence

$$\begin{aligned} u(t) = &\int_0^t \exp((t-s)A)\,P_-\,[G(x(s)) - G(y(s)) - DG(x(s)\,M(s)\,(x_- - y_-)]\,ds \\ &- \int_t^\infty \exp(-(s-t)A)\,P_+\,[G(x(s)) - G(y(s)) - DG(x(s))\,M(s)\,(x_- - y_-)]\,ds \end{aligned}$$

for all $t \geq 0$. We have shown above that

$$\|x(t) - y(t)\| \leq 2\Lambda\,\|x_- - y_-\| \leq 2\Lambda\delta$$

for all $t \geq 0$. This implies

$$\begin{aligned} &\|G(x(t)) - G(y(t)) - DG(x(t))\,M(t)\,(x_- - y_-)\| \\ &\leq \|G(x(t)) - G(y(t)) - DG(x(t))\,(x(t) - y(t))\| + \|DG(x(t))\,u(t)\| \\ &\leq \varepsilon\,\|x(t) - y(t)\| + \frac{3\alpha}{16\Lambda}\,\|u(t)\| \\ &\leq 2\Lambda\varepsilon\,\|x_- - y_-\| + \frac{3\alpha}{16\Lambda}\,\|u(t)\| \end{aligned}$$

for all $t \geq 0$. Putting these facts together, we obtain

$$\begin{aligned} \|u(t)\| \leq &\int_0^t \|\exp((t-s)A)\,P_-\|_{\mathrm{op}} \Big(2\Lambda\varepsilon\,\|x_- - y_-\| + \frac{3\alpha}{16\Lambda}\,\|u(s)\|\Big)\,ds \\ &+ \int_t^\infty \|\exp(-(s-t)A)\,P_-\|_{\mathrm{op}} \Big(2\Lambda\varepsilon\,\|x_- - y_-\| + \frac{3\alpha}{16\Lambda}\,\|u(s)\|\Big)\,ds \\ &\leq \int_0^t \Lambda\,e^{-\alpha(t-s)} \Big(2\Lambda\varepsilon\,\|x_- - y_-\| + \frac{3\alpha}{16\Lambda}\,\|u(s)\|\Big)\,ds \\ &+ \int_t^\infty \Lambda\,e^{-\alpha(s-t)} \Big(2\Lambda\varepsilon\,\|x_- - y_-\| + \frac{3\alpha}{16\Lambda}\,\|u(s)\|\Big)\,ds \\ &\leq \frac{4\Lambda^2\varepsilon}{\alpha}\,\|x_- - y_-\| + \frac{3}{8}\,\sup_{s \geq 0} \|u(s)\| \end{aligned}$$

for all $t \geq 0$. Thus, we conclude that

$$\sup_{t \geq 0} \|u(t)\| \leq \frac{4\Lambda^2\varepsilon}{\alpha}\,\|x_- - y_-\| + \frac{3}{8}\,\sup_{t \geq 0} \|u(t)\|,$$

hence

$$\sup_{t \geq 0} \|u(t)\| \leq \frac{8\Lambda^2 \varepsilon}{\alpha} \|x_- - y_-\|.$$

In particular,

$$\|u(0)\| \leq \frac{8\Lambda^2 \varepsilon}{\alpha} \|x_- - y_-\|.$$

This implies

$$\|\psi(x_-) - \psi(y_-) - P_+ M(0)\,(x_- - y_-)\| = \|P_+ u(0)\| \leq \frac{8\Lambda^3 \varepsilon}{\alpha} \|x_- - y_-\|.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that ψ is differentiable at x_- and $D\psi(x_-) = P_+ M(0)|_{U_-}$. This completes the proof. □

Corollary 4.6. *We have $D\psi(0) = 0$.*

Proof. If $x_- = 0$, then $x(t) = 0$ for all $t \geq 0$. This implies $M(t) = \exp(tA)\, P_-$, hence $D\psi(0) = P_+ M(0)|_{U_-} = 0$. □

Lemma 4.7. *The differential $D\psi(x_-)$ depends continuously on the point x_-.*

Proof. Fix a point $x_- \in U_-$ with $\|x_-\| \leq r$, and let $\varepsilon > 0$ be given. As above, we can find a real number $\delta > 0$ such that $\|DG(\xi) - DG(\eta)\|_{\text{op}} \leq \varepsilon$ for all points $\xi, \eta \in B_{2\Lambda r}(0)$ satisfying $\|\xi - \eta\| \leq 2\Lambda\delta$.

We now consider a point $y_- \in U_-$ with $\|y_-\| \leq r$ and $\|x_- - y_-\| \leq \delta$. Let $M : [0, \infty) \to \mathbb{R}^{n \times n}$ and $N : [0, \infty) \to \mathbb{R}^{n \times n}$ be bounded continuous functions satisfying the integral equations

$$\begin{aligned} M(t) &= \exp(tA) P_- + \int_0^t \exp((t-s)A)\, P_-\, DG(x(s))\, M(s)\, ds \\ &\quad - \int_t^\infty \exp(-(s-t)A)\, P_+\, DG(x(s))\, M(s)\, ds \end{aligned}$$

and

$$\begin{aligned} N(t) &= \exp(tA) P_- + \int_0^t \exp((t-s)A)\, P_-\, DG(y(s))\, N(s)\, ds \\ &\quad - \int_t^\infty \exp(-(s-t)A)\, P_+\, DG(y(s))\, N(s)\, ds. \end{aligned}$$

Then

$$\begin{aligned} M(t) - N(t) &= \int_0^t \exp((t-s)A)\, P_-\, [DG(x(s))\, M(s) - DG(y(s))\, N(s)]\, ds \\ &\quad - \int_t^\infty \exp(-(s-t)A)\, P_+\, [DG(x(s))\, M(s) - DG(y(s))\, N(s)]\, ds. \end{aligned}$$

Note that

$$\|x(t) - y(t)\| \leq 2\Lambda\, \|x_- - y_-\| \leq 2\Lambda\delta$$

for all $t \geq 0$. This implies

$$\begin{aligned}
&\|DG(x(t))\, M(t) - DG(y(t))\, N(t)\|_{\mathrm{op}} \\
&\leq \|(DG(x(t)) - DG(y(t)))\, M(t)\|_{\mathrm{op}} + \|DG(y(t))\, (M(t) - N(t))\|_{\mathrm{op}} \\
&\leq \varepsilon\, \|M(t)\|_{\mathrm{op}} + \frac{3\alpha}{16\Lambda}\, \|M(t) - N(t)\|_{\mathrm{op}}.
\end{aligned}$$

This gives

$$\begin{aligned}
&\|M(t) - N(t)\|_{\mathrm{op}} \\
&\leq \int_0^t \|\exp((t-s)A)\, P_-\|_{\mathrm{op}} \Big(\varepsilon\, \|M(s)\|_{\mathrm{op}} + \frac{3\alpha}{16\Lambda}\, \|M(s) - N(s)\|_{\mathrm{op}}\Big)\, ds \\
&+ \int_t^\infty \|\exp(-(s-t)A)\, P_+\|_{\mathrm{op}} \Big(\varepsilon\, \|M(s)\|_{\mathrm{op}} + \frac{3\alpha}{16\Lambda}\, \|M(s) - N(s)\|_{\mathrm{op}}\Big)\, ds \\
&\leq \int_0^t \Lambda\, e^{-\alpha(t-s)} \Big(\varepsilon\, \|M(s)\|_{\mathrm{op}} + \frac{3\alpha}{16\Lambda}\, \|M(s) - N(s)\|_{\mathrm{op}}\Big)\, ds \\
&+ \int_t^\infty \Lambda\, e^{-\alpha(s-t)} \Big(\varepsilon\, \|M(s)\|_{\mathrm{op}} + \frac{3\alpha}{16\Lambda}\, \|M(s) - N(s)\|_{\mathrm{op}}\Big)\, ds \\
&\leq \frac{2\Lambda\varepsilon}{\alpha} \sup_{s \geq 0} \|M(s)\|_{\mathrm{op}} + \frac{3}{8} \sup_{s \geq 0} \|M(s) - N(s)\|_{\mathrm{op}}
\end{aligned}$$

for all $t \geq 0$. Thus,

$$\sup_{t \geq 0} \|M(t) - N(t)\|_{\mathrm{op}} \leq \frac{2\Lambda\varepsilon}{\alpha} \sup_{t \geq 0} \|M(t)\|_{\mathrm{op}} + \frac{3}{8} \sup_{s \geq 0} \|M(t) - N(t)\|_{\mathrm{op}},$$

hence

$$\sup_{t \geq 0} \|M(t) - N(t)\|_{\mathrm{op}} \leq \frac{4\Lambda\varepsilon}{\alpha} \sup_{t \geq 0} \|M(t)\|_{\mathrm{op}}.$$

Putting $t = 0$, we obtain

$$\|P_+ M(0) - P_+ N(0)\|_{\mathrm{op}} \leq \frac{4\Lambda^2\varepsilon}{\alpha} \sup_{t \geq 0} \|M(t)\|_{\mathrm{op}}.$$

Since $\varepsilon > 0$ is arbitrary and $\sup_{t \geq 0} \|M(t)\|_{\mathrm{op}} < \infty$, the assertion follows. $\square$

Let W^s denote the graph of ψ, so that

$$W^s = \{x_- + \psi(x_-) : x_- \in U_-,\ \|x_-\| \leq r\}.$$

We have shown that W^s is a C^1 manifold which is tangential to the stable subspace U_- at the origin. We next show that any solution starting in W^s converges to the origin at an exponential rate:

Lemma 4.8. *Consider a vector $x_- \in U_-$ with $\|x_-\| \leq r$. Moreover, let $x_0 = \psi(x_-) + x_-$. Then the unique solution of the initial value problem*

$x'(t) = F(x(t))$, $x(0) = x_0$, is defined for all $t \geq 0$ and satisfies

$$\sup_{t \geq 0} e^{\frac{\alpha t}{2}} \, \|x(t)\| < \infty.$$

In other words, the solution $x(t)$ converges to the origin at an exponential rate.

Proof. By Lemma 4.3 and Lemma 4.4, there exists a unique function $x(t)$ such that $\|x(t)\| \leq 2\Lambda \, e^{-\frac{\alpha t}{2}} \, \|x_-\|$ and

$$\begin{aligned} \exp(-tA)\, x(t) = x_- &+ \int_0^t \exp(-sA)\, P_- G(x(s))\, ds \\ &- \int_t^\infty \exp(-sA)\, P_+ G(x(s))\, ds \end{aligned}$$

for all $t \geq 0$. This gives

$$\begin{aligned} \frac{d}{dt}(\exp(-tA)\, x(t)) &= \exp(-tA)\, P_- G(x(t)) + \exp(-tA)\, P_+ G(x(t)) \\ &= \exp(-tA)\, G(x(t)) \end{aligned}$$

for all $t \geq 0$. From this, we deduce that

$$x'(t) = Ax(t) + G(x(t)) = F(x(t))$$

for all $t \geq 0$. Therefore, $x(t)$ is a solution of the differential equation $x'(t) = F(x(t))$. Moreover, $P_- x(0) = x_-$ and $P_+ x(0) = \psi(x_-)$ by definition of ψ. Thus $x(0) = \psi(x_-) + x_- = x_0$. Since $x(t)$ converges exponentially to 0, the assertion follows. $\square$

Lemma 4.9. *Consider a vector $x_0 \in \mathbb{R}^n$. Write $x_0 = x_+ + x_-$, where $x_+ \in U_+$ and $x_- \in U_-$. Assume that $\|x_+\| \leq r$, $\|x_-\| \leq r$, and $x_+ \neq \psi(x_-)$. Then the solution of the initial value problem $x'(t) = F(x(t))$, $x(0) = x_0$, must leave the ball $B_{2\Lambda r}(0)$ at some point.*

Proof. Let $x(t)$ be the solution of the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$. Suppose that $x(t)$ remains in the ball $B_{2\Lambda r}(0)$ for all time. Using the identity

$$x'(t) = Ax(t) + G(x(t)),$$

we obtain

$$\exp(-tA)\, x(t) = x_0 + \int_0^t \exp(-sA)\, G(x(s))\, ds \tag{7}$$

for all $t \geq 0$. This implies

$$\exp(-tA)\, P_+ x(t) = x_+ + \int_0^t \exp(-sA)\, P_+ \, G(x(s))\, ds.$$

Since $\|x(t)\|$ is bounded and $\exp(-tA)\,P_+ \to 0$ as $t \to \infty$, we obtain

$$0 = x_+ + \int_0^\infty \exp(-sA)\,P_+ G(x(s))\,ds. \tag{8}$$

Subtracting (8) from (7), we obtain

$$\begin{aligned} \exp(-tA)\,x(t) = x_- &+ \int_0^t \exp(-sA)\,G(x(s))\,ds \\ &- \int_0^\infty \exp(-sA)\,P_+ G(x(s))\,ds, \end{aligned}$$

hence

$$\begin{aligned} \exp(-tA)\,x(t) = x_- &+ \int_0^t \exp(-sA)\,P_- G(x(s))\,ds \\ &- \int_t^\infty \exp(-sA)\,P_+ G(x(s))\,ds \end{aligned}$$

for all $t \geq 0$. Therefore, $x(t)$ is a solution of the integral equation (5). This implies $\psi(x_-) = P_+ x(0) = x_+$, contrary to our assumption. Thus, the solution of the initial value problem $x'(t) = F(x(t))$, $x(0) = x_0$, must eventually leave the ball $B_{2\Lambda r}(0)$. $\square$

To summarize, we have proved the following theorem:

Theorem 4.10. *Suppose that $F : U \to \mathbb{R}^n$ is of class C^1. Moreover, suppose that $0 \in U$ is a hyperbolic equilibrium point. Then there exists a submanifold W^s of class C^1 with the following properties:*

(i) *The origin lies on W^s and the tangent space to W^s at 0 is given by U_-. In particular, $\dim W^s = \dim U_-$.*

(ii) *If $x_0 \in W^s$ is sufficiently close to 0, then the unique solution of the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$ converges exponentially to 0.*

(iii) *If $x_0 \notin W^s$ is sufficiently close to 0, then the unique solution of the differential equation $x'(t) = F(x(t))$ with initial condition $x(0) = x_0$ will leave the ball $B_{2\Lambda r}(0)$ at some time in the future.*

Note that, in part (iii), the solution may re-enter the ball $B_{2\Lambda r}(0)$ at some later time, and it may still converge to the equilibrium point as $t \to \infty$, but it cannot do so without first leaving the ball $B_{2\Lambda r}(0)$.

4.3. Lyapunov's theorems

Theorem 4.11. *Suppose that $\bar{x}$ is an equilibrium point of the autonomous system $x'(t) = F(x(t))$. Moreover, suppose that L is a smooth function* *defined on some ball $B_r(\bar{x})$, which has a strict local minimum at $\bar{x}$. Finally, we assume that $\langle \nabla L(x), F(x)\rangle \leq 0$ for all $x \in B_r(\bar{x})$. Then $\bar{x}$ is stable.*

Proof. Let $\varepsilon \in (0, r)$ be given. Since L has a strict local minimum at $\bar{x}$, we have

$$L(\bar{x}) < \inf_{\partial B_\varepsilon(\bar{x})} L.$$

By continuity, we can find a real number $\delta \in (0, \varepsilon)$ such that

$$\sup_{B_\delta(\bar{x})} L < \inf_{\partial B_\varepsilon(\bar{x})} L.$$

We claim that $\varphi_t(x) \in B_\varepsilon(\bar{x})$ for all $x \in B_\delta(\bar{x})$ and all $t \geq 0$. To prove this, we argue by contradiction. Fix a point $x_0 \in B_\delta(\bar{x})$, and suppose that $\varphi_t(x_0) \notin B_\varepsilon(\bar{x})$ for some $t \geq 0$. Let

$$\tau = \inf\{t \geq 0 : \varphi_t(x_0) \notin B_\varepsilon(\bar{x})\}.$$

Clearly, $\varphi_\tau(x_0) \in \partial B_\varepsilon(\bar{x})$. Moreover, for $t \in (0, \tau)$, we have $\varphi_t(x_0) \in B_\varepsilon(\bar{x})$. This implies

$$\frac{d}{dt} L(\varphi_t(x_0)) = \langle \nabla L(\varphi_t(x_0)), F(\varphi_t(x_0))\rangle \leq 0$$

for all $t \in (0, \tau)$. Thus, we conclude that

$$\inf_{\partial B_\varepsilon(\bar{x})} L \leq L(\varphi_\tau(x_0)) \leq L(x_0) \leq \sup_{B_\delta(\bar{x})} L.$$

This contradicts our choice of δ. $\square$

Theorem 4.12. *Suppose that $\bar{x}$ is an equilibrium point of the autonomous system $x'(t) = F(x(t))$. Moreover, suppose that L is a smooth function defined on some ball $B_r(\bar{x})$, which has a strict local minimum at $\bar{x}$. Finally, we assume that $\langle \nabla L(x), F(x)\rangle < 0$ for all $x \in B_r(\bar{x}) \setminus \{\bar{x}\}$. Then $\bar{x}$ is asymptotically stable.*

Proof. It follows from Theorem 4.11 that $\bar{x}$ is stable. Suppose that $\bar{x}$ is not asymptotically stable. Let us fix a real number $\delta > 0$ such that $\varphi_t(x) \in B_{\frac{r}{2}}(\bar{x})$ for all $x \in B_\delta(\bar{x})$ and all $t \geq 0$. Since $\bar{x}$ is not asymptotically stable, we can find a point $x_0 \in B_\delta(\bar{x})$ such that $\limsup_{t\to\infty} \|\varphi_t(x_0) - \bar{x}\| > 0$. Consequently, we can find a sequence of times $s_k \to \infty$ such that $\liminf_{t\to\infty} \|\varphi_{s_k}(x_0) - \bar{x}\| > 0$. After passing to a subsequence if necessary, we may assume that the sequence $\varphi_{s_k}(x_0)$ converges to some point $y \in B_r(\bar{x}) \setminus \bar{x}$.

Since $\langle \nabla L(x), F(x)\rangle \leq 0$ for all $x \in B_r(\bar{x})$, the function $t \mapsto L(\varphi_t(x_0))$ is monotone decreasing. Hence, the limit $\lambda := \lim_{t\to\infty} L(\varphi_t(x_0))$ exists. This gives

$$L(\varphi_t(y)) = \lim_{k\to\infty} L(\varphi_t(\varphi_{s_k}(x_0))) = \lim_{k\to\infty} L(\varphi_{s_k+t}(x_0)) = \lambda.$$

Therefore, the function $L(\varphi_t(y))$ is constant. In particular,

$$\langle \nabla L(y), F(y) \rangle = \frac{d}{dt} L(\varphi_t(y))\Big|_{t=0} = 0,$$

which contradicts our assumption. $\square$

4.4. Gradient and Hamiltonian systems

Finding Lyapunov functions is a difficult task in general. However, there are certain classes of nonlinear system which always admit a monotone quantity. In the following, we assume that $U \subset \mathbb{R}^n$ is an open set and $F : U \to \mathbb{R}^n$ is continuously differentiable mapping.

Definition 4.13. We say that the system $x'(t) = F(x(t))$ is a gradient system if $F(x) = -\nabla V(x)$ for some real-valued function V.

There is a convenient criterion for deciding whether a given system is a gradient or Hamiltonian system. A necessary condition for the system $x'(t) = F(x(t))$ to be a gradient system is that $\frac{\partial F_j}{\partial x_i}(x) = \frac{\partial F_i}{\partial x_j}(x)$ for all $x \in U$. Moreover, if the domain U is simply connected, then this condition is a sufficient condition for $x'(t) = F(x(t))$ to be a gradient system.

Proposition 4.14. *Consider a gradient system of the form $x'(t) = F(x(t))$, where $F(x) = -\nabla V(x)$. Moreover, suppose that the function V attains a strict local minimum at $\bar{x}$. Then $\bar{x}$ is a stable equilibrium point.*

Proof. Since

$$\langle \nabla V(x), F(x) \rangle = -\|\nabla V(x)\|^2 \leq 0,$$

the function $V(x)$ is a Lyapunov function. Consequently, $\bar{x}$ is a stable equilibrium point. $\square$

Definition 4.15. Assume that n is even, and let

$$J = \begin{bmatrix} 0 & 1 & & & & & \\ -1 & 0 & & & & & \\ & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{bmatrix}.$$

We say that the system $x'(t) = F(x(t))$ is Hamiltonian if $F(x) = J\,\nabla H(x)$ for some real-valued function H.

Proposition 4.16. *Consider a gradient system of the form $x'(t) = F(x(t))$, where $F(x) = -J\,\nabla H(x)$. Moreover, suppose that the function H attains a strict local minimum at $\bar{x}$. Then $\bar{x}$ is a stable equilibrium point.*

Proof. Since J is anti-symmetric, we have

$$\langle \nabla H(x), F(x) \rangle = -\langle \nabla H(x), J\, \nabla H(x) \rangle = 0.$$

Therefore, the function $H(x)$ is a Lyapunov function. Consequently, $\bar{x}$ is a stable equilibrium point. □

4.5. Problems

Problem 4.1. Let U be an open set in $\mathbb{R}^n$ containing 0, and let $F : U \to \mathbb{R}^n$ be a smooth vector field. Assume that 0 is an equilibrium point of the system $x'(t) = F(x(t))$, and let $A = DF(0)$ be the differential of F at 0. Suppose that all eigenvalues of A have negative real part. Show that there exists a quadratic function $L(x) = \langle Sx, x \rangle$ such that $L(x)$ has a strict local minimum at 0 and

$$\langle \nabla L(x), F(x) \rangle < 0$$

if $x \in U \setminus \{0\}$ is sufficiently close to 0. Conclude from this that 0 is asymptotically stable. (Hint: Use Problem 2.6.)

Problem 4.2. Consider the system

$$\begin{aligned} x_1'(t) &= x_2(t)^2 + x_1(t)\, x_2(t) - 2 \\ x_2'(t) &= x_1(t)^2 + x_1(t)\, x_2(t) - 2. \end{aligned}$$

(i) Find all equilibrium points of this system. For each equilibrium point, decide whether or not it is hyperbolic. For each hyperbolic equilibrium point, decide whether it is a source, a sink, or a saddle.
(ii) Sketch the phase portrait of this system.
(iii) Find the stable and unstable curves for each saddle point. (Hint: Consider the functions $x_1(t) + x_2(t)$ and $x_1(t) - x_2(t)$.)

Problem 4.3. This example is taken from [**4**]. Let us consider the system

$$\begin{aligned} x_1'(t) &= x_1(t) + x_1(t)x_2(t) - \sqrt{x_1(t)^2 + x_2(t)^2}\, (x_1(t) + x_2(t)) \\ x_2'(t) &= x_2(t) - x_1(t)^2 + \sqrt{x_1(t)^2 + x_2(t)^2}\, (x_1(t) - x_2(t)). \end{aligned}$$

(i) Let us write $x_1(t) = r(t) \cos\theta(t)$ and $x_2(t) = r(t) \sin\theta(t)$. Show that $r(t)$ and $\theta(t)$ satisfy the differential equations

$$\begin{aligned} r'(t) &= r(t)\, (1 - r(t)) \\ \theta'(t) &= r(t)\, (1 - \cos\theta). \end{aligned}$$

(ii) Show that any solution that originates in a neighborhood of $(1, 0)$ will converge to $(1, 0)$ as $t \to \infty$.
(iii) Show that the equilibrium point $(1, 0)$ is unstable.

Chapter 5

Limit sets of dynamical systems and the Poincaré-Bendixson theorem

5.1. Positively invariant sets

Definition 5.1. Let us consider the ODE $x'(t) = F(x(t))$, where $F : U \to \mathbb{R}^n$ is continuously differentiable. We say that a set $A \subset U$ is invariant under the flow generated by this ODE if $\varphi_t(x) \in A$ for all $x \in A$ and all $t \in \mathbb{R}$ for which $\varphi_t(x)$ is defined. Moreover, we say that a set $A \subset U$ is positively invariant if $\varphi_t(x) \in A$ for all $x \in A$ and all $t \in [0, \infty)$ for which $\varphi_t(x)$ is defined.

In this section, we show that a set is positively invariant under the ODE $x'(t) = F(x(t))$ if the vector field F points inward along the boundary.

Theorem 5.2. *Let $F : U \to \mathbb{R}^n$ be continuously differentiable, and let $A \subset U$ be relatively closed so that $\bar{A} \cap U = A$. Then the following statements are equivalent:*

(i) *A is positively invariant under the ODE $x'(t) = F(x(t))$.*

(ii) *$\langle F(y), y - z\rangle \geq 0$ for all points $y \in A$, $z \in \mathbb{R}^n$ satisfying $\|y - z\| = \operatorname{dist}(z, A)$.*

Proof. We first show that (i) implies (ii). Assume that A is positively invariant. Moreover, suppose that $y \in A$ and $z \in \mathbb{R}^n$ are two points satisfying

$\|y - z\| = \operatorname{dist}(z, A)$. Let $x(t)$, $t \in [0, T)$, denote the unique solution of the ODE $x'(t) = F(x(t))$ with initial condition $x(0) = y$. Since A is positively invariant, we have

$$\|x(0) - z\| = \|y - z\| = \operatorname{dist}(z, A) \leq \|x(t) - z\|$$

for all $t \in [0, T)$. This implies

$$\begin{aligned} 0 \leq \frac{d}{dt}(\|x(t) - z\|^2)\Big|_{t=0} &= 2\,\langle x'(0), x(0) - z\rangle \\ &= 2\,\langle F(x(0)), x(0) - z\rangle \\ &= 2\,\langle F(y), y - z\rangle. \end{aligned}$$

This shows that $\langle F(y), y - z\rangle \geq 0$, as claimed.

We now show that (ii) implies (i). Let us assume that condition (ii) holds, and let $x(t)$, $t \in [0, T)$, be a solution of the ODE $x'(t) = F(x(t))$ with $x(0) \in A$. We claim that $x(t) \in A$ for all $t \in [0, T)$. To prove this, we argue by contradiction. Suppose that $x(t_0) \notin A$ for some $t_0 \in (0, T)$. Let

$$C = \{x(t) : t \in [0, t_0]\}.$$

Note that C is a compact subset of U. Let us fix positive real numbers r and L such that $\operatorname{dist}(C, \mathbb{R}^n \setminus U) > 2r$ and

$$\sup_{\operatorname{dist}(x,C) \leq r} \|DF(x)\|_{\mathrm{op}} < L.$$

By assumption, $x(t_0) \notin A$. We now define

$$\tau = \sup\{t \in [0, t_0] : e^{-Lt} \operatorname{dist}(x(t), A) \leq e^{-Lt_0}\,\delta\},$$

where δ is chosen so that $0 < \delta < \min\{\operatorname{dist}(x(t_0), A), r\}$. Clearly, $\tau \in (0, t_0)$. Since A is closed, we can find a point $y \in \bar{A}$ such that $\|x(\tau) - y\| = \operatorname{dist}(x(\tau), A)$. By definition of τ,

$$e^{-L\tau}\,\|x(\tau) - y\| = e^{-L\tau} \operatorname{dist}(x(t), A) = e^{-Lt_0}\,\delta$$

and

$$e^{-Lt}\,\|x(t) - y\| \geq e^{-Lt} \operatorname{dist}(x(t), A) \geq e^{-Lt_0}\,\delta$$

for all $t \in [\tau, t_0)$. From this, we deduce that

$$\frac{d}{dt}(e^{-2Lt}\,\|x(t) - y\|^2)\Big|_{t=\tau} \geq 0,$$

hence

$$\langle F(x(\tau)), x(\tau) - y\rangle \geq L\,\|x(\tau) - y\|^2.$$

Moreover, $\|x(\tau) - y\| = e^{-L(t_0 - \tau)}\,\delta < r$. Since $\operatorname{dist}(C, \mathbb{R}^n \setminus U) > 2r$, it follows that $y \in \bar{A} \cap U = A$. Hence, the assumptions imply that

$$\langle F(y), y - x(\tau)\rangle \geq 0.$$

Adding both inequalities, we conclude that

$$\langle F(x(\tau)) - F(y), x(\tau) - y\rangle \geq L\,\|x(\tau) - y\|^2.$$

By definition of L, we have $\|DF\|_{\text{op}} < L$ at each point on the line segment joining $x(\tau)$ and y. This implies

$$\langle F(x(\tau)) - F(y), x(\tau) - y\rangle \leq \|F(x(\tau)) - F(y)\|\,\|x(\tau) - y\| < L\,\|x(\tau) - y\|^2.$$

This is a contradiction. $\square$

Corollary 5.3. *Let $F : U \to \mathbb{R}^n$ be continuously differentiable, and let A be a closed subset of U with smooth boundary. Let $\nu(y)$ denote the outward-pointing unit normal vector at a point $y \in \partial A$. Then the following statements are equivalent:*

(i) *A is positively invariant under the ODE $x'(t) = F(x(t))$.*

(ii) *$\langle F(y), \nu(y)\rangle \leq 0$ for all points $y \in \partial A$.*

It is often useful to consider the domains with piecewise smooth boundary. For example, for planar domains we have the following result:

Corollary 5.4. *Let $F : U \to \mathbb{R}^2$ be continuously differentiable, and let A be a closed subset of U. We assume that the boundary ∂A is piecewise smooth. Let E denote the set of corners of ∂A. For each point $y \in \partial A \setminus E$, we denote by $\nu(y)$ the outward-pointing unit normal vector at y. Then the following statements are equivalent:*

(i) *A is positively invariant under the ODE $x'(t) = F(x(t))$.*

(ii) *$\langle F(y), \nu(y)\rangle \leq 0$ for all points $y \in \partial A \setminus E$.*

Proof. It is easy to see that (i) implies (ii). We now show that (ii) implies (i). By Theorem 5.2, it suffices to show that $\langle F(y), y - z\rangle \geq 0$ for all points $y \in A$, $z \in \mathbb{R}^n$ satisfying $\|y - z\| = \text{dist}(z, A)$. To verify this, we distinguish two cases:

Case 1: Suppose first that $y \notin E$. Clearly, y must lie on the boundary ∂A. Since the point y has smallest distance from z among all points in A, it follows that the vector $z - y = \alpha\,\nu(y)$ for some number $\alpha \geq 0$. This implies $\langle F(y), y - z\rangle = -\alpha\,\langle F(y), \nu(y)\rangle \leq 0$.

Case 2: Suppose finally that $y \in E$. By assumption, the point y lies on two boundary arcs, which we denote by Γ_1 and Γ_2. Let ν_1 denote the outward pointing unit normal vector field along Γ_1, and let ν_2 be the outward-pointing unit normal vector field along Γ_1. Since y has smallest distance from z among all points in A, we conclude that $z - y = \alpha_1\,\nu_1(y) + \alpha_2\,\nu_2(y_2)$ for some numbers $\alpha_1, \alpha_2 \geq 0$. Using our assumption (and the continuity of F), we conclude that $\langle F(y_1), \nu(y_1)\rangle \leq 0$ and $\langle F(y_2), \nu(y_2)\rangle \leq 0$. This implies

$\langle F(y), y - z \rangle = -\alpha_1 \langle F(y), \nu_1(y) \rangle - \alpha_2 \langle F(y), \nu_2(y) \rangle \leq 0$. This completes the proof. □

5.2. The ω-limit set of a trajectory

Definition 5.5. Let $F : U \to \mathbb{R}^n$ be continuously differentiable, and let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is defined for all $t \geq 0$. A point $y \in \mathbb{R}^n$ is an ω-limit point of the trajectory $x(t)$ if there exists a sequence of times $s_k \to \infty$ such that $\lim_{k \to \infty} x(s_k) = y$.

If we denote by Ω denote the set of all ω-limit points of $x(t)$, then

$$\Omega = \bigcap_{t \geq 0} \overline{\{x(s) : s \in [t, \infty)\}},$$

where $\overline{\{x(s) : s \in [t, \infty)\}}$ denotes the closure of the set $\{x(s) : s \in [t, \infty)\}$. Since intersections of closed sets are always closed, we conclude that Ω is a closed subset of $\mathbb{R}^n$.

Proposition 5.6. *If Ω is bounded and non-empty, then $\lim_{t \to \infty} \operatorname{dist}(x(t), \Omega) = 0$.*

Proof. Suppose this is false. Then there exists a sequence of times $t_k \to \infty$ such that $\operatorname{dist}(x(t_k), \Omega) > \varepsilon$ for all k. Let y be an arbitrary point in Ω. Since y is an ω-limit point of the trajectory $x(t)$, we can find a sequence of numbers s_k such that $s_k > t_k$ and $\|x(s_k) - y\| < \varepsilon$ for all k. This implies $\operatorname{dist}(x(s_k), \Omega) < \varepsilon$ for all k. By the intermediate value theorem, we can find a sequence of times $\tau_k \in (t_k, s_k)$ such that $\operatorname{dist}(x(\tau_k), \Omega) = \varepsilon$. Since Ω is bounded, it follows that the sequence $x(\tau_k)$ is bounded. By the Bolzano-Weierstrass theorem, we can find a sequence of integers $k_l \to \infty$ such that the sequence $x(\tau_{k_l})$ converges to some point $z \in \mathbb{R}^n$ as $l \to \infty$. Consequently, z is an ω-limit point and $\operatorname{dist}(z, \Omega) = \lim_{l \to \infty} \operatorname{dist}(x(\tau_{k_l}, \Omega) \geq \varepsilon$. This is a contradiction. □

Proposition 5.7. *If Ω is bounded, then Ω is connected.*

Proof. Suppose that Ω is not connected. Then there exist non-empty closed sets A_1 and A_2 such that $A_1 \cup A_2 = \Omega$ and $A_1 \cap A_2 = \emptyset$. Since Ω is bounded, the sets A_1 and A_2 are compact. Since A_1 and A_2 are disjoint, it follows that $\operatorname{dist}(A_1, A_2) > 0$.

Let us fix two arbitrary points $y_1 \in A_1$ and $y_2 \in A_2$. Since y_1 is an ω-limit point of $x(t)$, we can find a sequence of times $s_{k,1} \to \infty$ such that $\lim_{k \to \infty} x(s_{k,1}) = y_1$. Similarly, we can find a sequence of times $s_{k,2} \to \infty$ such that $\lim_{k \to \infty} x(s_{k,2}) = y_2$. In particular, $\lim_{k \to \infty} \operatorname{dist}(x(s_{k,1}), A_1) = 0$

and $\lim_{k\to\infty} \text{dist}(x(s_{k,2}), A_1) = \text{dist}(y_2, A_1) \geq \text{dist}(A_1, A_2)$. By the intermediate value theorem, we can find a sequence of times $\tau_k \to \infty$ such that

$$\text{dist}(x(\tau_k), A_1) = \frac{1}{2}\,\text{dist}(A_1, A_2).$$

Using the triangle inequality, we obtain

$$\text{dist}(x(\tau_k), A_2) \geq \text{dist}(A_1, A_2) - \text{dist}(x(\tau_k), A_1) \geq \frac{1}{2}\,\text{dist}(A_1, A_2).$$

This implies

$$\text{dist}(x(\tau_k), \Omega) = \min\{\text{dist}(x(\tau_k), A_1), \text{dist}(x(\tau_k), A_2)\} = \frac{1}{2}\,\text{dist}(A_1, A_2).$$

This contradicts Proposition 5.6. □

Proposition 5.8. *The set $\Omega \cap U$ is an invariant set.*

Proof. Consider a $y \in \Omega \cap U$. Then there exists a sequence of times $s_k \to \infty$ such that $\lim_{k\to\infty} x(s_k) = y$. Hence, if $t \in \mathbb{R}$ is fixed, then

$$\lim_{k\to\infty} x(s_k + t) = \lim_{k\to\infty} \varphi_t(x(s_k)) = \varphi_t(y).$$

Thus, $\varphi_t(y) \in \Omega$ for all $t \in \mathbb{R}$. This proves the assertion. □

Finally, we show that certain equilibrium points cannot arise as ω-limit points.

Proposition 5.9. *Let $x(t)$, $t \geq 0$, be a solution of the differential equation $x'(t) = F(x(t))$ and let $\bar{x}$ be an ω-limit point of $x(t)$. Moreover, we assume that $\bar{x}$ is a stable equilibrium point for the system $x'(t) = -F(x(t))$. Then $x(t) = \bar{x}$ for all $t \geq 0$.*

Proof. As usual, we denote by φ_t the flow generated by the differential equation $x'(t) = F(x(t))$. Let us fix a real number $\varepsilon > 0$. By assumption, we can find a real number $\delta > 0$ such that $\varphi_{-t}(x) \in B_\varepsilon(\bar{x})$ for all $x \in B_\delta(\bar{x})$ and all $t \geq 0$. Moreover, since $\bar{x}$ is an ω-limit point, we can find a sequence of times $s_k \to \infty$ such that $\lim_{k\to\infty} x(s_k) = \bar{x}$. In particular, $x(s_k) \in B_\delta(\bar{x})$ if k is sufficiently large. This implies $x(0) = \varphi_{-s_k}(x(s_k)) \in B_\varepsilon(\bar{x})$. Since $\varepsilon > 0$ is arbitrary, we conclude that $x(0) = \bar{x}$. From this, the assertion follows. □

As an application, we can prove a global version of Lyapunov's second theorem:

Proposition 5.10. *Consider the differential equation $x'(t) = F(x(t))$, where $F : U \to \mathbb{R}^n$ is continuously differentiable. We assume that there exists a function $L : U \to \mathbb{R}^n$ with the following properties:*

(i) *For each* $\mu \in \mathbb{R}$*, the sublevel set* $\{x \in U : L(x) \leq \mu\}$ *is a compact subset of* U*.*

(ii) *We have* $\langle \nabla L(x), F(x) \rangle < 0$ *for all points* $x \in U \setminus E$*, where* E *consists of isolated points.*

Then, given any point $x_0 \in U$*, the function* $\varphi_t(x_0)$ *is defined for all* $t \geq 0$ *and converges to an equilibrium point.*

Proof. Fix a point $x_0 \in U$, and let $\mu = L(x_0)$. Since $\langle \nabla L(x), F(x) \rangle \leq 0$ for all $x \in U$, we conclude that the function $t \mapsto L(\varphi_t(x_0))$ is monotone decreasing. In particular, $L(\varphi_t(x_0)) \leq \mu$ for all $t \geq 0$. Since the sublevel set $\{x \in U : L(x) \leq \mu\}$ is a compact subset of U, we conclude that the solution exists for all $t \geq 0$. Since the function $t \mapsto L(\varphi_t(x_0))$ is monotone decreasing, the limit $\lambda := \lim_{t \to \infty} L(\varphi_t(x_0))$ exists.

Let Ω denote the ω-limit set of the trajectory $\varphi_t(x_0)$. Given any point $y \in \Omega$, we can find a sequence of times $s_k \to \infty$ such that $\lim_{k \to \infty} \varphi_{s_k}(x_0) = y$. This implies

$$L(\varphi_t(y)) = \lim_{k \to \infty} L(\varphi_t(\varphi_{s_k}(x_0))) = \lim_{k \to \infty} L(\varphi_{s_k + t}(x_0)) = \lambda.$$

Consequently, the function $L(\varphi_t(y))$ is constant. This implies

$$\langle \nabla L(y), F(y) \rangle = \frac{d}{dt} L(\varphi_t(y)) \Big|_{t=0} = 0.$$

Consequently,

$$\Omega \subset \{x \in U : \langle L(x), F(x) \rangle = 0\} \subset E.$$

In particular, Ω consists of at most finitely many points. Since Ω is connected, we conclude that Ω consists of at most a single point. On the other hand, since the trajectory $\varphi_t(x_0)$ is contained in a compact set, the set Ω is non-empty by the Bolzano-Weierstrass theorem. Thus, Ω consists of exactly one point. Moreover, this point must be an equilibrium point since Ω is a positively invariant set. Hence, the assertion follows from Proposition 5.6. $\square$

5.3. ω-limit sets of planar dynamical systems

In this section, we will consider a planary dynamical system of the form $x'(t) = F(x(t))$, where $F : \mathbb{R}^2 \times \mathbb{R}^2$. Our goal is to analyze the ω-limit set of a given solution $x(t)$ of the ODE $x'(t) = F(x(t))$.

Definition 5.11. A line segment $S = \{\lambda z_0 + (1 - \lambda) z_1 : \lambda \in (0, 1)\}$ is said to be transversal if, for each point $x \in \bar{S}$, the vector $F(x)$ is not parallel to the vector $z_1 - z_0$.

Lemma 5.12. *Let S be a transversal line segment, and let $\bar{x}$ be an arbitrary point in S. Then there exists an open neighborhood U of $\bar{x}$ and a smooth function $h : U \to \mathbb{R}$ such that $h(\bar{x}) = 0$ and $\varphi_{h(x)}(x) \in S$ for all points $x \in U$.*

Proof. This follows immediately from the implicit function theorem. □

Lemma 5.13. *Let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is defined on some time interval J, and let S be a transversal line segment. Then the set $E = \{t \in J : x(t) \in S\}$ is discrete.*

Proof. Suppose this is false. Then there exists a number $\tau \in \mathbb{R}$ and a sequence of numbers $t_k \in E \setminus \{\tau\}$ such that $\lim_{k\to\infty} t_k = \tau$. Clearly, $x(t_k) \in S$ and $x(\tau) = \lim_{k\to\infty} x(t_k) \in \bar{S}$. Hence, the vector $\frac{x(\tau)-x(t_k)}{\tau-t_k}$ is tangential to S. Consequently, the limit $\lim_{k\to\infty} \frac{x(\tau)-x(t_k)}{\tau-t_k} = x'(\tau) = F(x(\tau))$ is also tangential to S. This contradicts the fact that S is a transversal line segment. □

The following monotonicity property plays a fundamental role in the proof of the Poincaré-Bendixson theorem:

Lemma 5.14. *Let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is not periodic, and let $S = \{\lambda z_0 + (1-\lambda) z_1 : \lambda \in (0,1)\}$ be a transversal line segment. Let us consider three times $t_0 < t_1 < t_2$ with the property that $x(t) \in S$ for $t \in \{t_0, t_1, t_2\}$ and $x(t) \notin S$ for all $t \in (t_0, t_1) \cup (t_1, t_2)$. Moreover, let us write $x(t_i) = \lambda_i z_0 + (1-\lambda_i) z_1$ where $\lambda_i \in [0,1]$ for $i = 0, 1, 2$. Then either $\lambda_0 < \lambda_1 < \lambda_2$ or $\lambda_0 > \lambda_1 > \lambda_2$.*

Proof. Since $x(t)$ is not periodic, the numbers $\lambda_0, \lambda_1, \lambda_2$ are all distinct. Suppose now that the assertion is false. Then we either have $\lambda_1 > \max\{\lambda_0, \lambda_2\}$ or $\lambda_1 < \min\{\lambda_0, \lambda_2\}$. Without loss of generality, we may assume that $\lambda_1 > \max\{\lambda_0, \lambda_2\}$. (Otherwise, we switch the roles of z_0 and z_1.) Moreover, we may assume that $\lambda_0 < \lambda_2$. (Otherwise, we replace F by $-F$ and t_i by $-t_{2-i}$.) Consequently, $\lambda_0 < \lambda_2 < \lambda_1$.

Let

$$\Gamma = \{x(t) : t \in [t_0, t_1]\} \cup \{\lambda z_0 + (1-\lambda) z_1 : \lambda \in (\lambda_0, \lambda_1)\}.$$

In other words, Γ is the union of the trajectory from $x(t_0)$ to $x(t_1)$ with a line segment from $x(t_0)$ to $x(t_1)$. By assumption, $x(t) \notin S$ for all $t \in (t_0, t_1)$, so Γ is free of self-intersections. By the Jordan curve theorem, the complement $\mathbb{R}^2 \setminus \Gamma$ has exactly two connected components. Let us denote these components by D_1 and D_2.

By assumption, the vector $F(\lambda z_0 + (1-\lambda) z_1)$ cannot be parallel to S for any $\lambda \in [\lambda_0, \lambda_1]$. After switching the roles of D_1 and D_2 if necessary, we may assume that the vector $F(\lambda z_0 + (1-\lambda) z_1)$ points into D_1 for each $\lambda \in (\lambda_0, \lambda_1)$. Hence, at each point on the boundary $\partial D_1 = \Gamma$, the vector field F points inward or is tangential to ∂D_1. By Corollary 5.4, the closure $\bar{D}_1$ is a positively invariant set. Since $x(t_1) \in \bar{D}_1$, we conclude that $x(t) \in \bar{D}_1$ for all $t > t_1$. On the other hand, since the point $x(t_2)$ lies on the line segment $\{\lambda z_0 + (1-\lambda) z_1 : \lambda \in (\lambda_0, \lambda_1)\}$, we have $x(t_2 - \varepsilon) \notin \bar{D}_1$ if $\varepsilon > 0$ is sufficiently small. This is a contradiction. □

Lemma 5.15. *Let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is defined for all $t \geq 0$ and is not periodic. Moreover, let Ω denote its ω-limit set, and let $S = \{\lambda z_0 + (1-\lambda) z_1 : s \in [0,1]\}$ be a transversal line segment. Then the intersection $\Omega \cap S$ consists of at most one point.*

Proof. We argue by contradiction. Suppose that the intersection $\Omega \cap S$ contains two distinct points y_1 and y_2. Since $y_1 \in \Omega$, we can find a sequence of real numbers $s_{k,1} \to \infty$ such that $\lim_{k \to \infty} x(s_{k,1}) = y_1$. Moreover, there exists a sequence of real numbers $s_{k,2} \to \infty$ such that $\lim_{k \to \infty} x(s_{k,2}) = y_2$. Using Lemma 5.12, we can find a sequence of real numbers $\tilde{s}_{k,1}$ such that $\tilde{s}_{k,1} - s_{k,1} \to 0$ and $x(\tilde{s}_{k,1}) \in S$ for all k. Similarly, we obtain a sequence of real numbers $\tilde{s}_{k,2}$ such that $\tilde{s}_{k,2} - s_{k,2} \to 0$ and $x(\tilde{s}_{k,2}) \in S$ for all k.

By Lemma 5.13, the set $E = \{t \geq 0 : x(t) \in S\}$ is discrete. Moreover, since $\tilde{s}_{k,1}, \tilde{s}_{k,2} \in E$ for all k, it follows that E is unbounded. Let us write $E = \{t_k : k \in \mathbb{N}\}$, where t_k is an increasing sequence of times going to infinity. Since $x(t_k) \in S$, we may write $x(t_k) = \lambda_k z_0 + (1-\lambda_k) z_1$ for some $\lambda_k \in [0,1]$. By Lemma 5.14, the sequence λ_k is either monotone increasing or monotone decreasing. In either case, the limit $\lim_{k \to \infty} \lambda_k$ exists. This implies that the limit $\lim_{k \to \infty} x(t_k)$ exists. On the other hand, the sequences $\{x(\tilde{s}_{k,1}) : k \in \mathbb{N}\}$ and $\{x(\tilde{s}_{k,2} : k \in \mathbb{N}\}$ are subsequences of the sequence $\{x(t_k) : k \in \mathbb{N}\}$, and we have $\lim_{k \to \infty} x(\tilde{s}_{k,1}) = y_1$ and $\lim_{k \to \infty} x(\tilde{s}_{k,2}) = y_2$. Thus, $y_1 = y_2$, contrary to our assumption. □

Lemma 5.16. *Let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is defined for all $t \geq 0$ and is not periodic. Moreover, let Ω denote its ω-limit set. Assume that Ω is bounded, non-empty, and contains no equilibrium points. Then every point in Ω lies on a periodic orbit.*

Proof. Let us consider an arbitrary point $y_0 \in \Omega$, and let $y(t)$ be the unique maximal solution with $y(0) = y_0$. By Proposition 5.8, we have $y(t) \in \Omega$ for all t. Since Ω is a compact set, we conclude that the solution $y(t)$ is defined for all $t \in \mathbb{R}$. Let z be an ω-limit point of the solution $y(t)$. Since $y(t) \in \Omega$ for all $t \in \mathbb{R}$, we conclude that $z \in \Omega$. Consequently, z cannot be an equilibrium point. Hence, we can find a transversal line segment S such that $z \in S$. By Lemma 5.15, the set $\Omega \cap S$ consists of at most one point. Therefore, $\Omega \cap S = \{z\}$.

Since z is an ω-limit point of $y(t)$, we can find a sequence $s_k \to \infty$ such that $\lim_{k\to\infty} y(s_k) = z$. Using Lemma 5.12, we can find another sequence of times $\tilde{s}_k$ such that $\tilde{s}_k - s_k \to 0$ and $y(\tilde{s}_k) \in S$ for all k. This implies $y(\tilde{s}_k) \in \Omega \cap S$ for all k. Consequently, $y(\tilde{s}_k) = z$ for all k. From this, we deduce that the solution $y(t)$ is periodic. □

Poincaré-Bendixson Theorem. *Let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$, which is defined for all $t \geq 0$. Let Ω be the set of ω-limit points of the trajectory $x(t)$. Suppose that Ω is bounded, non-empty, and contains no equilibrium points. Then there exists a periodic solution $y(t)$ of the ODE $y'(t) = F(y(t))$ such that $\{y(t) : t \in \mathbb{R}\} = \Omega$.*

Proof. If $x(t)$ is periodic, the assertion is trivial. In the following, we will assume that $x(t)$ is not periodic. By Lemma 5.16, there exists a periodic solution $y(t)$ of the ODE $y'(t) = F(y(t))$ such that $A := \{y(t) : t \in \mathbb{R}\} \subset \Omega$.

We claim that the set $\Omega \setminus A$ is closed. To see this, we consider a sequence of points $z_k \in \Omega \setminus A$ such that $\lim_{k\to\infty} z_k = \bar{z}$. Clearly, $\bar{z} \in \Omega$ since Ω is closed. In particular, $\bar{z}$ cannot be an equilibrium point. Consequently, we can find a transversal line segment S such that $\bar{z} \in S$. By Lemma 5.15, the set $\Omega \cap S$ consists of at most one point. Thus, $\Omega \cap S = \{\bar{z}\}$.

By Lemma 5.12, there exists an open neighborhood U of $\bar{z}$ and a smooth function $h : U \to \mathbb{R}$ such that $h(\bar{z}) = 0$ and $\varphi_{h(z)}(z) \in S$ for all points $z \in U$. Since $\lim_{k\to\infty} z_k = \bar{z}$, we have $z_k \in U$ if k is sufficiently large. Then $\varphi_{h(z_k)}(z_k) \in S$. Since Ω is an invariant set and $z_k \in \Omega$, it follows that $\varphi_{h(z_k)}(z_k) \in \Omega \cap S$ if k is sufficiently large. From this, we deduce that $\varphi_{h(z_k)}(z_k) = \bar{z}$ for k sufficiently large. Since $z_k \notin A$, we conclude that $\bar{z} \notin A$. This shows that the set $\Omega \setminus A$ is closed.

To summarize, we have shown that Ω can be written as a disjoint union of the closed sets A and $\Omega \setminus A$. Since A is connected, we must have $\Omega \setminus A = \emptyset$. This completes the proof. □

5.4. Stability of periodic solutions and the Poincaré map

We now return to the n-dimensional case. Suppose that $x(t)$ is a non-constant periodic solution of the ODE $x'(t) = F(x(t))$, and let $\Gamma = \{x(t) : t \in \mathbb{R}\}$. Our goal in this section is to analyze whether this periodic orbit is stable; that is, whether a nearby solution $y(t)$ will converge to Γ as $t \to \infty$.

To simplify the notation, we will assume that $x(0) = 0$. Since 0 is not an equilibrium point, we can find an $(n-1)$-dimensional subspace $S \subset \mathbb{R}^n$ such that $F(0) \notin S$. By the implicit function theorem, we can find an open neighborhood U of 0 and a smooth function $h : U \to \mathbb{R}$ such that $h(0) = T$ and $\varphi_{h(y)}(y) \in S$ for all points $y \in U$. We next define a map $P : S \cap U \to S$

by

$$P(y) = \varphi_{h(y)}(y) \in S$$

for $y \in S \cap U$. Since $\varphi_T(0) = 0$ and $h(0) = T$, we obtain $P(0) = 0$. The map P is called the Poincaré map.

Theorem 5.17. *Suppose that the eigenvalues of $DP(0)$ all lie inside the unit circle in $\mathbb{C}$. If y_0 is sufficiently close to 0, then $\mathrm{dist}(\varphi_t(y_0), \Gamma) \to 0$ as $t \to \infty$.*

Proof. Let $A = DP(0)$ denote the differential of P at the origin. By assumption, the eigenvalues of A all lie inside the unit circle in $\mathbb{C}$. Using the $L+N$ decomposition, it is easy to see that $\lim_{k\to\infty} A^k = 0$ (cf. Problem 2.7 above). Let us fix a positive integer m such that $\|A^m\|_{\mathrm{op}} < \frac{1}{2}$.

For abbreviation, let $\tilde{T} = mT$. By the implicit function theorem, we can find an open neighborhood $\tilde{U}$ of 0 and a smooth function $\tilde{h} : \tilde{U} \to \mathbb{R}$ such that $\tilde{h}(0) = \tilde{T}$ and $\varphi_{\tilde{h}(y)}(y) \in S$ for all points $y \in \tilde{U}$. We now define a map $\tilde{P} : S \cap \tilde{U} \to S$ by

$$\tilde{P}(y) = \varphi_{\tilde{h}(y)}(y) \in S$$

for $y \in S \cap \tilde{U}$. Note that

$$\tilde{P}(y) = \underbrace{P \circ \ldots \circ P}_{m \text{ times}}(y)$$

if $y \in S$ is sufficiently close to the origin. This implies

$$D\tilde{P}(0) = DP(0)^m = A^m.$$

Consequently, $\|D\tilde{P}(0)\|_{\mathrm{op}} < \frac{1}{2}$ by our choice of m. Consequently, we can find a real number $\delta > 0$ such that $B_\delta(0) \subset \tilde{U}$ and $\|\tilde{P}(y)\| \leq \frac{1}{2}\|y\|$ for all points $y \in S \cap B_\delta(0)$.

We now consider an initial point $y_0 \in S$. We assume that y_0 is sufficiently close to the origin such that $y_0 \in \tilde{U}$ and $\|\varphi_{\tilde{h}(y_0)}(y_0)\| < \delta$. We inductively define a sequence of numbers t_k such that $t_0 = \tilde{h}(y_0)$ and $t_{k+1} = t_k + \tilde{h}(\varphi_{t_k}(y_0))$. Clearly,

$$\varphi_{t_{k+1}}(y_0) = \varphi_{\tilde{h}(\varphi_{t_k}(y_0))}(\varphi_{t_k}(y_0)) = \tilde{P}(\varphi_{t_k}(y_0)).$$

Using induction on k, we can show that $\|\varphi_{t_k}(y_0)\| < 2^{-k}\,\delta$. Therefore, $\varphi_{t_k}(y_0) \to 0$ as $k \to \infty$. This implies

$$t_{k+1} - t_k = \tilde{h}(\varphi_{t_k}(y_0)) \to \tilde{h}(0) = \tilde{T}$$

as $k \to \infty$. Since $\varphi_{t_k}(y_0) \to 0$ as $k \to \infty$, we conclude that

$$\sup_{t\in[t_k,t_{k+1}]} \|\varphi_t(y_0) - \varphi_{t-t_k}(0)\| = \sup_{t\in[0,t_{k+1}-t_k]} \|\varphi_t(\varphi_{t_k}(y_0)) - \varphi_t(0)\| \to 0$$

as $k \to \infty$. In particular,

$$\sup_{t \in [t_k, t_{k+1}]} \operatorname{dist}(\varphi_t(y_0), \Gamma) \to 0$$

as $k \to \infty$. From this, the assertion follows. $\square$

Finally, we describe how the differential of the Poincaré map is related to the differential of φ_T.

Proposition 5.18. *Let us define a linear transformation $Q : \mathbb{R}^n \to S$ by $Qy = y$ for all $y \in S$ and $Q\,F(0) = 0$. Then*

$$DP(0)\,y = Q\,D\varphi_T(0)\,y$$

for all $y \in S$.

Proof. Fix a vector $y \in S$. Using the chain rule, we obtain

$$\frac{d}{ds} P(sy)\Big|_{s=0} = \frac{d}{ds} \varphi_{h(sy)}(sy)\Big|_{s=0} = D\varphi_T(0)\,y + \kappa\,F(0),$$

where $\kappa = \frac{d}{ds} h(sy)\big|_{s=0}$. On the other hand, since $P(sy) \in S$ for all s, we conclude that $\frac{d}{ds} P(sy)\big|_{s=0} \in S$. Thus, we conclude that

$$\frac{d}{ds} P(sy)\Big|_{s=0} = Q\,(D\varphi_T(0)\,y + \kappa\,F(0)) = Q\,D\varphi_T(0)\,y.$$

From this, the assertion follows. $\square$

5.5. Problems

Problem 5.1. Let $p(x)$ be a polynomial of odd degree whose leading coefficient is positive. Moreover, let (x_0, y_0) be an arbitrary point in $\mathbb{R}^2$. Suppose that $(x(t), y(t))$ is the unique maximal solution of the system

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -y(t)^3 - p(x) \end{aligned}$$

with the initial condition $(x(0), y(0)) = (x_0, y_0)$.
(i) Show that $\sup_{t \geq 0} (x(t)^2 + y(t)^2) < \infty$. (Hint: Look for a monotone quantity of the form $f(x) + g(y)$.)
(ii) Show that the solution $(x(t), y(t))$ is defined for all $t \geq 0$.
(iii) Suppose that $(\bar{x}, \bar{y})$ is an ω-limit point of the trajectory $(x(t), y(t))$. Show that $p(\bar{x}) = \bar{y} = 0$.
(iv) Show that the ω-limit set of the trajectory $(x(t), y(t))$ consists of exactly one point.

Problem 5.2. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable mapping. Assume that $x(t)$ is a solution of the system $x'(t) = F(x(t))$ which is defined for all $t \geq 0$, and let Ω be its ω-limit set. We assume that Ω is bounded and

$$\{x \in \Omega : \|x\| \leq r\} = \{\lambda v : \lambda \in [0, r]\}$$

for some positive real number r and some unit vector v. The goal of this problem is to show that every point on the line segment $\{\lambda v : \lambda \in [0, r]\}$ is an equilibrium point.
(i) Fix a number $\lambda \in (0, r)$. Consider a sequence of times $s_k \to \infty$ such that $x(s_k) \to rv$, and define $\tau_k = \sup\{t \in [0, s_k] : \|x(t)\| \leq r \text{ and } \langle x(t), v\rangle \leq \lambda\}$. Show that $\tau_k \to \infty$, $x(\tau_k) \to \lambda v$, and $\langle F(x(\tau_k)), v\rangle \geq 0$. Deduce from this that $\langle F(\lambda v), v\rangle \geq 0$.
(ii) Fix a number $\lambda \in (0, r)$. Consider a sequence of times $s_k \to \infty$ such that $x(s_k) \to 0$, and define $\tau_k = \sup\{t \in [0, s_k] : \|x(t)\| \geq r \text{ or } \langle x(t), v\rangle \geq \lambda\}$. Show that $\tau_k \to \infty$, $x(\tau_k) \to \lambda v$, and $\langle F(x(\tau_k)), v\rangle \leq 0$. Deduce from this that $\langle F(\lambda v), v\rangle \leq 0$.
(iii) Show that $F(\lambda v) = 0$ for all $\lambda \in (0, r)$. In other words, every point on this line segment is an equilibrium point.

Problem 5.3. Let F be a vector field on $\mathbb{R}^2$, and let $x(t)$ be a solution of the ODE $x'(t) = F(x(t))$ which is defined for all $t \in \mathbb{R}$ and is not periodic. Suppose that $\bar{x} \in \mathbb{R}^2$ is both an α-limit point and an ω-limit point of the trajectory $x(t)$. In other words, there exists a family of times $\{s_k : k \in \mathbb{Z}\}$ such that $\lim_{k\to\infty} x(s_k) = \lim_{k\to-\infty} x(s_k) = \bar{x}$. Show that $\bar{x}$ is an equilibrium point. (Hint: Consider a transversal line segment passing through $\bar{x}$ and use the monotonicity property.)

Problem 5.4. Let us consider the unique maximal solution of the system

$$\begin{aligned} x_1'(t) &= \frac{(1 - x_1(t)^2)^2}{1 + x_1(t)^2}\,(x_1(t) + (1 - x_1(t)^2)\,x_2(t)) \\ x_2'(t) &= -x_1(t) + (1 - x_1(t)^2)\,x_2(t) \end{aligned}$$

with initial condition $(x_1(0), x_2(0)) = (0, 1)$.
(i) Let us write $y_1(t) = \frac{x_1(t)}{1 - x_1(t)^2}$ and $y_2(t) = x_2(t)$. Show that

$$\begin{aligned} y_1'(t) &= \frac{2}{1 + \sqrt{1 + 4y_1(t)^2}}\,(y_1(t) + y_2(t)) \\ y_2'(t) &= \frac{2}{1 + \sqrt{1 + 4y_1(t)^2}}\,(y_2(t) - y_1(t)). \end{aligned}$$

(ii) Show that the ω-limit set of $(x_1(t), x_2(t))$ is a union of two parallel lines.

Problem 5.5. Consider the system

$$x_1'(t) = (1 - x_1(t)^2)\,(x_1(t) + 2\,x_2(t))$$
$$x_2'(t) = (1 - x_2(t)^2)\,(x_2(t) - 2\,x_1(t)).$$

(i) Show that the square $Q = [-1,1] \times [-1,1]$ is an invariant set for this system.
(ii) Let $(x_1(t), x_2(t))$ denote the unique maximal solution of this system with initial condition $(x_1(0), x_2(0)) = (\frac{1}{2}, 0)$, and let Ω denote its ω-limit set. Show that $\Omega \subset \partial Q$. (Hint: Consider the function $(1 - x_1(t)^2)(1 - x_2(t)^2)$.)
(iii) Show that Ω cannot consist of a single point. (Hint: Suppose that $(x_1(t), x_2(t))$ converges to an equilibrium point, and use the stable manifold theorem to arrive at a contradiction.)
(iv) Let $L_1 = \{1\} \times [-1,1]$, $L_2 = [-1,1] \times \{1\}$, $L_3 = \{-1\} \times [-1,1]$, and $L_4 = [-1,1] \times \{-1\}$. Show that $\Omega = \bigcup_{j \in J} L_j$ for some non-empty set $J \subset \{1,2,3,4\}$.
(v) Show that $\Omega = \partial Q$. (Hint: Use Problem 5.2.)

Chapter 6

Ordinary differential equations in geometry, physics, and biology

6.1. Delaunay's surfaces in differential geometry

Let Σ be a surface of revolution in $\mathbb{R}^3$ so that

$$\Sigma = \{(r(t)\cos s, r(t)\sin s, t) : t \in I\}$$

for some function $r : I \to (0, \infty)$. At each point on Σ, we have two curvature radii. The reciprocals of the curvature radii are referred to as the prinicpal curvatures, and their sum is referred to as the mean curvature of Σ.

In 1841, C. Delaunay investigated surfaces of revolution which have constant mean curvature 2. The mean curvature of a surface of revolution is given by

$$H = -\frac{r(t)r''(t) - (1 + r'(t)^2)}{r(t)\,(1 + r'(t)^2)^{\frac{3}{2}}}.$$

Hence, the condition $H = 2$ is equivalent to the differential equation

$$r(t)r''(t) = (1 + r'(t)^2) - 2r(t)\,(1 + r'(t)^2)^{\frac{3}{2}}.$$

If we put $x(t) = r(t)$ and $y(t) = r'(t)$, we obtain the system

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= (1 + y(t)^2)\left(\frac{1}{x(t)} - 2\sqrt{1 + y(t)^2}\right) \end{aligned}$$

in the half-plane $U = \{(x, y) \in \mathbb{R}^2 : x > 0\}$. In order to analyze this system, we consider the function

$$L(x, y) = x^2 - \frac{x}{\sqrt{1 + y^2}}.$$

A straightforward calculation yields

$$\frac{\partial L}{\partial x}(x, y) = 2x - \frac{1}{\sqrt{1 + y^2}}$$

and

$$\frac{\partial L}{\partial y}(x, y) = \frac{xy}{\sqrt{1 + y^2}\,(1 + y^2)}.$$

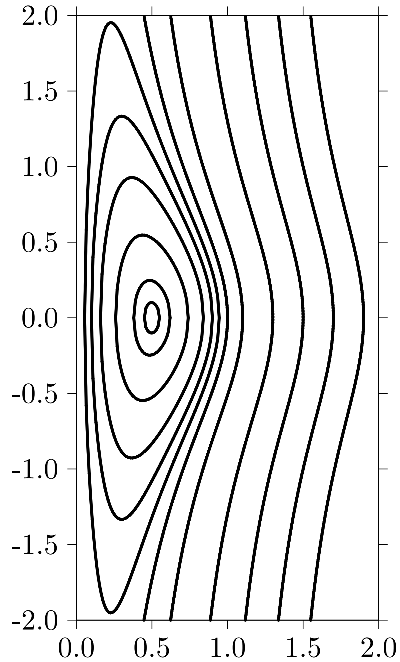
This implies

$$\begin{aligned}
&y\,\frac{\partial L}{\partial x}(x, y) + (1 + y^2)\left(\frac{1}{x} - 2\sqrt{1 + y^2}\right)\frac{\partial L}{\partial y}(x, y) \\
&= 2xy - \frac{y}{\sqrt{1 + y^2}} + \left(\frac{1}{x} - 2\sqrt{1 + y^2}\right)\frac{xy}{\sqrt{1 + y^2}} \\
&= 0.
\end{aligned}$$

Hence, if $(x(t), y(t))$ is a solution

$$\begin{aligned}
x'(t) &= y(t) \\
y'(t) &= (1 + y(t)^2)\left(\frac{1}{x(t)} - 2\sqrt{1 + y(t)^2}\right),
\end{aligned}$$

then $\frac{d}{dt}L(x(t), y(t)) = 0$. In other words, every solution curve is contained in a level curve of the function $L(x, y)$. The level curves are shown in the figure below:



The system above has only one equilibrium point, which is $(\frac{1}{2}, 0)$. The coefficient matrix of the linearized system at $(\frac{1}{2}, 0)$ is given by

$$\begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix},$$

which has eigenvalues $2i$ and $-2i$. Nonetheless, the point $(\frac{1}{2}, 0)$ is a stable equilibrium point. To see this, we apply Lyapunov's theorem. We first observe that the gradient of L at the point $(\frac{1}{2}, 0)$ vanishes, and the Hessian of L at the point $(\frac{1}{2}, 0)$ is given by

$$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Since this matrix is positive definite, we conclude that L has a strict local minimum at $(\frac{1}{2}, 0)$. Therefore, the point $(\frac{1}{2}, 0)$ is stable by Lyapunov's theorem. However, the point $(\frac{1}{2}, 0)$ is clearly not asymptotically stable, since every trajectory is contained in a level curve of L. Note that the surface of revolution associated with the constant solution $(\frac{1}{2}, 0)$ is a cylinder of radius $\frac{1}{2}$.

We next fix a real number $\frac{1}{2} < a < 1$, and let $(x(t), y(t))$ be the unique maximal solution satisfying the initial condition $(x(0), y(0)) = (a, 0)$. The function $(x(t), y(t))$ is defined on some interval J (which may be infinite).

Since L is constant along any solution curve, we must have

$$L(x(t), y(t)) = L(a, 0) = -a\,(1-a)$$

for all $t \in J$. This implies

$$-x(t)\,(1-x(t)) \leq x(t)^2 - \frac{x(t)}{\sqrt{1+y(t)^2}} = -a(1-a)$$

for all $t \in J$. From this it follows that

$$1 - a \leq x(t) \leq a$$

for all $t \in J$. Moreover,

$$\begin{aligned} \sqrt{1+y(t)^2} &= \frac{x(t)}{a\,(1-a) + x(t)^2} \\ &= \frac{1}{2\sqrt{a(1-a)}} \left(1 - \frac{(\sqrt{a(1-a)} - x(t))^2}{a(1-a) + x(t)^2}\right) \\ &\leq \frac{1}{2\sqrt{a(1-a)}}, \end{aligned}$$

hence

$$y(t)^2 \leq \frac{1}{4a(1-a)} - 1 = \frac{(2a-1)^2}{4a(1-a)}.$$

From this, we deduce that the solution is defined for all times, i.e. $J = \mathbb{R}$. (Otherwise, we could find a sequence of times $t_k \in J$ such that $\lim_{k\to\infty} x(t_k) = 0$ or $\lim_{k\to\infty} x(t_k)^2 + y(t_k)^2 = \infty$. That would contradict the previous inequalities.)

In the next step, we claim that the function $(x(t), y(t))$ is periodic. To prove this, let

$$T = \inf\{t > 0 : y(t) = 0\}.$$

(Note that T might be infinite. We will later show that this is not the case.) Note that $y(t) < 0$ for $t \in (0, T)$. Using the identity $L(x(t), y(t)) = -a\,(1-a)$, we obtain

$$y(t)^2 = \frac{x(t)^2}{(a(1-a) + x(t)^2)^2} - 1 = \frac{(a^2 - x(t)^2)\,(x(t)^2 - (1-a)^2)}{(a(1-a) + x(t)^2)^2}.$$

Therefore, $1 - a < x(t) < a$ and

$$y(t) = -\frac{\sqrt{a^2 - x(t)^2}\,\sqrt{x(t)^2 - (1-a)^2}}{a(1-a) + x(t)^2}$$

for all $t \in (0, T)$. This gives

$$\begin{aligned}
\tau &= \int_0^\tau \frac{x'(t)}{y(t)}\,dt \\
&= -\int_0^\tau \frac{a(1-a)+x(t)^2}{\sqrt{a^2-x(t)^2}\,\sqrt{x(t)^2-(1-a)^2}}\,x'(t)\,dt \\
&= \int_{x(\tau)}^a \frac{a(1-a)+x^2}{\sqrt{a^2-x^2}\,\sqrt{x^2-(1-a)^2}}\,dx \\
&\leq \int_{1-a}^a \frac{a(1-a)+x^2}{\sqrt{a^2-x^2}\,\sqrt{x^2-(1-a)^2}}\,dx
\end{aligned}$$

for $\tau \in (0, T)$. Since the integral

$$\int_{1-a}^a \frac{a(1-a)+x^2}{\sqrt{a^2-x^2}\,\sqrt{x^2-(1-a)^2}}\,dx$$

is finite, it follows that T is finite. We next observe that $y(T) = 1-a$ and $x(T) = 1-a$ by definition of T. Repeating this argument, we obtain $x(2T) = a$ and $y(2T) = 0$. In particular, $x(2T) = x(0)$ and $y(2T) = y(0)$. Hence, it follows from the uniqueness theorem that $x(t+2T) = x(t)$ and $y(t+2T) = y(t)$ for all $t \in \mathbb{R}$. This shows that the function $(x(t), y(t))$ is periodic with period

$$2T = 2\int_{1-a}^a \frac{a(1-a)+x^2}{\sqrt{a^2-x^2}\,\sqrt{x^2-(1-a)^2}}\,dx.$$

Finally, we observe that the unique solution with the initial condition $(x(0), y(0)) = (1, 0)$ is given by $x(t) = \sqrt{1-t^2}$ and $y(t) = -\frac{t}{\sqrt{1-t^2}}$. The surface of revolution associated with this solution is a sphere of radius 1.

6.2. The mathematical pendulum

Let us considered an idealized pendulum. Let θ denote the angle between the pendulum and the vertical axis. If we neglect friction, the angle θ satisfies the differential equation

$$\theta''(t) = -\frac{g}{l}\sin\theta(t),$$

where l is the length of the pendulum and g denotes the acceleration due to gravity. By a suitable choice of units, we can arrange that $\frac{g}{l} = 1$, so

$$\theta''(t) = -\sin\theta(t).$$

This differential equation is equivalent to the system

$$\begin{aligned}
x'(t) &= y(t) \\
y'(t) &= -\sin x(t).
\end{aligned}$$

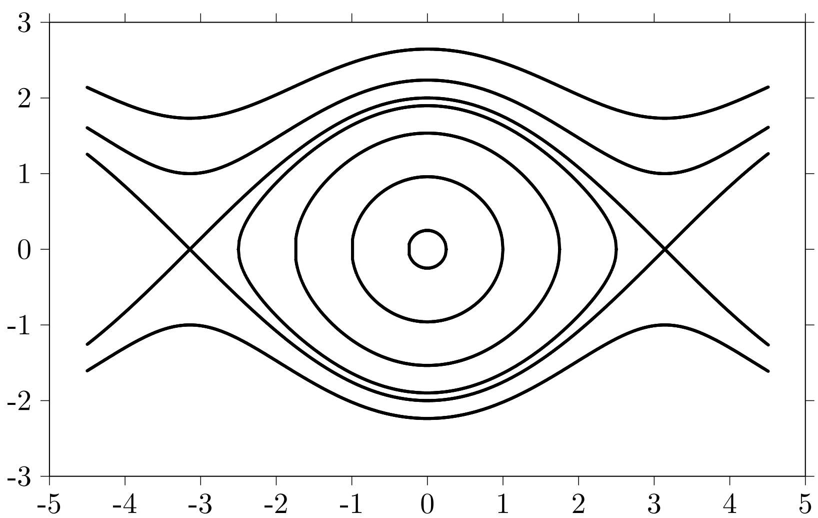
This system is Hamiltonian, and the Hamiltonian function is

$$H(x,y) = \frac{1}{2}\,y^2 - \cos x.$$

In particular,

$$\frac{1}{2}\,y(t)^2 - \cos x(t) = \text{constant}$$

for every solution $(x(t), y(t))$. This reflects the law of conservation of energy. The level curves of the function $H(x,y)$ are shown below:



The system above has infinitely many equilibrium points, which are of the form $(\pi k, 0)$ with $k \in \mathbb{Z}$. The point $(0,0)$ is a strict local minimum of the function $H(x,y)$. Consequently, the point $(0,0)$ is a stable equilibrium by Lyapunov's theorem. The point $(0,0)$ is not asymptotically stable, however.

We next consider the equilibrium point $(\pi, 0)$. The coefficient matrix of the linearized system at $(\pi, 0)$ is given by

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

This matrix has eigenvalues 1 and -1, and the associated eigenvectors are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. By the stable manifold theorem, we can find a curve W^s passing through the point $(\pi, 0)$ such that $\lim_{t\to\infty} \varphi_t(x,y) = (\pi, 0)$ for every point $(x,y) \in W^s$. Since $H(x,y)$ is constant along a solution of the ODE, we have

$$H(x,y) = \lim_{t\to\infty} H(\varphi_t(x,y)) = H(\pi, 0) = 1$$

for all points $(x, y) \in W^s$. Consequently,

$$\begin{aligned} W^s \subset & \{(x,y) : \frac{1}{2}\, y^2 - \cos x = 1\} \\ &= \{(x,y) : y = 2\cos\frac{x}{2}\} \cup \{(x,y) : y = -2\cos\frac{x}{2}\}. \end{aligned}$$

Since W^s is tangent to the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ at the point $(\pi, 0)$, we must have

$$W^s \subset \{(x,y) : y = 2\cos\frac{x}{2}\}.$$

Similarly, we can find a curve W^u passing through the point $(\pi, 0)$ such that $\lim_{t \to -\infty} \varphi_t(x, y) = (\pi, 0)$ for every point $(x, y) \in W^s$. A similar argument as above gives

$$W^u \subset \{(x,y) : y = -2\cos\frac{x}{2}\}.$$

We next fix a real number $0 < a < \pi$, and let $(x(t), y(t))$ be the unique maximal solution satisfying the initial condition $(x(0), y(0)) = (a, 0)$. The function $(x(t), y(t))$ is defined on some interval J (which may be infinite). Then

$$\frac{1}{2}\, y(t)^2 - \cos x(t) = -\cos a$$

for all $t \in J$. In particular, $|y(t)| \leq 2$ for all $t \in J$. Since $x'(t) = y(t)$, it follows that $|x(t)| \leq 2\,|t|$ for all $t \in J$. In particular, the solution $(x(t), y(t))$ cannot approach infinity in finite time. Consequently, the solution is defined for all t, i.e. $J = \mathbb{R}$.

We next show that the function $(x(t), y(t))$ is periodic. Let

$$T = \inf\{t > 0 : y(t) = 0\}.$$

(Note that T might be infinite. We will later show that this is not the case.) Clearly, $y(t) \neq 0$ and $\cos x(t) \neq \cos a$ for $t \in (0, T)$. Since the functions $x(t)$ and $y(t)$ are continuous, we have $-a < x(t) < a$ and $y(t) < 0$ for all $t \in (0, T)$. Moreover,

$$y(t) = -\sqrt{2\,(\cos x(t) - \cos a)}$$

for $t \in (0, T)$. Therefore, we obtain

$$\begin{aligned}
\tau &= \int_0^\tau \frac{1}{y(t)}\, x'(t)\, dt \\
&= -\int_0^\tau \frac{1}{\sqrt{2\,(\cos x(t) - \cos a)}}\, x'(t)\, dt \\
&= \int_{x(\tau)}^a \frac{1}{\sqrt{2\,(\cos x - \cos a)}}\, dx \\
&\leq \int_{-a}^a \frac{1}{\sqrt{2\,(\cos x - \cos a)}}\, dx
\end{aligned}$$

for all $\tau \in (0, T)$. Since the integral

$$\int_{-a}^a \frac{1}{\sqrt{2\,(\cos x - \cos a)}}\, dx$$

is finite, we conclude that $T < \infty$. By definition of T, $y(T) = 0$ and $\cos x(T) = \cos a$. Since the function $x(t)$ is decreasing for $t \in (0, T)$, we conclude that $x(T) = -a$. This gives

$$T = \int_{-a}^a \frac{1}{\sqrt{2\,(\cos x - \cos a)}}\, dx.$$

Repeating this argument, we obtain $x(2T) = a$ and $y(2T) = 0$. In particular, $x(2T) = x(0)$ and $y(2T) = y(0)$. Hence, it follows from the uniqueness theorem that $x(t+2T) = x(t)$ and $y(t+2T) = y(t)$ for all $t \in \mathbb{R}$. This shows that the function $(x(t), y(t))$ is periodic with period

$$T = 2\int_{-a}^a \frac{1}{\sqrt{2\,(\cos x - \cos a)}}\, dx.$$

6.3. Kepler's problem

Consider the following system of two coupled second order differential equations:

$$\begin{aligned}
x_1''(t) &= -\frac{x_1(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} \\
x_2''(t) &= -\frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}}.
\end{aligned}$$

This system describes the motion of a point mass in a central force field, where the force is proportional to r^{-2}.

Proposition 6.1. *Suppose that $(x_1(t), x_2(t))$ is a solution of the system of differential equations given above. Then we have the following conserved* *quantities:*

$$L = x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t) = \text{constant},$$

$$A_1 = -(x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_2'(t) + \frac{x_1(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}} = \text{constant},$$

$$A_2 = (x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_1'(t) + \frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}} = \text{constant}.$$

Proof. We compute

$$\begin{aligned}
&\frac{d}{dt}(x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t)) \\
&= x_1(t)\,x_2''(t) - x_2(t)\,x_1''(t) \\
&= -x_1(t)\,\frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} + x_2(t)\,\frac{x_1(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} \\
&= 0.
\end{aligned}$$

This proves the first statement. Since the function $x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t)$ is constant, we obtain

$$\begin{aligned}
&\frac{d}{dt}\Big(- (x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_2'(t) + \frac{x_1(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}}\Big) \\
&= -(x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_2''(t) \\
&+ \frac{x_1'(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}} - \frac{x_1(t)^2\,x_1'(t) + x_1(t)x_2(t)\,x_2'(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} \\
&= -(x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_2''(t) \\
&- (x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,\frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} \\
&= 0.
\end{aligned}$$

This proves the second statement. Finally, if we replace $(x_1(t), x_2(t))$ by $(x_2(t), x_1(t))$, we obtain

$$\frac{d}{dt}\Big((x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t))\,x_1'(t) + \frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}}\Big) = 0.$$

This completes the proof. $\square$

We note that the first identity reflects the conservation of angular momentum. These conservation laws hold for any central force field. The conservation laws for A_1 and A_2 are much more subtle, and are special to Kepler's problem. The vector (A_1, A_2) is called the Runge-Lenz vector.

Corollary 6.2. *Suppose that $(x_1(t), x_2(t))$ is a solution of the system of differential equations given above. Then*

$$x_1(t)^2 + x_2(t)^2 = (A_1 x_1(t) + A_2 x_2(t) + L^2)^2.$$

In particular, the path $t \mapsto (x_1(t), x_2(t))$ is contained in a conic section.

Proof. We compute

$$\begin{aligned} A_1 x_1(t) + A_2 x_2(t) &= -(x_1(t)\, x_2'(t) - x_2(t)\, x_1'(t))^2 + (x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}} \\ &= -L^2 + (x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}. \end{aligned}$$

This implies

$$x_1(t)^2 + x_2(t)^2 = (A_1 x_1(t) + A_2 x_2(t) + L^2)^2.$$

In other words, the path $t \mapsto (x_1(t), x_2(t))$ is contained in the set

$$\Gamma = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = (A_1 x_1 + A_2 x_2 + L^2)^2\}.$$

If $A_1^2 + A_2^2 < 1$, then Γ is an ellipse with principle axes $\frac{L^2}{\sqrt{1-A_1^2-A_2^2}}$ and $\frac{L^2}{1-A_1^2-A_2^2}$. If $A_1^2 + A_2^2 = 1$, then Γ is a parabola. Finally, if $A_1^2 + A_2^2 > 1$, then Γ is a hyperbola. □

Finally, let us derive a formula for the position $(x_1(t), x_2(t))$ as a function of t. For simplicity, we only consider the case when the orbit is an ellipse.

Proposition 6.3. *Suppose that $(x_1(t), x_2(t))$ is a solution of the system of differential equations given above. Moreover, suppose that $A_1 = \varepsilon$ and $A_2 = 0$, where $0 \leq \varepsilon < 1$ and A_1 and A_2 are defined as in Proposition 6.1. Then we may write*

$$\begin{aligned} x_1(t) &= \frac{L^2}{1-\varepsilon^2}\,(\cos\theta(t) + \varepsilon), \\ x_2(t) &= \frac{L^2}{\sqrt{1-\varepsilon^2}}\,\sin\theta(t), \end{aligned}$$

where $\theta(t)$ satisfies Kepler's equation

$$\theta(t) + \varepsilon\,\sin\theta(t) = \Big(\frac{L^2}{1-\varepsilon^2}\Big)^{-\frac{3}{2}}\, t + \text{constant}.$$

In particular, the solution $(x_1(t), x_2(t))$ is periodic with period $2\pi\,\big(\frac{L^2}{1-\varepsilon^2}\big)^{\frac{3}{2}}$.

Proof. By Corollary 6.2,

$$x_1^2 + x_2^2 = (\varepsilon\, x_1 + L^2)^2.$$

Rearranging terms gives

$$(1-\varepsilon^2)\Big(x_1 - \frac{L^2\,\varepsilon}{1-\varepsilon^2}\Big)^2 + x_2^2 = \frac{L^4}{1-\varepsilon^2}.$$

Consequently, we can find a function function $\theta(t)$ such that

$$x_1(t) = \frac{L^2}{1-\varepsilon^2}\,(\cos\theta(t) + \varepsilon)$$
$$x_2(t) = \frac{L^2}{\sqrt{1-\varepsilon^2}}\,\sin\theta(t).$$

Differentiating these identities with respect to t gives

$$\begin{aligned} L &= x_1(t)\,x_2'(t) - x_2(t)\,x_1'(t) \\ &= \frac{L^4}{(1-\varepsilon^2)^{\frac{3}{2}}}\,(\cos^2\theta(t) + \varepsilon\,\cos\theta(t))\,\theta'(t) + \frac{L^4}{(1-\varepsilon^2)^{\frac{3}{2}}}\,\sin^2\theta(t)\,\theta'(t) \\ &= \frac{L^4}{(1-\varepsilon^2)^{\frac{3}{2}}}\,(1 + \varepsilon\,\cos\theta(t))\,\theta'(t). \end{aligned}$$

Thus, we conclude that

$$(1 + \varepsilon\,\cos\theta(t))\,\theta'(t) = \Big(\frac{L^2}{1-\varepsilon^2}\Big)^{-\frac{3}{2}}.$$

Integrating this equation with respect to t yields

$$\theta(t) + \varepsilon\,\sin\theta(t) = \Big(\frac{L^2}{1-\varepsilon^2}\Big)^{-\frac{3}{2}}\,t + \text{constant}.$$

Hence, if we define $T = 2\pi\,\big(\frac{L^2}{1-\varepsilon^2}\big)^{\frac{3}{2}}$, then $\theta(t+T) = \theta(t) + 2\pi$. This implies $(x_1(t+T), x_2(t+T)) = (x_1(t), x_2(t))$. Thus, the solution $(x_1(t), x_2(t))$ is periodic with period $T = 2\pi\,\big(\frac{L^2}{1-\varepsilon^2}\big)^{\frac{3}{2}}$. □

6.4. Predator-prey models

In this section, we analyze a model for the growth of the populations of two species, one of which preys on the other. This model was proposed by Volterra and Lotka in the 1920s. Let $x(t)$ denote the size of the prey population at time t, and let $y(t)$ denote the size of the predator population at time t. The dynamics of $x(t)$ and $y(t)$ is modeled by the differential equations

$$x'(t) = a\,x(t) - \alpha\,x(t)\,y(t)$$
$$y'(t) = -c\,y(t) + \gamma\,x(t)\,y(t),$$

where a, α, c, γ are positive constants.

This system has two equilibrium points, $(0,0)$ and $(\frac{c}{\gamma}, \frac{a}{\alpha})$. The coefficient matrix of the linearized system at the point $(0,0)$ is given by

$$\begin{bmatrix} a & 0 \\ 0 & -c \end{bmatrix}.$$

Therefore, the point $(0,0)$ is a saddle point. The stable curve is given by $\{x = 0\}$, and the unstable curve is $\{y = 0\}$.

Similarly, the coefficient matrix of the linearized system at the point $(\frac{c}{\gamma}, \frac{a}{\alpha})$ is given by

$$\begin{bmatrix} 0 & \frac{a\gamma}{\alpha} \\ -\frac{\alpha c}{\gamma} & 0 \end{bmatrix}.$$

The eigenvalues of this matrix are $\sqrt{ac}\, i$ and $-\sqrt{ac}\, i$, so we need additional arguments to decide whether the equilibrium point $(\frac{c}{\gamma}, \frac{a}{\alpha})$ is stable.

We next consider the function

$$L(x,y) = \gamma\, x - c\, \log x + \alpha\, y - a \log y$$

for $x, y > 0$. We claim that this function is a Lyapunov function. It is clear that the function $x \mapsto \gamma\, x - c\, \log x$ attains its global minimum at the point $\frac{c}{\gamma}$. Similarly, the function $y \mapsto \alpha\, y - a\, \log y$ attains its global minimum at the point $\frac{a}{\alpha}$. Therefore, the function L attains its global minimum at the point $(\frac{c}{\gamma}, \frac{a}{\alpha})$. Moreover, this is a strict minimum.

Suppose now that $(x(t), y(t))$ is a solution of the system of differential equations considered above satisfying $x(t), y(t) > 0$. Then

$$\begin{aligned}
&\frac{d}{dt} L(x(t), y(t)) \\
&= \Big(\gamma - \frac{c}{x(t)}\Big)\, x'(t) + \Big(\alpha - \frac{a}{y(t)}\Big)\, y'(t) \\
&= \Big(\gamma - \frac{c}{x(t)}\Big)\, (a\, x(t) - \alpha\, x(t)\, y(t)) + \Big(\alpha - \frac{a}{y(t)}\Big)\, (-c\, y(t) + \gamma\, x(t)\, y(t)) \\
&= (\gamma\, x(t) - c)\, (a - \alpha\, y(t)) + (\alpha\, y(t) - a)\, (-c + \gamma\, x(t)) \\
&= 0.
\end{aligned}$$

Therefore, the function $L(x,y)$ is a conserved quantity. By Lyapunov's theorem, the equilibrium point $(\frac{c}{\gamma}, \frac{a}{\alpha})$ is stable.

Finally, it is not difficult to show that each level set of the function L is smooth curve which lies in a compact subset of the quadrant $\{(x,y) \in \mathbb{R}^2 : x, y > 0\}$. Hence, every solution that originates in this quadrant is either constant or periodic.

6.5. Mathematical models for the spread of infectious diseases

In this section, we discuss a model for the spread of infectious diseases which is due to Kermack and McKendrick. This model relies on several assumptions. First, we assume that the incubation period can be neglected, so that any infected person can immediately infect others. Second, we assume that any person who has recovered from the disease gains permanent immunity. Let us denote by $x(t)$ the number of persons who are susceptible to the disease. Moreover, let $y(t)$ be the number of persons who are currently infected. In other words, $x(t)$ is the number of persons who have not contracted the disease prior to time t, and $y(t)$ is the number of persons who have been infected but have not yet recovered. Finally, let $z(t)$ denote the number of persons who have recovered from the disease.

The dynamics of $x(t), y(t), z(t)$ can be modeled by the following system of differential equations:

$$\begin{aligned} x'(t) &= -\beta\, x(t)\, y(t) \\ y'(t) &= \beta\, x(t)\, y(t) - \gamma\, y(t) \\ z'(t) &= \gamma\, y(t). \end{aligned}$$

Here, β and γ are positive constants. This system is an example of what is called an SIR-model, and was first proposed by Kermack and McKendrick.

We note that the sum $x(t) + y(t) + z(t)$ is constant. Hence, it is enough to solve the two-dimensional system

$$\begin{aligned} x'(t) &= -\beta\, x(t)\, y(t) \\ y'(t) &= \beta\, x(t)\, y(t) - \gamma\, y(t). \end{aligned}$$

To that end, we consider the quantity

$$L(x, y) = x(t) - \frac{\gamma}{\beta} \log x(t) + y(t).$$

A straightforward calculation gives

$$\begin{aligned} \frac{d}{dt} L(x(t), y(t)) &= \Big(1 - \frac{\gamma}{\beta\, x(t)}\Big)\, x'(t) + y'(t) \\ &= -(\beta\, x(t) - \gamma)\, y(t) + \beta\, x(t)\, y(t) - \gamma\, y(t) \\ &= 0. \end{aligned}$$

Thus, $L(x, y)$ is a conserved quantity.

If $x(0) \leq \frac{\gamma}{\beta}$, then the function $x(t)$ is monotone decreasing and converges to 0 as $t \to \infty$. On the other hand, if $x(0) > \frac{\gamma}{\beta}$, then the number of infected persons will increase at first. The epidemic will reach its peak when $x(t) = \frac{\gamma}{\beta}$. Afterwards, the number of infected persons will decrease, and will converge to 0 as $t \to \infty$. For that reason, the ratio $\frac{\gamma}{\beta}$ is referred to as the epidemiological threshold.

6.6. A mathematical model of glycolysis

In this section, we discuss Sel'kov's model for glycolysis. Our treatment closely follows [**5**], Section 7.3. The differential equations governing Sel'kov's model are as follows:

$$\begin{aligned} x'(t) &= -x(t) + a\,y(t) + x(t)^2\,y(t) \\ y'(t) &= b - a\,y(t) - x(t)^2\,y(t). \end{aligned}$$

Here, the functions $x(t)$ and $y(t)$ describe the concentrations, at time t, of two chemicals (adenosine diphosphate and fructose-6-phosphate), and a and b are positive constants. We will focus on the case when the initial values $x(0) = x_0$ and $y(0) = y_0$ are positive.

Proposition 6.4. *Given any real number $\lambda \geq \frac{b}{a}$, the set*

$$A = \{(x,y) \in \mathbb{R}^2 : x \leq 0,\ 0 \leq y \leq \lambda,\ \textit{and}\ x + y \leq \lambda + b\}$$

is positively invariant.

Proof. It suffices to show that the vector field

$$F(x,y) = (-x + ay + x^2 y, b - ay - x^2 y)$$

is inward-pointing along the boundary of A. In other words, we need to show that

$$\langle F(x,y), \nu \rangle \leq 0$$

for every point $(x,y) \in \partial A$, where ν denotes the outward-pointing unit normal vector to ∂A.

The boundary of A consists of four line segments:

$$\partial A = L_1 \cup L_2 \cup L_3 \cup L_4,$$

where

$$\begin{aligned} L_1 &= \{(x,y) \in \mathbb{R}^2 : x = 0 \text{ and } 0 \leq y \leq \lambda\} \\ L_2 &= \{(x,y) \in \mathbb{R}^2 : y = 0 \text{ and } 0 \leq x \leq \lambda + b\} \\ L_3 &= \{(x,y) \in \mathbb{R}^2 : x + y = \lambda + b \text{ and } b \leq x \leq \lambda + b\} \\ L_4 &= \{(x,y) \in \mathbb{R}^2 : y = \lambda \text{ and } 0 \leq x \leq b\}. \end{aligned}$$

Step 1: Consider a point $(x,y) \in L_1$. The outward-pointing unit normal vector to ∂A at (x,y) is $\nu = (-1, 0)$. This implies

$$\langle F(x,y), \nu \rangle = \langle F(0,y), \nu \rangle = \langle (ay, b - ay), (-1, 0) \rangle = -ay \leq 0$$

since $y \geq 0$.

Step 2: We next consider a point $(x,y) \in L_2$. In this case, the outward-pointing unit-normal vector is given by $\nu = (0, -1)$. Therefore,

$$\langle F(x,y), \nu \rangle = \langle F(x,0), \nu \rangle = \langle (-x, b), (0, -1) \rangle = -b \leq 0.$$

Step 3: Consider now a point $(x, y) \in L_3$. In this case, the outward-pointing unit-normal vector is $\nu = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Hence,

$$\begin{aligned}
&\langle F(x,y), \nu\rangle \\
&= \Big\langle (-x(t) + a\,y(t) + x(t)^2\,y(t), b - a\,y(t) - x(t)^2\,y(t)), \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\Big)\Big\rangle \\
&= \frac{1}{\sqrt{2}}\,(b - x) \leq 0
\end{aligned}$$

since $x \geq b$.

Step 4: Finally, let (x, y) be a point on the line segment L_4. The outward-pointing unit normal vector to ∂A is $\nu = (0, 1)$. From this it follows that

$$\begin{aligned}
\langle F(x,y), \nu\rangle &= \langle F(x,\lambda), \nu\rangle \\
&= \langle (-x + a\lambda + x^2\lambda, b - a\lambda - x^2\lambda), (0,1)\rangle \\
&= b - a\lambda - x^2\lambda \\
&\leq b - a\lambda \leq 0
\end{aligned}$$

since $\lambda \geq \frac{b}{a}$.

Therefore, the region A is positively invariant as claimed. $\square$

We next discuss the asymptotic behavior of the solution as $t \to \infty$.

Proposition 6.5. *Let (x_0, y_0) be a pair of positive real numbers. Moreover, let $(x(t), y(t))$ be the unique maximal solution of the system above with $x(0) = x_0$ and $y(0) = y_0$, and let Ω be the set of all ω-limit points of $(x(t), y(t))$. Then Ω is bounded and non-empty.*

Proof. Consider the region

$$A = \{(x, y) \in \mathbb{R}^2 : x \geq 0,\ 0 \leq y \leq \lambda, \text{ and } x + y \leq \lambda + b\},$$

where $\lambda = \max\{x_0 + y_0 - b, y_0, \frac{b}{a}\}$. It follows from Proposition 6.4 that $(x(t), y(t)) \in A$ for all $t \geq 0$. Since A is a bounded region, it follows from the Bolzano-Weierstrass theorem that the solution $(x(t), y(t))$ has at least one ω-limit point. This shows that Ω is non-empty. On the other hand, since $(x(t), y(t)) \in A$ for all $t \geq 0$, it follows that $\Omega \subset A$, which implies that Ω is bounded. $\square$

Finally, we analyze the equilibrium points:

Proposition 6.6. *The point $(b, \frac{b}{a+b^2})$ is the only equilibrium point of the system. If $b^4 + (2a-1)\,b^2 + (a^2 + a) < 0$, then both eigenvalues of the matrix $DF(b, \frac{b}{a+b^2})$ have positive real part. Similarly, if $b^4 + (2a-1)\,b^2 + (a^2 + a) > 0$, then both eigenvalues of the matrix $DF(b, \frac{b}{a+b^2})$ have negative real part.*

Proof. Suppose that $(\bar{x}, \bar{y})$ is an equilibrium point. Then

$$-\bar{x} + a\,\bar{y} + \bar{x}^2\,\bar{y} = 0$$

and

$$b - a\,\bar{y} - \bar{x}^2\,\bar{y} = 0.$$

Adding both identities, we obtain

$$b - \bar{x} = 0,$$

hence $\bar{x} = b$. Substituting this into the first equation, we obtain

$$(a + b^2)\,\bar{y} = b,$$

hence

$$\bar{y} = \frac{b}{a + b^2}.$$

The differential of F is given by

$$DF(x, y) = \begin{bmatrix} -1 + 2xy & a + x^2 \\ -2xy & -a - x^2 \end{bmatrix}.$$

The trace and determinant of this matrix are

$$\operatorname{tr} \begin{bmatrix} -1 + 2xy & a + x^2 \\ -2xy & -a - x^2 \end{bmatrix} = -1 + 2xy - x^2 - a$$

and

$$\det \begin{bmatrix} -1 + 2xy & a + x^2 \\ -2xy & -a - x^2 \end{bmatrix} = a + x^2.$$

In particular, the determinant of $DF(x, y)$ is always positive. This shows that the matrix $DF(b, \frac{b}{a+b^2})$ cannot have two real eigenvalues with opposite signs. The sign of the trace of $DF(b, \frac{b}{a+b^2})$ indicates whether the eigenvalues of the matrix $DF(b, \frac{b}{a+b^2})$ have positive or negative real part. □

Theorem 6.7. *Let (x_0, y_0) be a pair of positive real numbers such that $(x_0, y_0) \neq (b, \frac{b}{a+b^2})$. Moreover, let $(x(t), y(t))$ be the unique maximal solution with $x(0) = x_0$ and $y(0) = y_0$. Then the following holds:*

(i) *If $b^4 + (2a - 1)\,b^2 + (a^2 + a) < 0$, then the solution $(x(t), y(t))$ approaches a periodic solution as $t \to \infty$.*

(ii) *If $b^4 + (2a - 1)\,b^2 + (a^2 + a) > 0$, then either $\lim_{t\to\infty}(x(t), y(t)) = (b, \frac{b}{a+b^2})$ or the solution $(x(t), y(t))$ converges to a periodic solution as $t \to \infty$.*

Proof. As above, let Ω be the set of all ω-limit points of $(x(t), y(t))$. We first assume that $b^4 + (2a - 1)\,b^2 + (a^2 + a) < 0$. In this case, both eigenvalues of the matrix $DF(b, \frac{b}{a+b^2})$ have positive real part. By Theorem 4.2, the equilibrium point $(b, \frac{b}{a+b^2})$ is asymptotically stable after a time reversal $t \to -t$. Using Proposition 5.9, we conclude that $(b, \frac{b}{a+b^2}) \notin \Omega$. In particular, the set Ω contains no equilibrium points. Hence, the Poincaré-Bendixson theorem implies that the solution $(x(t), y(t))$ approaches a periodic solution as $t \to \infty$.

We now assume that $b^4 + (2a - 1)\,b^2 + (a^2 + a) > 0$. In this case, both eigenvalues of the matrix $DF(b, \frac{b}{a+b^2})$ have negative real part. By Theorem 4.2, the equilibrium point $(b, \frac{b}{a+b^2})$ is asymptotically stable. If $(b, \frac{b}{a+b^2}) \notin \Omega$, it follows from the Poincaré-Bendixson theorem that the solution $(x(t), y(t))$ converges to a periodic solution as $t \to \infty$. On the other hand, if $(b, \frac{b}{a+b^2}) \in \Omega$, then we have $\lim_{t\to\infty}(x(t), y(t)) = (b, \frac{b}{a+b^2})$ since $(b, \frac{b}{a+b^2})$ is asymptotically stable. $\square$

6.7. Problems

Problem 6.1. Let $(x(t), y(t))$ be the unique solution of the differential equations

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -\sin x(t) \end{aligned}$$

with initial values $x(0) = 0$ and $y(0) = 2$.
(i) Show that

$$x'(t) = 2\,\cos\frac{x(t)}{2}.$$

(ii) Using Problem 1.5, conclude that

$$\frac{x(t)}{2} = \arctan(e^t) - \arctan(e^{-t})$$

and

$$y(t) = \frac{4}{e^t + e^{-t}}.$$

What can you say about the asymptotic behavior of $(x(t), y(t))$ as $t \to \infty$?

Problem 6.2. Let $(x_1(t), x_2(t))$ be a solution of the system

$$\begin{aligned} x_1''(t) &= -\frac{x_1(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}} \\ x_2''(t) &= -\frac{x_2(t)}{(x_1(t)^2 + x_2(t)^2)^{\frac{3}{2}}}. \end{aligned}$$

(i) Show that

$$E = \frac{1}{2}\,(x_1'(t)^2 + x_2'(t)^2) - \frac{1}{(x_1(t)^2 + x_2(t)^2)^{\frac{1}{2}}} = \text{constant}.$$

(ii) Show that $A_1^2 + A_2^2 = 2EL^2 + 1$, where L, A_1, A_2 denote the conserved quantities from Proposition 6.1.

Chapter 7

Sturm-Liouville theory

7.1. Boundary value problems for linear differential equations of second order

In this chapter, we will study linear differential equations of second order with variable coefficients.

Definition 7.1. Let $p(t)$ and $q(t)$ be two continuously differentiable functions that are defined on some interval $[a, b]$. We then consider the following linear differential equation of second order for the unknown function $u(t)$:

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + q(t)\,u(t) = 0. \tag{9}$$

The differential equation (9) is called a Sturm-Liouville equation. Moreover, we say that (9) is a regular Sturm-Liouville equation if $p(t) > 0$ for all $t \in [a, b]$.

In the following, we will only consider regular Sturm-Liouville equations.

Proposition 7.2. *Suppose that (9) is a regular Sturm-Liouville equation. Then the set of all functions $u(t)$ satisfying (9) is a vector space of dimension 2. In other words, there exist two linearly independent solutions $w_0(t)$ and $w_1(t)$ of (9), and any other solution of (9) can be written as a linear combination of $w_0(t)$ and $w_1(t)$ with constant coefficients.*

Proof. The differential equation (9) is equivalent to the non-autonomous linear system

$$\begin{aligned} x_1'(t) &= \frac{1}{p(t)}\,x_2(t) \\ x_2'(t) &= -q(t)\,x_1(t). \end{aligned}$$

By Proposition 3.10, there exists a unique solution of (9) which satisfies the initial condition $u(a) = 1$ and $u'(a) = 0$. Similarly, there exists a unique solution of (9) which satisfies the initial conditions $u(a) = 0$ and $u'(a) = 1$. Let us denote these solutions by $w_0(t)$ and $w_1(t)$. It follows from Proposition 3.10 that the functions $w_0(t)$ and $w_1(t)$ are defined on entire interval $[a, b]$.

It is clear that $w_0(t)$ and $w_1(t)$ are linearly indepedent (i.e. neither function is a constant multiple of the other). It remains to show that any solution of (9) can be written as a linear combination of $w_0(t)$ and $w_1(t)$. To see this, let $u(t)$ be an arbitrary solution of (9). Then the function $v(t) = u(t) - u(a)\, w_0(t) - u'(a)\, w_1(t)$ is a solution of (9), and we have $v(a) = u(a) - u(a)\, w_0(a) - u'(a)\, w_1(a) = 0$ and $v'(a) = u'(a) - u(a)\, w_0'(a) - u'(a)\, w_1'(a) = 0$. Hence, the uniqueness theorem implies that $v(t) = 0$ for all $t \in [a, b]$. Thus, we conclude that $u(t) = u(a)\, w_0(t) + u'(a)\, w_1(t)$ for all $t \in [a, b]$. This shows that $u(t)$ can be written as a linear combination of $w_0(t)$ and $w_1(t)$ with constant coefficients. □

So far, we have focused on initial value problems for systems of ordinary differential equations. However, in many situations, it is more natural to impose boundary conditions instead. For example, we can prescribe the values of the function u at the endpoints of the interval $[a, b]$. Alternatively, we may prescribe the values of $u'(t)$ at the endpoints of the interval $[a, b]$. The following list shows the most common boundary conditions:

- Dirichlet boundary conditions: $u(a) = A$, $u(b) = B$.
- Neumann boundary conditions: $u'(a) = A$, $u'(b) = B$.
- Mixed Dirichlet-Neumann boundary condition: $u(a) = A$, $u'(b) = B$.
- Periodic boundary condition: $u(a) = u(b)$, $u'(a) = u'(b)$.

Here, A and B are given real numbers.

A Sturm-Liouville equation together with a set of boundary conditions is called a Sturm-Liouville system. While the initial value problem for a regular Sturm-Liouville system always has exactly one solution, a boundary value problem may have infinitely many solutions or no solutions at all. To see this, we first consider the differential equation

$$\frac{d^2}{dt^2} u(t) + u(t) = 0$$

with the Dirichlet boundary conditions $u(0) = 0$ and $u(\pi) = 0$. This equation has infinitely many solutions. In fact, every constant multiple of the function $\sin t$ is a solution.

We next consider the differential equation

$$\frac{d^2}{dt^2} u(t) + u(t) = 0$$

with the Dirichlet boundary conditions $u(0) = 0$ and $u(\pi) = 1$. This equation has no solution. In fact, if $u(t)$ is a solution of $\frac{d^2}{dt^2} u(t) + u(t) = 0$ with $u(0) = 0$, then $u(t)$ must be a constant multiple of $\sin t$. But then $u(\pi) = 0$ since $\sin \pi = 0$. Therefore, there is no solution of the differential equation $\frac{d^2}{dt^2} u(t) + u(t) = 0$ with $u(0) = 0$ and $u(\pi) = 1$.

In the sequel, we will mostly focus on the Dirichlet boundary value problem. The following result gives a necessary and sufficient condition for the Dirichlet boundary value problem to admit a unique solution:

Proposition 7.3. *Suppose that (9) is a regular Sturm-Liouville equation. Then the following statements are equivalent:*

- (i) *Given any pair of real numbers (A, B), there exists a unique solution of (9) such that $u(a) = A$ and $u(b) = B$.*
- (ii) *Every solution of (9) satisfying $u(a) = u(b) = 0$ vanishes identically.*

Proof. It is clear that (i) implies (ii). To prove the reverse implication, we consider the set V of all solutions of the differential equation (9). By Proposition 7.2, V is a vector space of dimension 2. We next consider the linear transformation $L : V \to \mathbb{R}^2$, which assigns to every function $u \in V$ the pair $(u(a), u(b)) \in \mathbb{R}^2$. If condition (ii) holds, then L has trivial nullspace. Since V has the same dimension as $\mathbb{R}^2$, it follows that L is invertible. This implies that statement (i) holds. $\square$

Thus, in order to decide whether the Dirichlet boundary value problem has a unique solution, it suffices to study solutions of the differential equation (9) which satisfy the boundary condition $u(a) = u(b) = 0$.

We now develop an important method for studying the solutions of a regular Sturm-Liouville equation

$$\frac{d}{dt}\Big[p(t) \, \frac{d}{dt} u(t)\Big] + q(t) \, u(t) = 0.$$

Suppose that $u(t)$ is a solution of (9) which is not identically zero. Our strategy is to rewrite (9) as a system of two differential equations of first order, and pass to polar coordinates:

$$\begin{aligned} &p(t) \, \frac{d}{dt} u(t) = r(t) \, \cos\theta(t) \\ &u(t) = r(t) \, \sin\theta(t). \end{aligned}$$

This change of variables is known as the Prüfer substitution.

Clearly,

$$\begin{aligned} 0 &= \frac{d}{dt}(r(t)\,\sin\theta(t)) - \frac{1}{p(t)}\,r(t)\,\cos\theta(t) \\ &= r'(t)\,\sin\theta(t) + r(t)\,\theta'(t)\,\cos\theta(t) - \frac{1}{p(t)}\,r(t)\cos\theta(t). \end{aligned}$$

Moreover, since $u(t)$ is a solution of (9), we have

$$\begin{aligned} 0 &= \frac{d}{dt}(r(t)\,\cos\theta(t)) + q(t)\,r(t)\,\sin\theta(t) \\ &= r'(t)\,\cos\theta(t) - r(t)\,\theta'(t)\,\sin\theta(t) + q(t)\,r(t)\,\sin\theta(t). \end{aligned}$$

This gives a system of two linear equations for the two unknowns $r'(t)$ and $\theta'(t)$. Solving this system yields

$$r'(t) = \Big(\frac{1}{p(t)} - q(t)\Big)\,r(t)\,\cos\theta(t)\,\sin\theta(t) \tag{10}$$

and

$$\theta'(t) = q(t)\,\sin^2\theta(t) + \frac{1}{p(t)}\,\cos^2\theta(t). \tag{11}$$

This is a system of two nonlinear differential equations for $r(t)$ and $\theta(t)$. A key observation is that the differential equation for the function $\theta(t)$ does not make any reference to the function $r(t)$. Hence, we can first look for a solution $\theta(t)$ of (11). Once $\theta(t)$ is known, the function $r(t)$ is determined by the formula

$$r(t) = r(a)\,\exp\bigg(\int_a^t \Big(\frac{1}{p(s)} - q(s)\Big)\,\cos\theta(s)\,\sin\theta(s)\,ds\bigg).$$

Moreover, if $u(t)$ satisfies a boundary condition, we obtain additional restrictions on $r(t)$ and $\theta(t)$. For example, if u satisfies the Dirichlet boundary conditions $u(a) = u(b) = 0$, then $\sin\theta(a) = \sin\theta(b) = 0$. Similarly, the Neumann boundary condition $u'(a) = u'(b) = 0$ is equivalent to $\cos\theta(a) = \cos\theta(b) = 0$. Moreover, the mixed Dirichlet-Neumann boundary $u(a) = u'(b) = 0$ leads to the equation $\sin\theta(a) = \cos\theta(b) = 0$. In all these cases, the boundary condition for $u(t)$ leads to a set of endpoint conditions for the function $\theta(t)$. This is true for many boundary conditions, with the notable exception of periodic boundary conditions.

7.2. The Sturm comparison theorem

In this section, we describe a method for comparing the phase functions of two Sturm-Liouville systems. To that end, we need the following comparison principle:

Proposition 7.4. *Suppose that $\theta(t)$ and $\tilde{\theta}(t)$ satisfy the differential inequalities*

$$\theta'(t) \geq q(t)\,\sin^2\theta(t) + \frac{1}{p(t)}\,\cos^2\theta(t)$$

and

$$\tilde{\theta}'(t) \leq q(t)\,\sin^2\tilde{\theta}(t) + \frac{1}{p(t)}\,\cos^2\tilde{\theta}(t),$$

where $p(t)$ is a positive function. If $\theta(a) \geq \tilde{\theta}(a)$, then $\theta(t) \geq \tilde{\theta}(t)$ for all $t \in [a,b]$. Moreover, if $\theta(b) = \tilde{\theta}(b)$, then $\theta(t) = \tilde{\theta}(t)$ for all $t \in [a,b]$.

Proof. We can find a positive constant $L > 0$ such that

$$\begin{aligned}\frac{d}{dt}(\theta(t) - \tilde{\theta}(t)) &\geq q(t)\,(\sin^2\theta(t) - \sin^2\tilde{\theta}(t)) + \frac{1}{p(t)}\,(\cos^2\theta(t) - \cos^2\tilde{\theta}(t)) \\ &\geq -L\,|\theta(t) - \tilde{\theta}(t)|\end{aligned}$$

for all $t \in [a,b]$.

Suppose now that there exists a real number $t_1 \in [a,b]$ such that $\theta(t_1) < \tilde{\theta}(t_1)$. By assumption, $t_1 \in (a,b]$. We now define

$$t_0 = \sup\{t \in [a,t_1) : \theta(t) \geq \tilde{\theta}(t)\}.$$

It is easy to see that $\theta(t_0) = \tilde{\theta}(t_0)$ and $\theta(t) < \tilde{\theta}(t)$ for all $t \in (t_0,t_1]$. This implies

$$\frac{d}{dt}(\theta(t) - \tilde{\theta}(t)) \geq L\,(\theta(t) - \tilde{\theta}(t))$$

for all $t \in (t_0,t_1]$. Consequently, the function $e^{-Lt}\,(\theta(t) - \tilde{\theta}(t))$ is monotone increasing on the interval $[t_0,t_1]$. Since $\theta(t_0) - \tilde{\theta}(t_0) = 0$, we conclude that $\theta(t) - \tilde{\theta}(t) \geq 0$ for all $t \in [t_0,t_1]$. This contradicts our choice of t_1. Thus, we conclude that $\theta(t) \geq \tilde{\theta}(t)$ for all $t \in [a,b]$.

Consequently,

$$\frac{d}{dt}(\theta(t) - \tilde{\theta}(t)) \geq -L\,(\theta(t) - \tilde{\theta}(t))$$

for all $t \in [a,b]$. Therefore, the function $e^{Lt}\,(\theta(t) - \tilde{\theta}(t))$ is nonnegative and monotone increasing on the interval $[a,b]$. Hence, if $\theta(b) - \tilde{\theta}(b) = 0$, then $\theta(t) - \tilde{\theta}(t) = 0$ for all $t \in [a,b]$. This completes the proof. □

Corollary 7.5. *Suppose that $\theta(t)$ satisfies the differential inequality*

$$\theta'(t) \geq q(t)\,\sin^2\theta(t) + \frac{1}{p(t)}\,\cos^2\theta(t),$$

where $p(t)$ is a positive function. If $\theta(a) \geq \pi k$, then $\theta(b) > \pi k$.

Proof. Let us define $\tilde{\theta}(t) = \pi k$. Since $\tilde{\theta}(t)$ is constant, we have

$$\tilde{\theta}'(t) < \tilde{q}(t) \sin^2 \tilde{\theta}(t) + \frac{1}{\tilde{p}(t)} \cos^2 \tilde{\theta}(t).$$

Hence, Proposition 7.4 implies that $\theta(b) \geq \pi k$. Moreover, the inequality is be strict. In fact, if $\theta(b) = \pi k$, then Proposition 7.4 implies that $\theta(t) = \pi k$ for all $t \in [a, b]$, which is impossible. $\square$

Corollary 7.6. *Suppose that $0 < p(t) \leq \tilde{p}(t)$ and $q(t) \geq \tilde{q}(t)$ for all $t \in [a, b]$. Moreover, suppose that $u(t)$ and $\tilde{u}(t)$ are non-trivial solutions of the Sturm-Liouville equations*

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + q(t)\,u(t) = 0$$

and

$$\frac{d}{dt}\Big[\tilde{p}(t)\,\frac{d}{dt}\tilde{u}(t)\Big] + \tilde{q}(t)\,\tilde{u}(t) = 0.$$

Finally, suppose that $\theta(t)$ and $\tilde{\theta}(t)$ denote the associated phase functions. If $\theta(a) \geq \tilde{\theta}(a)$, then $\theta(t) \geq \tilde{\theta}(t)$ for all $t \in [a, b]$. Moreover, if $\theta(b) = \tilde{\theta}(b)$, then $\theta(t) = \tilde{\theta}(t)$ for all $t \in [a, b]$.

Proof. Let us write

$$\begin{aligned} p(t)\,\frac{d}{dt}u(t) &= r(t)\,\cos\theta(t) \\ u(t) &= r(t)\,\sin\theta(t) \end{aligned}$$

and

$$\begin{aligned} \tilde{p}(t)\,\tilde{u}'(t) &= \tilde{r}(t)\,\cos\tilde{\theta}(t) \\ \tilde{u}(t) &= \tilde{r}(t)\,\sin\tilde{\theta}(t). \end{aligned}$$

The functions $\theta(t)$ and $\tilde{\theta}(t)$ satisfy the differential equations

$$\theta'(t) = q(t)\,\sin^2\theta(t) + \frac{1}{p(t)}\,\cos^2\theta(t)$$

and

$$\begin{aligned} \tilde{\theta}'(t) &= \tilde{q}(t)\,\sin^2\tilde{\theta}(t) + \frac{1}{\tilde{p}(t)}\,\cos^2\tilde{\theta}(t) \\ &\leq q(t)\,\sin^2\tilde{\theta}(t) + \frac{1}{p(t)}\,\cos^2\tilde{\theta}(t). \end{aligned}$$

Hence, the assertion follows from Proposition 7.4. $\square$

As another consequence of the comparison principle, we obtain the following result:

Sturm Comparison Theorem. *Suppose that* $0 < p(t) \leq \tilde{p}(t)$ *and* $q(t) \geq \tilde{q}(t)$ *for all* $t \in [a,b]$. *Moreover, suppose that* $u(t)$ *and* $\tilde{u}(t)$ *are non-trivial solutions of the Sturm-Liouville equations*

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + q(t)\,u(t) = 0$$

and

$$\frac{d}{dt}\Big[\tilde{p}(t)\,\frac{d}{dt}\tilde{u}(t)\Big] + \tilde{q}(t)\,\tilde{u}(t) = 0.$$

If $\tilde{u}(a) = \tilde{u}(b) = 0$, *then there exists a real number* $\tau \in (a,b]$ *and* $u(\tau) = 0$. *In other words, between any two zeroes of the function* $\tilde{u}(t)$ *there is at least one zero of the function* $u(t)$.

Proof. As above, we write

$$\begin{aligned} p(t)\,\frac{d}{dt}u(t) &= r(t)\,\cos\theta(t) \\ u(t) &= r(t)\,\sin\theta(t) \end{aligned}$$

and

$$\begin{aligned} \tilde{p}(t)\,\tilde{u}'(t) &= \tilde{r}(t)\,\cos\tilde{\theta}(t) \\ \tilde{u}(t) &= \tilde{r}(t)\,\sin\tilde{\theta}(t). \end{aligned}$$

We can arrange that $\theta(a) \in [0,\pi)$. Moreover, since $\tilde{u}(a) = 0$, we may assume that $\tilde{\theta}(a) = 0$. Since $\tilde{u}(b) = 0$, it follows that $\tilde{\theta}(b)$ must be an integer multiple of π. Moreover, $\tilde{\theta}(b) > 0$ by Corollary 7.5. Consequently, $\tilde{\theta}(b) \geq \pi$. On the other hand, since $\theta(a) \geq \tilde{\theta}(a)$, it follows from Corollary 7.6 that $\theta(b) \geq \tilde{\theta}(b) \geq \pi$. Since $\theta(a) < \pi$, the intermediate value theorem implies the existence of a real number $\tau \in (a,b]$ such that $\theta(\tau) = \pi$. This implies $u(\tau) = 0$, as claimed. □

7.3. Eigenvalues and eigenfunctions of Sturm-Liouville systems

Suppose that $p(t)$, $q(t)$, and $\rho(t)$ are three functions defined on some time interval $[a,b]$. We assume that $p(t)$ and $q(t)$ are continuously differentiable, and $\rho(t)$ is continuous. Moreover, we assume that $p(t)$ and $\rho(t)$ are positive for all $t \in [a,b]$. In this section, we will consider the Sturm-Liouville system of the form

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\,u(t) = 0, \qquad u(a) = u(b) = 0, \tag{12}$$

where λ is a constant.

Definition 7.7. We say that λ is said to be an eigenvalue of the Sturm-Liouville (12) if the system (12) admits a non-trivial solution. If $u(t)$ is a non-trivial solution of (12), we say that $u(t)$ is an eigenfunction with eigenvalue λ.

Our goal in this section is to analyze the eigenvalues of (12). To that end, we consider an arbitrary number $\lambda \in \mathbb{R}$. It follows from the basic existence and uniqueness theorem that the initial value problem

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\,u(t) = 0, \qquad u(a) = 0,\, u'(a) = 1$$

has a unique solution. We will denote this solution by $u_\lambda(t)$.

Proposition 7.8. *A real number λ is an eigenvalue of (12) if and only if $u_\lambda(b) = 0$. In this case, the function $u_\lambda(t)$ is an eigenfunction, and any other eigenfunction is a constant multiple of the function $u_\lambda(t)$.*

Proof. Suppose that $u(t)$ is an eigenfunction with eigenvalue λ. The functions $u(t)$ and $u'(a)\,u_\lambda(t)$ satisfy the same differential equation with the same initial values. Hence, it follows from the existence and uniqueness theorem that $u(t) = u'(a)\,u_\lambda(t)$ for all $t \in [a,b]$. Since $u(b) = 0$, it follows that $u'(a)\,u_\lambda(b) = 0$. On the other hand, since $u(t)$ is a non-trivial solution, we must have $u'(a) \neq 0$. Therefore, $u_\lambda(b) = 0$, and $u(t)$ is a constant multiple of $u(t)$. Conversely, if $u_\lambda(b) = 0$, then the function $u_\lambda(t)$ is an eigenfunction with eigenvalue λ. $\square$

In order to study the zeroes of the function $u_\lambda(t)$, we use the Prüfer substitution. Let us write

$$p(t)\,\frac{d}{dt}u_\lambda(t) = r_\lambda(t)\,\cos\theta_\lambda(t)$$
$$u_\lambda(t) = r_\lambda(t)\,\sin\theta_\lambda(t),$$

where $\theta_\lambda(a) = 0$. It was shown above that the function $\theta_\lambda(t)$ satisfies the differential equation

$$\theta_\lambda'(t) = (q(t) + \lambda\,\rho(t))\,\sin^2\theta_\lambda(t) + \frac{1}{p(t)}\,\cos^2\theta_\lambda(t).$$

Using Corollary 7.5, we obtain $\theta_\lambda(t) > 0$ for all $t \in (a,b]$.

Proposition 7.9. *A real number λ is an eigenvalue of (12) if and only if $\theta_\lambda(b) = \pi n$ for some positive integer n.*

Proof. By Proposition 7.8, λ is an eigenvalue of (12) if and only if $u_\lambda(b) = 0$. This is equivalent to saying that $\theta_\lambda(b) = \pi n$ for some integer n. Finally, n must be a positive integer since $\theta_\lambda(b) > 0$. $\square$

Proposition 7.10. *If $\theta_\lambda(b) = \pi n$, then the function $u_\lambda(t)$ has exactly $n-1$ zeroes in the interval (a,b).*

Proof. For each $k \in \{0, 1, \ldots, n\}$, we define

$$\tau_k = \inf\{t \in [a,b] : \theta_\lambda(t) \geq \pi k\}.$$

It is easy to see that $a = \tau_0 < \tau_1 < \ldots < \tau_{n-1} < \tau_n = b$ and $\theta_\lambda(t) < \pi k$ for all $t \in [a, \tau_k)$. Moreover, $\theta_\lambda(\tau_k) = \pi k$ for each $k \in \{0, 1, \ldots, n\}$. Using Corollary 7.5, we conclude that $\theta_\lambda(t) > \pi k$ for all $t \in (\tau_k, b]$. Putting these facts together, we conclude that $\pi k < \theta_\lambda(t) < \pi(k+1)$ for all $t \in (\tau_k, \tau_{k+1})$. Consequently, $u_\lambda(\tau_k) = 0$ and $u_\lambda(t) \neq 0$ for all $t \in (\tau_k, \tau_{k+1})$. Therefore, the function $u_\lambda(t)$ has exactly $n-1$ zeroes in the interval (a,b). □

As an example, let us consider the eigenvalue problem

$$\frac{d^2}{dt^2} u(t) + \lambda\, u(t) = 0, \qquad u(a) = u(b) = 0.$$

In this case, the function $u_\lambda(t)$ is given by

$$u_\lambda(t) = \begin{cases} \frac{1}{\sqrt{\lambda}} \sin(\sqrt{\lambda}\,(t-a)) & \text{if } \lambda > 0 \\ t - a & \text{if } \lambda = 0 \\ \frac{1}{\sqrt{-\lambda}} \sinh(\sqrt{-\lambda}\,(t-a)) & \text{if } \lambda < 0. \end{cases}$$

In particular, $u_\lambda(b) = 0$ if and only if λ is positive and $\sqrt{\lambda}\,(t-a) = \pi n$ for some positive integer n. Therefore, the n-th eigenvalue is given by

$$\lambda_n = \frac{\pi^2 n^2}{(b-a)^2},$$

and the associated eigenfunction is given by

$$\text{constant} \cdot \sin\Big(\pi n\, \frac{t-a}{b-a}\Big).$$

Note that $\lambda_n = O(n^2)$.

We now return to the general case. In view of Proposition 7.9, the problem of finding the eigenvalues of (12) comes down to finding all numbers λ for which $\theta_\lambda(b) = \pi n$, where n is a positive integer. This is a nonlinear equation in λ. We will show that this equation has exactly one solution for every positive integer n. This is a consequence of the following theorem:

Theorem 7.11. *The function $\lambda \mapsto \theta_\lambda(b)$ is continuous and strictly increasing. Moreover,*

$$\lim_{\lambda \to -\infty} \theta_\lambda(b) = 0$$

and

$$\lim_{\lambda \to \infty} \theta_\lambda(b) = \infty.$$

Proof. The proof involves several steps:

Step 1: Using Theorem 3.6, it is easy to show that the function $(\lambda, t) \mapsto \theta_\lambda(t)$ is continuous (see also Problem 3.3). In particular, the function $\lambda \mapsto \theta_\lambda(b)$ is continuous.

Step 2: We next show that the function $\lambda \mapsto \theta_\lambda(b)$ is strictly increasing. Suppose that λ and μ are two real numbers such that $\lambda > \mu$. The function $u_\lambda(t)$ satisfies the differential equation

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u_\lambda(t)\Big] + (q(t) + \lambda\,\rho(t))\,u_\lambda(t) = 0.$$

Similarly, the function $u_\mu(t)$ satisfies the differential equation

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u_\mu(t)\Big] + (q(t) + \mu\,\rho(t))\,u_\mu(t) = 0.$$

Since $\lambda > \mu$ and $\rho(t)$ is positive, we have

$$q(t) + \lambda\,\rho(t) > q(t) + \mu\,\rho(t)$$

for all $t \in [a, b]$. Moreover, $\theta_\lambda(a) = \theta_\mu(a) = 0$. By Corollary 7.6, we have $\theta_\lambda(b) \geq \theta_\mu(b)$. We claim that $\theta_\lambda(b) > \theta_\mu(b)$. Indeed, if $\theta_\lambda(b) = \theta_\mu(b)$, then Corollary 7.6 implies that $\theta_\lambda(t) = \theta_\mu(t)$ for all $t \in [a, b]$. Using the differential equations

$$\theta_\lambda'(t) = (q(t) + \lambda\,\rho(t))\,\sin^2\theta_\lambda(t) + \frac{1}{p(t)}\,\cos^2\theta_\lambda(t)$$

and

$$\theta_\mu'(t) = (q(t) + \mu\,\rho(t))\,\sin^2\theta_\mu(t) + \frac{1}{p(t)}\,\cos^2\theta_\mu(t),$$

we conclude that $\lambda = \mu$, contrary to our assumption. Consequently, $\theta_\lambda(b) > \theta_\mu(b)$. Therefore, the function $\lambda \mapsto \theta_\lambda(b)$ is strictly increasing.

Step 3: We now show that $\lim_{\lambda\to\infty}\theta_\lambda(b) = \infty$. Suppose that a positive integer n is given. The function

$$\hat{u}(t) = \sin\Big(\pi n\,\frac{t-a}{b-a}\Big)$$

satisfies the differential equation

$$\frac{d}{dt}\Big[p_{\max}\,\frac{d}{dt}\hat{u}(t)\Big] + \frac{\pi^2 n^2}{(b-a)^2}\,p_{\max}\,\hat{u}(t) = 0.$$

Let us write

$$\begin{aligned} p_{\max}\,\frac{d}{dt}\hat{u}(t) &= \hat{r}(t)\,\cos\hat{\theta}(t) \\ \hat{u}(t) &= \hat{r}(t)\,\sin\hat{\theta}(t), \end{aligned}$$

where $\hat{\theta}(a) = 0$.

Suppose that λ is chosen such that $\lambda > 0$ and

$$\lambda\,\rho_{\min} \geq \frac{\pi^2 n^2}{(b-a)^2}\,p_{\max} - q_{\min}.$$

Then

$$q(t) + \lambda\,\rho(t) \geq \frac{\pi^2 n^2}{(b-a)^2}\,p_{\max}$$

for all $t \in [a,b]$. Moreover,

$$p(t) \leq p_{\max}$$

for all $t \in [a,b]$. Hence, it follows from Corollary 7.6 that $\theta_\lambda(b) \geq \hat{\theta}(b)$. Since $\hat{u}(b) = 0$, the number $\hat{\theta}(b)$ must be an integer multiple of π. Moreover, since the function $\hat{u}(t)$ has exactly $n-1$ zeroes between a and b, we have $\hat{\theta}(b) = \pi n$ by Proposition 7.10. Putting these facts together, we obtain $\theta_\lambda(b) \geq \hat{\theta}(b) = \pi n$. Since n is arbitrary, we conclude that $\lim_{\lambda\to\infty} \theta_\lambda(b) = \infty$.

Step 4: It remains to show that $\lim_{\lambda\to-\infty} \theta_\lambda(b) = 0$. Suppose that a number $\varepsilon \in (0,\pi)$ is given. Let us choose a real number λ such that $\lambda < 0$ and

$$(q_{\max} + \lambda\,\rho_{\min})\,\sin^2\varepsilon + \frac{1}{p_{\min}}\,\cos^2\varepsilon < 0.$$

We claim that $\theta_\lambda(b) \leq \varepsilon$. Suppose this is false. In this case, we define

$$\tau = \inf\{t \in (a,b] : \theta_\lambda(b) > \varepsilon\}.$$

Clearly, $\theta_\lambda(\tau) = \varepsilon$ and $\theta_\lambda'(\tau) \geq 0$. On the other hand,

$$\begin{aligned}
\theta_\lambda'(\tau) &= (q(\tau) + \lambda\,\rho(\tau))\,\sin^2\theta_\lambda(\tau) + \frac{1}{p(\tau)}\,\cos^2\theta_\lambda(\tau) \\
&= (q(\tau) + \lambda\,\rho(\tau))\,\sin^2\varepsilon + \frac{1}{p(\tau)}\,\cos^2\varepsilon \\
&\leq (q_{\max} + \lambda\,\rho_{\min})\,\sin^2\varepsilon + \frac{1}{p_{\min}}\,\cos^2\varepsilon \\
&< 0
\end{aligned}$$

by our choice of λ. This is a contradiction. Consequently, $\theta_\lambda(b) \leq \varepsilon$. Since $\varepsilon > 0$ can be chosen arbitrarily small, we conclude that $\lim_{\lambda\to-\infty} \theta_\lambda(b) = 0$. This completes the proof of Theorem 7.11. □

Theorem 7.12. *There exists an increasing sequence of real numbers $\lambda_1 < \lambda_2 < \ldots$ with the following properties:*

(i) *The real number λ_n is the unique solution of the equation $\theta_\lambda(b) = \pi n$.*

(ii) *$\lambda_n \to \infty$ as $n \to \infty$.*

(iii) *A real number λ is an eigenvalue if and only if there exists a positive integer n such that $\lambda = \lambda_n$.*

(iv) *For every positive integer n, the function $u_{\lambda_n}(t)$ is an eigenfunction with eigenvalue λ_n. Moreover, the function $u_{\lambda_n}(t)$ has exactly $n-1$ zeroes in the open interval (a,b).*

Proof. Let n be a positive integer. It follows from Theorem 7.11 that $\lim_{\lambda\to-\infty}\theta_\lambda(b) = 0$ and $\lim_{\lambda\to\infty}\theta_\lambda(b) = \infty$. Hence, by the intermediate value theorem, there exists a real number λ_n such that $\theta_{\lambda_n}(b) = \pi n$. Since the function $\lambda \mapsto \theta_\lambda(b)$ is strictly increasing, we have $\theta_\lambda(b) < \pi n$ for $\lambda < \lambda_n$ and $\theta_\lambda(b) > \pi n$ for $\lambda > \lambda_n$. Therefore, λ_n is the only solution of the equation $\theta_\lambda(b) = \pi n$. This proves (i).

In order to prove (ii), we observe that $\theta_{\lambda_n}(b) = \pi n$ by definition of λ_n. This implies $\lim_{n\to\infty}\theta_{\lambda_n}(b) = \infty$. Since the function $\lambda \mapsto \theta_{\lambda_n}(b)$ is continuous, this can only happen if $\lim_{n\to\infty}\lambda_n = \infty$.

Finally, (iii) and (iv) follow immediately from Proposition 7.9 and Proposition 7.10. This completes the proof of Theorem 7.12. $\square$

7.4. The Liouville normal form

Consider the eigenvalue problem

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\,u(t) = 0, \qquad u(a) = u(b) = 0. \tag{13}$$

We continue to assume that $p(t), \rho(t) > 0$. We say that a system is in Liouville normal form if $p(t) = \rho(t) = 1$ for all $t \in [a,b]$.

We claim that every system is equivalent to one that is in Liouville normal form. To explain this, let

$$T = \int_a^b \sqrt{\frac{\rho(s)}{p(s)}}\,ds.$$

We then consider the system

$$\frac{d^2}{d\tau^2}w(\tau) + (Q(\tau) + \lambda)\,w(\tau) = 0, \qquad w(0) = w(T) = 0, \tag{14}$$

where $Q : [0,T] \to \mathbb{R}$ is defined by

$$Q\bigg(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\bigg) = \rho(t)^{-1}\,\bigg\{q(t) + (p(t)\,\rho(t))^{\frac{1}{4}}\,\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}(p(t)\,\rho(t))^{-\frac{1}{4}}\Big]\bigg\}.$$

The systems (13) and (14) are equivalent in the following sense:

Theorem 7.13. *Suppose that $u : [a,b] \to \mathbb{R}$ and $w : [0,T] \to \mathbb{R}$ are related by*

$$u(t) = (p(t)\rho(t))^{-\frac{1}{4}}\, w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big).$$

Then u is a solution of (13) if and only if w is a solution of (14).

Proof. Differentiating the identity

$$u(t) = (p(t)\rho(t))^{-\frac{1}{4}}\, w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big)$$

with respect to t gives

$$\begin{aligned} p(t)\, \frac{d}{dt} u(t) &= (p(t)\rho(t))^{\frac{1}{4}}\, \frac{d}{d\tau} w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &\quad + p(t)\, \frac{d}{dt}(p(t)\rho(t))^{-\frac{1}{4}}\, w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big). \end{aligned}$$

This implies

$$\begin{aligned} \frac{d}{dt}\Big[p(t)\, \frac{d}{dt} u(t)\Big] &= \rho(t)\, (p(t)\rho(t))^{-\frac{1}{4}}\, \frac{d^2}{d\tau^2} w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &\quad + \frac{d}{dt}(p(t)\rho(t))^{\frac{1}{4}}\, \frac{d}{d\tau} w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &\quad + (p(t)\rho(t))^{\frac{1}{2}}\, \frac{d}{dt}(p(t)\rho(t))^{-\frac{1}{4}}\, \frac{d}{d\tau} w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &\quad + \frac{d}{dt}\Big[p(t)\, \frac{d}{dt}(p(t)\rho(t))^{-\frac{1}{4}}\Big]\, w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &= \rho(t)\, (p(t)\rho(t))^{-\frac{1}{4}}\, \frac{d^2}{d\tau^2} w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big) \\ &\quad + \frac{d}{dt}\Big[p(t)\, \frac{d}{dt}(p(t)\rho(t))^{-\frac{1}{4}}\Big]\, w\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\Big). \end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
&\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + q(t)\,u(t) \\
&= \rho(t)\,(p(t)\rho(t))^{-\frac{1}{4}}\,\frac{d^2}{d\tau^2}w\bigg(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\bigg) \\
&+ \rho(t)\,(p(t)\rho(t))^{-\frac{1}{4}}\,Q\bigg(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\bigg)\,w\bigg(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\bigg).
\end{aligned}$$

From this, the assertion follows easily. □

In other words, the two systems (13) and (14) have the same eigenvalues, and there is a one-to-one correspondence between eigenfunctions of (13) and eigenfunctions of (14).

7.5. Asymptotic behavior of eigenvalues of a Sturm-Liouville system

In this section, we analyze the asymptotic behavior of the eigenvalues of a Sturm-Liouville system. We first consider a system which is in Liouville normal form:

Proposition 7.14. *Consider the system*

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\,u(t) = 0, \qquad u(a) = u(b) = 0.$$

The n-th eigenvalue of this system can be bounded by

$$\frac{\pi^2 n^2}{(b-a)^2} - q_{\max} \leq \lambda_n \leq \frac{\pi^2 n^2}{(b-a)^2} - q_{\min}.$$

Proof. As before, we denote by $u_\lambda(t)$ the unique solution of the initial value problem

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\,u(t) = 0, \qquad u(a) = 0,\, u'(a) = 1.$$

As usual, we write

$$\begin{aligned}
\frac{d}{dt}u_\lambda(t) &= r_\lambda(t)\,\cos\theta_\lambda(t) \\
u_\lambda(t) &= r_\lambda(t)\,\sin\theta_\lambda(t),
\end{aligned}$$

where $\theta_\lambda(a) = 0$. Moreover, we consider the function

$$\hat{u}(t) = \sin\Big(\pi n\,\frac{t-a}{b-a}\Big).$$

The function $\hat{u}(t)$ is a solution of the Sturm-Liouville equation

$$\frac{d^2}{dt^2}\hat{u}(t) + \frac{\pi^2 n^2}{(b-a)^2}\,\hat{u}(t) = 0.$$

We may write

$$\begin{aligned}\frac{d}{dt}\hat{u}(t) &= \hat{r}(t)\,\cos\hat{\theta}(t)\\ \hat{u}(t) &= \hat{r}(t)\,\sin\hat{\theta}(t),\end{aligned}$$

where $\hat{\theta}(a) = 0$.

We first define

$$\underline{\lambda} = \frac{\pi^2 n^2}{(b-a)^2} - q_{\max} \geq 0.$$

Then

$$q(t) + \underline{\lambda} \leq \frac{\pi^2 n^2}{(b-a)^2}$$

for all $t \in [a,b]$. Hence, Corollary 7.6 implies that $\theta_{\underline{\lambda}}(t) \leq \hat{\theta}(t)$ for all $t \in [a,b]$. In particular,

$$\theta_{\underline{\lambda}}(b) \leq \hat{\theta}(b) = \pi n = \theta_{\lambda_n}(b).$$

Since the function $\lambda \mapsto \theta_\lambda(b)$ is strictly monotone increasing, we conclude that

$$\lambda_n \geq \underline{\lambda} = \frac{\pi^2 n^2}{(b-a)^2} - q_{\max}.$$

In the next step, we define

$$\overline{\lambda} = \frac{\pi^2 n^2}{(b-a)^2} - q_{\min}.$$

Then

$$q(t) + \overline{\lambda} \geq \frac{\pi^2 n^2}{(b-a)^2}$$

for all $t \in [a,b]$. By Corollary 7.6, we have $\theta_{\overline{\lambda}}(t) \geq \hat{\theta}(t)$ for all $t \in [a,b]$. Hence, we obtain

$$\theta_{\overline{\lambda}}(b) \geq \hat{\theta}(b) = \pi n = \theta_{\lambda_n}(b).$$

Since the function $\lambda \mapsto \theta_\lambda(b)$ is strictly monotone increasing, we conclude that

$$\lambda_n \leq \overline{\lambda} = \frac{\pi^2 n^2}{(b-a)^2} - q_{\min}.$$

This completes the proof of Proposition 7.14. $\square$

Corollary 7.15. *Consider the system*

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\, u(t) = 0, \qquad u(a) = u(b) = 0.$$

If λ_n *denotes the* n*-th eigenvalue of this system, then*

$$\lambda_n = \frac{\pi^2 n^2}{(b-a)^2} + O(1).$$

Combining Corollary 7.15 with Theorem 7.13, we can draw the following conclusion:

Theorem 7.16. *Consider the system*

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\, u(t) = 0, \qquad u(a) = u(b) = 0.$$

Let λ_n *be the* n*-th eigenvalue of this system. Then*

$$\lambda_n = \frac{\pi^2 n^2}{T^2} + O(1),$$

where

$$T = \int_a^b \sqrt{\frac{\rho(t)}{p(t)}}\, dt.$$

7.6. Asymptotic behavior of eigenfunctions

We now analyze the asymptotic behavior of the n-th eigenfunction of a Sturm-Liouville system. In view of the discussion above, it suffices to consider systems in Liouville normal form.

Proposition 7.17. *Suppose that* $u : [a,b] \to \mathbb{R}$ *is a solution of the equation*

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\, u(t) = 0$$

with $u(a) = 0$*. Moreover, suppose that* u *is normalized such that*

$$\int_a^b u(t)^2\, dt = 1.$$

If λ *is large, then*

$$u(t) = \pm\sqrt{\frac{2}{b-a}}\, \sin(\sqrt{\lambda}\,(t-a)) + O\Big(\frac{1}{\sqrt{\lambda}}\Big).$$

Proof. Let us fix a large number $\lambda > 0$. The differential equation

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\, u(t) = 0$$

can be rewritten as

$$\frac{d}{dt}\Big[\lambda^{-\frac{1}{2}}\,\frac{d}{dt}u(t)\Big] + (\lambda^{-\frac{1}{2}}\,q(t) + \lambda^{\frac{1}{2}})\,u(t) = 0. \tag{15}$$

We now perform a Prüfer substitution for (15). In other words, we write

$$\begin{aligned} \lambda^{-\frac{1}{2}}\,\frac{d}{dt}u(t) &= r(t)\,\cos\theta(t) \\ u(t) &= r(t)\,\sin\theta(t), \end{aligned}$$

where $\theta(a) = 0$. Without loss of generality, we may assume that $r(a) > 0$. (Otherwise, we replace $u(t)$ by $-u(t)$.) The functions $r(t)$ and $\theta(t)$ satisfy the differential equations

$$r'(t) = -\lambda^{-\frac{1}{2}}\,q(t)\,r(t)\,\cos\theta(t)\,\sin\theta(t)$$

and

$$\theta'(t) = (\lambda^{-\frac{1}{2}}\,q(t) + \lambda^{\frac{1}{2}})\,\sin^2\theta(t) + \lambda^{\frac{1}{2}}\,\cos^2\theta(t).$$

This implies

$$\frac{d}{dt}\log r(t) = -\lambda^{-\frac{1}{2}}\,q(t)\,\cos\theta(t)\,\sin\theta(t)$$

and

$$\frac{d}{dt}(\theta(t) - \lambda^{\frac{1}{2}}\,t) = \lambda^{-\frac{1}{2}}\,q(t).$$

Consequently,

$$\Big|\frac{d}{dt}\log r(t)\Big| \leq C_1\,\lambda^{-\frac{1}{2}}$$

and

$$\Big|\frac{d}{dt}(\theta(t) - \lambda^{\frac{1}{2}}\,t)\Big| \leq C_1\,\lambda^{-\frac{1}{2}},$$

where $C_1 = \sup_{t\in[a,b]}|q(t)|$. Integrating these inequalities gives

$$\Big|\log\frac{r(t)}{r(a)}\Big| \leq C_2\,\lambda^{-\frac{1}{2}}$$

and

$$|\theta(t) - \lambda^{\frac{1}{2}}\,(t-a)| \leq C_2\,\lambda^{-\frac{1}{2}}$$

for all $t \in [a,b]$. This implies

$$\Big|\frac{r(t)}{r(a)} - 1\Big| \leq C_3\,\lambda^{-\frac{1}{2}}$$

and

$$|\sin\theta(t) - \sin(\lambda^{\frac{1}{2}}\,(t-a))| \leq C_3\,\lambda^{-\frac{1}{2}}$$

for all $t \in [a,b]$. Putting these facts together, we obtain

$$\begin{aligned}
\Big|\frac{1}{r(a)}\,u(t) - \sin(\lambda^{\frac{1}{2}}\,(t-a))\Big| &= \Big|\frac{r(t)}{r(a)}\,\sin\theta(t) - \sin(\lambda^{\frac{1}{2}}\,(t-a))\Big| \\
&\leq \Big|\frac{r(t)}{r(a)} - 1\Big| + |\sin\theta(t) - \sin(\lambda^{\frac{1}{2}}\,(t-a))| \qquad (16) \\
&\leq 2\,C_3\,\lambda^{-\frac{1}{2}}.
\end{aligned}$$

We now estimate the term $r(a)$. To that end, we observe that

$$\Big|\frac{1}{r(a)^2}\,u(t)^2 - \sin^2(\lambda^{\frac{1}{2}}\,(t-a))\Big| \leq 8\,C_3\,\lambda^{-\frac{1}{2}}$$

for all $t \in [a,b]$. We now integrate this inequality over the interval $[a,b]$. Using the identities

$$\int_a^b u(t)^2\,dt = 1$$

and

$$\int_a^b \sin^2(\lambda^{\frac{1}{2}}\,(t-a))\,dt = \frac{b-a}{2} - \frac{1}{2\lambda^{\frac{1}{2}}}\,\cos(\lambda^{\frac{1}{2}}\,(b-a))\,\sin(\lambda^{\frac{1}{2}}\,(b-a)),$$

we obtain

$$\Big|\frac{1}{r(a)^2} - \frac{b-a}{2}\Big| \leq C_4\,\lambda^{-\frac{1}{2}}.$$

Consequently,

$$\Big|r(a) - \sqrt{\frac{2}{b-a}}\Big| \leq C_5\,\lambda^{-\frac{1}{2}}.$$

In particular, $r(a) \leq C_6$, where C_6 is a uniform constant that does not depend on λ. Using (16), we obtain

$$|u(t) - r(a)\,\sin(\lambda^{\frac{1}{2}}\,(t-a))| \leq C_7\,\lambda^{-\frac{1}{2}},$$

hence

$$\Big|u(t) - \sqrt{\frac{2}{b-a}}\,\sin(\lambda^{\frac{1}{2}}\,(t-a))\Big| \leq C_8\,\lambda^{-\frac{1}{2}}.$$

This completes the proof. □

Corollary 7.18. *Consider the system*

$$\frac{d^2}{dt^2}u(t) + (q(t)+\lambda)\,u(t) = 0, \qquad u(a) = u(b) = 0.$$

Let λ_n be the n-th eigenvalue of this system. Moreover, suppose that u_n is the corresponding eigenfunction, normalized such that

$$\int_a^b u_n(t)^2\,dt = 1.$$

For n large, we have the asymptotic expansion

$$u_n(t) = \pm\sqrt{\frac{2}{b-a}}\,\sin\Big(\pi n\,\frac{t-a}{b-a}\Big) + O\Big(\frac{1}{n}\Big).$$

Proof. By Proposition 7.17,

$$u_n(t) = \pm\sqrt{\frac{2}{b-a}}\,\sin(\sqrt{\lambda_n}\,(t-a)) + O\Big(\frac{1}{\sqrt{\lambda_n}}\Big).$$

Moreover, it follows from Corollary 7.15 that

$$\sqrt{\lambda_n} = \frac{\pi n}{b-a} + O\Big(\frac{1}{n}\Big).$$

Putting these facts together, the assertion follows. □

Combining Corollary 7.18 with Theorem 7.13, we can draw the following conclusion:

Theorem 7.19. *Consider the system*

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\,u(t) = 0, \qquad u(a) = u(b) = 0.$$

Let λ_n be the n-th eigenvalue of this system. Moreover, suppose that u_n is the corresponding eigenfunction, normalized such that

$$\int_a^b \rho(t)\,u_n(t)^2\,dt = 1.$$

For n large, we have

$$u_n(t) = \pm\sqrt{\frac{2}{T}}\,(p(t)\rho(t))^{-\frac{1}{4}}\,\sin\Big(\frac{\pi n}{T}\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\Big) + O\Big(\frac{1}{n}\Big),$$

where

$$T = \int_a^b \sqrt{\frac{\rho(s)}{p(s)}}\,ds.$$

Proof. We may write

$$u_n(t) = (p(t)\rho(t))^{-\frac{1}{4}}\,w_n\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\Big)$$

for some function $w_n : [0,T] \to \mathbb{R}$. The function w_n satisfies

$$\frac{d^2}{d\tau^2}w_n(\tau) + (Q(\tau) + \lambda_n)\,w_n(\tau) = 0, \qquad w_n(0) = w_n(T) = 0.$$

Using the substitution rule, we obtain

$$\begin{aligned}\int_0^T w_n(\tau)^2\,d\tau &= \int_a^b \sqrt{\frac{\rho(t)}{p(t)}}\, w_n\Big(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\Big)^2\,dt \\ &= \int_a^b \rho(t)\,u_n(t)^2\,dt \\ &= 1.\end{aligned}$$

By Corollary 7.18, the function w_n satisfies

$$w_n(\tau) = \pm\sqrt{\frac{2}{T}}\,\sin\Big(\frac{\pi n}{T}\,\tau\Big) + O\Big(\frac{1}{n}\Big)$$

for n large. Thus,

$$u_n(t) = \pm\sqrt{\frac{2}{T}}\,(p(t)\rho(t))^{-\frac{1}{4}}\,\sin\bigg(\frac{\pi n}{T}\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\,ds\bigg) + O\Big(\frac{1}{n}\Big)$$

for n large. □

7.7. Orthogonality and completeness of eigenfunctions

Let us consider a Sturm-Liouville system of the form

$$\frac{d^2}{dt^2}u(t) + (q(t)+\lambda)\,u(t) = 0, \qquad u(a) = u(b) = 0.$$

Let us denote by $\lambda_1 < \lambda_2 < \ldots$ the eigenvalues of this Sturm-Liouville system, and let $u_1(t), u_2(t), \ldots$ denote the associated eigenfunctions. We assume that the eigenfunctions are normalized such that

$$\int_a^b u_n(t)^2\,dt = 1$$

for all n.

Proposition 7.20. *We have*

$$\int_a^b u_m(t)\,u_n(t)\,dt = 0$$

for $m \neq n$.

Proof. We compute

$$\begin{aligned} 0 &= \int_a^b \frac{d}{dt}\Big(u_m(t)\,\frac{d}{dt}u_n(t) - u_n(t)\,\frac{d}{dt}u_m(t)\Big)\,dt \\ &= \int_a^b \Big(u_m(t)\,\frac{d^2}{dt^2}u_n(t) - u_n(t)\,\frac{d^2}{dt^2}u_m(t)\Big)\,dt \\ &= \int_a^b (\lambda_m - \lambda_n)\,u_m(t)\,u_n(t)\,dt. \end{aligned}$$

Since $\lambda_m \neq \lambda_n$, the assertion follows. □

Corollary 7.21. *For every continuous function v, we have*

$$\sum_{n=1}^{\infty} \Big(\int_a^b u_n(t)\,v(t)\,dt\Big)^2 \leq \int_a^b v(t)^2\,dt.$$

Proof. Let

$$w = v - \sum_{n=1}^{m} \Big(\int_a^b u_n(t)\,v(t)\,dt\Big)\,u_n.$$

Then

$$0 \leq \int_a^b w(t)^2\,dt = \int_a^b v(t)^2\,dt - \sum_{n=1}^{m} \Big(\int_a^b u_n(t)\,v(t)\,dt\Big)^2.$$

Since m is arbitrary, the assertion follows. □

In the remainder of this section, we establish a completeness property of the set of eigenfunctions. More precisely, we will show that

$$\int_a^b v(t)^2\,dt = \sum_{n=m+1}^{\infty} \Big(\int_a^b u_n(t)\,v(t)\,dt\Big)^2$$

for every continuous function v. In order to prove this, we will follow the arguments in Birkhoff and Rota's book [**2**]. Let

$$\hat{u}_n(t) = \sqrt{\frac{2}{b-a}}\,\sin\Big(\pi n\,\frac{t-a}{b-a}\Big).$$

It is a well known fact that every continuous function on the interval $[a,b]$ can be represented by a Fourier series. Moreover, for every continuous function v we have

$$\int_a^b v(t)^2\,dt = \sum_{n=m+1}^{\infty} \Big(\int_a^b \hat{u}_n(t)\,v(t)\,dt\Big)^2.$$

This relation is known as Parseval's identity. Moreover, Corollary 7.18 implies that

$$\int_a^b (u_n(t) - \hat{u}_n(t))^2\,dt = O\Big(\frac{1}{n^2}\Big),$$

hence

$$\sum_{n=1}^{\infty} \int_a^b (u_n(t) - \hat{u}_n(t))^2 \, dt < \infty.$$

Proposition 7.22. *Fix an integer m such that*

$$\sum_{n=m+1}^{\infty} \int_a^b (u_n(t) - \hat{u}_n(t))^2 \, dt \leq \frac{1}{4}.$$

Moreover, suppose that v is a function satisfying

$$\int_a^b \hat{u}_n(t) \, v(t) \, dt = 0$$

for $n = 1, \ldots, m$. Then

$$\int_a^b v(t)^2 \, dt \leq 4 \sum_{n=m+1}^{\infty} \bigg(\int_a^b u_n(t) \, v(t) \, dt \bigg)^2.$$

Proof. Using Parseval's identity for Fourier series, we obtain

$$\begin{aligned}
\int_a^b v(t)^2 \, dt &= \sum_{n=m+1}^{\infty} \bigg(\int_a^b \hat{u}_n(t) \, v(t) \, dt \bigg)^2 \\
&\leq 2 \sum_{n=m+1}^{\infty} \bigg(\int_a^b u_n(t) \, v(t) \, dt \bigg)^2 \\
&\quad + 2 \sum_{n=m+1}^{\infty} \bigg(\int_a^b (u_n(t) - \hat{u}_n(t)) \, v(t) \, dt \bigg)^2 \\
&\leq 2 \sum_{n=m+1}^{\infty} \bigg(\int_a^b u_n(t) \, v(t) \, dt \bigg)^2 \\
&\quad + 2 \sum_{n=m+1}^{\infty} \bigg(\int_a^b (u_n(t) - \hat{u}_n(t))^2 \, dt \bigg) \bigg(\int_a^b v(t)^2 \, dt \bigg) \\
&\leq 2 \sum_{n=m+1}^{\infty} \bigg(\int_a^b u_n(t) \, v(t) \, dt \bigg)^2 + \frac{1}{2} \int_a^b v(t)^2 \, dt.
\end{aligned}$$

From this, the assertion follows. □

Corollary 7.23. *Fix an integer m such that*

$$\sum_{n=m+1}^{\infty} \int_a^b (u_n(t) - \hat{u}_n(t))^2 \, dt \leq \frac{1}{4}.$$

Moreover, suppose that $v \in \operatorname{span}\{u_1, \ldots, u_m\}$ is a function satisfying

$$\int_a^b \hat{u}_n(t) \, v(t) \, dt = 0$$

for $n = 1, \ldots, m$. *Then* $v = 0$.

Proposition 7.24. *Fix a continuous function* v *and a real number* $\varepsilon > 0$. *Then there exists an integer* $m > 0$ *and a function* $w \in \operatorname{span}\{u_1, \ldots, u_m\}$ *such that*

$$\int_a^b (v(t) - w(t))^2 \, dt \leq \varepsilon.$$

Proof. Let us fix an integer m such that

$$\sum_{n=m+1}^{\infty} \int_a^b (u_n(t) - \hat{u}_n(t))^2 \, dt \leq \frac{1}{4}$$

and

$$\sum_{n=m+1}^{\infty} \left(\int_a^b u_n(t) \, v(t) \, dt \right)^2 \leq \frac{\varepsilon}{4}.$$

By Corollary 7.23, we can find a function $w \in \operatorname{span}\{u_1, \ldots, u_m\}$ such that

$$\int_a^b \hat{u}_n(t) \, (v(t) - w(t)) \, dt = 0$$

for $n = 1, \ldots, m$. Using Proposition 7.22, we obtain

$$\begin{aligned} \int_a^b (v(t) - w(t))^2 \, dt &\leq 4 \sum_{n=m+1}^{\infty} \left(\int_a^b u_n(t) \, (v(t) - w(t)) \, dt \right)^2 \\ &= 4 \sum_{n=m+1}^{\infty} \left(\int_a^b u_n(t) \, v(t) \, dt \right)^2 \\ &\leq \varepsilon, \end{aligned}$$

as claimed. □

Corollary 7.25. *For every continuous function* v, *we have*

$$\sum_{n=1}^{\infty} \left(\int_a^b u_n(t) \, v(t) \, dt \right)^2 = \int_a^b v(t)^2 \, dt.$$

Proof. Given any $\varepsilon > 0$, we can find an integer $m > 0$ and a function $w \in \operatorname{span}\{u_1, \ldots, u_m\}$ such that

$$\sum_{n=1}^{m} \left(\int_a^b u_n(t) \, (v(t) - w(t)) \, dt \right)^2 \leq \int_a^b (v(t) - w(t))^2 \, dt \leq \varepsilon.$$

This gives

$$\begin{aligned}\sqrt{\int_a^b v(t)^2\,dt} &\leq \sqrt{\int_a^b w(t)^2\,dt} + \sqrt{\varepsilon} \\ &= \sqrt{\sum_{n=1}^m \bigg(\int_a^b u_n(t)\,w(t)\,dt\bigg)^2} + \sqrt{\varepsilon} \\ &\leq \sqrt{\sum_{n=1}^m \bigg(\int_a^b u_n(t)\,v(t)\,dt\bigg)^2} + 2\sqrt{\varepsilon}.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the asserion follows. □

Combining Corollary 7.18 with Theorem 7.13, we obtain the following result:

Theorem 7.26. *Consider the system*

$$\frac{d}{dt}\Big[p(t)\,\frac{d}{dt}u(t)\Big] + (q(t) + \lambda\,\rho(t))\,u(t) = 0, \qquad u(a) = u(b) = 0.$$

Let λ_n be the n-th eigenvalue of this system. Moreover, suppose that u_n is the corresponding eigenfunction, normalized such that

$$\int_a^b \rho(t)\,u_n(t)^2\,dt = 1.$$

Then

$$\sum_{n=1}^\infty \bigg(\int_a^b \rho(t)\,u_n(t)\,v(t)\,dt\bigg)^2 = \int_a^b \rho(t)\,v(t)^2\,dt$$

for every continuous function v.

7.8. Problems

Problem 7.1. Suppose that a and b are real numbers such that $a < b$. For which real numbers λ does the boundary value problem

$$\frac{d^2}{dt^2}u(t) + \lambda\,u(t) = 0, \qquad u(a) = u'(b) = 0$$

have a non-trivial solution?

Problem 7.2. Let a and b be two real numbers such that $a < b$, and let $q(t)$ be a continuous function defined on $[a, b]$. Suppose that the boundary value problem

$$\frac{d^2}{dt^2}u(t) + (q(t) + \lambda)\,u(t) = 0, \qquad u(a) = u'(b) = 0$$

has a non-trivial solution $u(t)$. Moreover, suppose that $q(t) \leq 0$ for all $t \in [a, b]$. Show that $\lambda \geq \frac{\pi^2}{4(b-a)^2}$.

Problem 7.3. Let $T = \int_a^b \sqrt{\frac{\rho(s)}{p(s)}}\, ds$. Moreover, suppose that $u : [a, b] \to \mathbb{R}$ and $w : [0, T] \to \mathbb{R}$ are related by

$$u(t) = (p(t)\rho(t))^{-\frac{1}{4}}\, w\bigg(\int_a^t \sqrt{\frac{\rho(s)}{p(s)}}\, ds\bigg).$$

Show that

$$\int_a^b \rho(t)\, u(t)^2\, dt = \int_0^T w(\tau)^2\, d\tau.$$

Problem 7.4. This problem is concerned with a generalization of Proposition 7.20. Suppose that $u(t)$ is a solution of the Sturm-Liouville system

$$\frac{d}{dt}\Big[p(t)\, \frac{d}{dt} u(t)\Big] + (q(t) + \lambda\, \rho(t))\, u(t) = 0, \qquad u(a) = u(b) = 0.$$

Moreover, let $v(t)$ be a solution of the Sturm-Liouville system

$$\frac{d}{dt}\Big[p(t)\, \frac{d}{dt} v(t)\Big] + (q(t) + \mu\, \rho(t))\, v(t) = 0, \qquad v(a) = v(b) = 0.$$

Show that

$$\int_a^b \rho(t)\, u(t)\, v(t)\, dt = 0$$

if $\lambda \neq \mu$.

Problem 7.5. Suppose that $u(t)$ is a solution of the Sturm-Liouville system

$$\frac{d}{dt}\Big[p(t)\, \frac{d}{dt} u(t)\Big] + (q(t) + \lambda\, \rho(t))\, u(t) = 0, \qquad u(a) = u(b) = 0.$$

We assume that $p(t)$, $q(t)$, and $\rho(t)$ are real-valued functions, and $p(t), \rho(t) > 0$. Finally, we assume that λ is a complex number with $\mathrm{Im}(\lambda) \neq 0$. Show that u vanishes identically.

Problem 7.6. Let $p(t)$ and $q(t)$ be positive functions on $[a, b]$ with the property that the product $p(t)q(t)$ is monotone increasing. Moreover, suppose that $u(t)$ is a solution of the differential equation

$$\frac{d}{dt}\Big[p(t)\, \frac{d}{dt} u(t)\Big] + q(t)\, u(t) = 0.$$

Finally, let t_k be an increasing sequence of times such that u attains a local extremum at t_k. Show that the sequence $u(t_k)^2$ is monotone decreasing. This result is known as the Sonin-Pólya-Butlewski theorem. (Hint: Consider the function $u(t)^2 + \frac{p(t)\, u'(t)^2}{q(t)}$.)

Bibliography

[1] H. Amann, *Ordinary differential equations: an introduction to nonlinear analysis,* Walter de Gruyter, Berlin, 1990

[2] G. Birkhoff and G.-C. Rota, *Ordinary differential equations,* 4th edition, Wiley, New York, 1989

[3] M.W. Hirsch, S. Smale, and R.L. Devaney, *Differential equations, dynamical systems, and an introduction to chaos,* 2nd edition, Elsevier/Academic Press, Amsterdam, 2004

[4] L. Perko, *Differential equations and dynamical systems,* 3rd edition, Springer-Verlag, New York, 2001

[5] S. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering,* Perseus Book Publishing, 1994

[6] G. Teschl, *Ordinary differential equations and dynamical systems,* Graduate Studies of Mathematics, vol. 140, American Mathematical Society, Providence RI, 2012