Math 6410 : Ordinary Differential Equations

Chee Han Tan

Last modified : January 13, 2018

Contents

Preface

These notes are largely based on **Math 6410: Ordinary Differential Equations** course, taught by Paul Bressloff in Fall 2015, at the University of Utah. Additional examples or remarks or results from other sources are added as we see fit, mainly to facilitate our understanding. These notes are by no means accurate or applicable, and any mistakes here are of course our own. Please report any typographical errors or mathematical fallacy to me by email tan@math.utah.edu.

Chapter 1

Initial Value Problems

1.1 Introduction

We begin by studying the initial value problem (IVP)

$$\begin{cases} \dot{x} = f(x), & x \in U, \quad f\colon U \longrightarrow \mathbb{R}^n \\ x(0) = x_0 \in U, & x = (x_1, \ldots, x_n)^T \end{cases} \tag{IVP}$$

where U is an open subset of $\mathbb{R}^n$. Unless stated otherwise, f will be as smooth as we need them to be. First order ODEs are of fundamental importance, due to the fact that any nth order ODE can be written as a system of n first order ODEs.

Example 1.1.1 (Newtonian Dynamics)**.** Consider $m\ddot{x} = -\gamma\dot{x} + F(x)$, where $-\gamma\dot{x}$ is the damping term and $F(x)$ some external force. This can be written as a system of first order ODEs:

$$\begin{cases} \dot{x} &= y = f_1(x, y), \\ \dot{y} &= -\dfrac{\gamma}{m}y + \dfrac{F(x)}{m} = f_2(x, y). \end{cases}$$

- Observe that we have a conservative system if $\gamma = 0$. Indeed, let $V(x) = -\int^x F(s)\,ds$ be the potential function. Multiplying the ODE by $\dot{x}$ yields

  $$m\dot{x}\ddot{x} = -\left(\frac{dV}{dx}\right)\dot{x} \implies \frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2\right) = -\frac{d}{dt}V(x(t)).$$

  Thus, we obtain the well-known conservation of energy: $\dfrac{1}{2}m\dot{x}^2 + V(x) = E$.

- Let $p = m\dot{x}$ be the momentum, and define the Hamiltonian as

  $$H = H(x, p) = \frac{1}{2m}p^2 + V(x).$$

  We have the following Hamilton's equations:

  $$\dot{x} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial x}.$$

An important property of Hamiltonian is they are conserved over time, *i.e.*

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial p}\dot{p} = 0.$$

Example 1.1.2 (Classical Predator-Prey Model)**.** Let $N(t)$ be the amount of bacteria, and assume that N is large enough so that we can treat it as a continuous variable. One simple model is the following:

$$\dot{N} = NK\left(1 - \frac{N}{N_{max}}\right) - d, \quad N(0) = N_0.$$

Non-dimensionalising the system with the dimensionless parameters

$$x = \frac{N}{N_{max}}, \quad \tau = Kt, \quad D = \frac{d}{KN_{max}}$$

yields the dimensionless equation $\dot{x} = (1-x)x - D$. This is a bifurcation problem, with bifurcation paramter D. If $0 < D < 1/4$, then there eixsts 2 fixed points; otherwise there is no fixed points.

From a geometric perspective, it seems rather natural to think of solution to (IVP) as a function of initial value $x(0) = x_0$ and time t. This motivates the definition of a flow.

Definition 1.1.3. Let $U \subset \mathbb{R}^n$ and $f \in C^1(U)$. Given $x_0 \in U$, let $\phi_t(x_0) = \phi(x_0, t)$ be the solution of (IVP) with $t \in I(x_0)$, where $I(x_0)$ is the maximal time interval over which a unique solution exists. Such set of mappings ϕ_t is called the **flow** of the vector field f.

- For a fixed $x_0 \in U$, $\phi(x_0, \cdot)\colon I(x_0) \longrightarrow U$ defines a curve or ***trajectory***, $\Gamma(x_0)$ of the system through x_0. Specifically,

$$\Gamma(x_0) = \Big\{x \in U \colon x = \phi_t(x_0),\, t \in I(x_0)\Big\}.$$

- If x_0 varies over a set $M \subset U$, then the flow $\phi_t \colon M \longrightarrow U$ defines the motion of a cloud of points in M.

Theorem 1.1.4 (Time-shift invariance)**.** *For any $x_0 \in U$, if $t \in I(x_0)$ and $s \in I(\phi_t(x_0))$, $s > 0$, then $s + t \in I(x_0)$ and $\phi_{s+t}(x_0)$ can be expressed as $\phi_{s+t}(x_0) = \phi_s(\phi_t(x_0))$.*

Proof. Let $I(x_0) = (\alpha, \beta)$ be the maximal time interval over which there exists a unique solution. Consider the function $x\colon (\alpha, s+t) \longrightarrow U$ defined by

$$x(r) = \begin{cases} \phi(r, x_0) & \text{if} \quad r \in (\alpha, t], \\ \phi(r - t, \phi_t(x_0)) & \text{if} \quad r \in [t, s+t]. \end{cases}$$

We see that $x(r)$ is a solution of the IVP on $(\alpha, s+t)$. Hence, $s + t \in I(x_0)$ and uniqueness of solution gives

$$\phi_{s+t}(x_0) = x(s+t) = \phi(s, \phi_t(x_0)) = \phi_s(\phi_t(x_0)).$$

■

Remark 1.1.5. The result is trivial if $s = 0$. The result holds for $s < 0$ by evolving backwards in time. A corollary of the above theorem is that $\phi_{-t}(\phi_t(x_0)) = \phi_t(\phi_{-t}(x_0)) = x_0$, i.e. we have an abelian group structure here!

Remark 1.1.6. There exists ODEs in which solution does not exist globally, even if $f \in C^1(U)$. Consider the ODE $\dot{x} = f(x) = \dfrac{1}{x}$, where $f \in C^1(U)$ with $U = \{x \in \mathbb{R} \colon x > 0\}$. For an initial condition $x(0) = x_0 \in U$, a solution to this IVP is given by $\phi_t(x_0) = \sqrt{2t + x_0^2}$, with $I(x_0) = \left(\dfrac{-x_0^2}{2}, \infty\right)$.

1.2 Planar Dynamics

Consider the general autonomous second order system

$$\begin{cases} \dot{x} &= X(x, y) \\ \dot{y} &= Y(x, y) \end{cases}$$

Assuming existence and uniqueness, then paths only cross at fixed points.

1.3 Existence and Uniqueness

In this section, we focus on establishing sufficient conditions such that there exists a unique solution to (IVP). Not suprisingly, regularity of $f \colon U \longrightarrow \mathbb{R}^n$ plays an important role, as we shall see below.

(a) Continuity of f is not sufficient to guarantee uniqueness of a solution.

- Consider $\dot{x} = 3x^{2/3}$, $x(0) = 0$. One solution is the trivial solution $x(t) \equiv 0$ for all $t \geq 0$. Another solution is obtained by using separation of variable:

$$\frac{1}{3}\int_0^x \frac{1}{s^{2/3}}\, ds = t \implies x(t) = t^3.$$

- Observe that f is continuous but not differentiable at the origin.

(b) A solution can become unbounded at some finite time $t = T$, *i.e.* finite-time blow up.

- Consider $\dot{x} = x^2$, $x(0) = 1$. We have two branches of solution, $x(t) = \dfrac{1}{1-t}$, depending on the initial condition.
- In this case, our solution is only defined on $t \in (-\infty, 1)$, and $\lim\limits_{t \to 1^-} x(t) = \infty$.
- The other branch is defined on $t \in (1, \infty)$, but is not reached by the given initial condition.

Definition 1.3.1.

(a) We say that $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is ***differentiable*** at $x_0 \in \mathbb{R}^n$ if there exists a linear transformation $Df(x_0) \in \mathcal{L}(\mathbb{R}^n)$, called the $n \times n$ **Jacobian matrix** such that

$$\lim_{|h|\to 0} \frac{|f(x_0+h) - f(x_0) - Df(x_0)h|}{|h|} = 0.$$

(b) Let U be an open subset of $\mathbb{R}^n$. A function $f\colon U \longrightarrow \mathbb{R}^n$ is said to satisfy a ***Lipschitz condition*** on U if there exists $K > 0$ such

$$|f(x) - f(y)| \leq K|x-y| \quad \text{for all } \; x, y \in U.$$

(c) A function $f\colon U \longrightarrow \mathbb{R}^n$ is said to be ***locally Lipschitz*** on U if for every $x_0 \in U$, there exists an ε-neighbourhood $N_\varepsilon(x_0) \subset U$ and $K_0 > 0$ such that

$$|f(x) - f(y)| \leq K_0|x-y| \quad \text{for all } \; x, y \in N_\varepsilon(x_0).$$

Theorem 1.3.2. $C^1(U) \implies$ *locally Lipschitz on* U*.*

Proof. Choose an arbitrary $x_0 \in U$ and suppose $x, y \in B_\varepsilon(x_0) \subset U$, with $K = \max\limits_{x\in B_\varepsilon(x_0)} \|Df(x)\|$. Convexity of $B_\varepsilon(x_0)$ means that $x + su \in B_\varepsilon(x_0)$ for $0 \leq s \leq 1$, with $u = y - x$. Let $F(s) = f(x+su)$, then $F'(s) = Df(x+su)\cdot u$ and

$$\begin{aligned} |f(y) - f(x)| = |F(1) - F(0)| = \left|\int_0^1 F'(s)\,ds\right| &\leq \int_0^1 |Df(x+su)\cdot u|\,ds \\ &\leq K\int_0^1 |x-y|\,ds = K|x-y|. \end{aligned}$$

∎

1.3.1 Picard's Method of Successive Approximation

Let $x(t)$ be a solution to (IVP), that is $x(t)$ is a continuous function satisfying the integral equation

$$x(t) = x_0 + \int_0^t f(x(s))\,ds,$$

and vice-versa. Successive approximations to the solution of the integral equation are defined by the sequence of functions

$$u_0(t) = x_0, \quad u_{k+1}(t) = x_0 + \int_0^t f(u_k(s))\,ds. \tag{1.3.1}$$

Example 1.3.3. Consider $\dot{x} = Ax$ with $x(0) = x_0$.

$$u_1(t) = x_0 + \int_0^t Ax_0\,ds \qquad\qquad = x_0(1 + At).$$

$$\begin{aligned}
u_2(t) &= x_0 + \int_0^t A[x_0(1+As)]\,ds &&= x_0\left(1 + At + \frac{A^2t^2}{2}\right). \\
&\vdots \qquad\qquad \vdots && \qquad \vdots \\
u_k(t) &= x_0\left(1 + At + \ldots + \frac{(At)^k}{k!}\right) &&\longrightarrow x_0 e^{At} \text{ as } k \longrightarrow \infty.
\end{aligned}$$

Theorem 1.3.4 (Fundamental Theorem of Existence and Uniqueness of ODEs)**.**
Let U be an open subset of $\mathbb{R}^n$ containing x_0, and $f \in C^1(U)$. There exists $a > 0$ such that the (IVP) *has a unique solution on $I = [-a, a]$.*

Proof. We prove that the sequence of successive approximations $(u_k)_{k=1}^{\infty}$ given by (1.3.1) converges to a solution. Using completeness of $C^0[-a,a]$ with respect to the uniform norm, it suffices to show that (u_k) is a Cauchy sequence. It is clear from (1.3.1) and $f \in C^1(U)$ that $(u_k) \in C^0(U)$, in particular $(u_k) \in C^0[-a,a]$. The proof is divided into four parts.

(A) Theorem 1.3.2 implies that f is locally Lipschitz on U. Given $x_0 \in U$, choose $b = \varepsilon/2 > 0$ and consider $N_b(x_0) \subset N_\varepsilon(x_0)$ on which f is Lipschitz with Lipschitz constant $K > 0$. Let $M = \max\limits_{N_b(x_0)} |f(x)|$. **We need to choose $\boldsymbol{a > 0}$ such that $\boldsymbol{(u_k) \in N_b(x_0)}$ for all $\boldsymbol{k \geq 1}$.** Suppose $a > 0$ is sufficiently small such that

$$\max_{t\in[-a,a]} |u_k(t) - x_0| \leq b. \tag{1.3.2}$$

A direct estimate shows that

$$|u_{k+1}(t) - x_0| \leq \int_0^t |f(u_k(s))|ds \leq Ma \quad \forall t \in [-a,a].$$

By choosing $0 < a \leq b/M$, it follows from induction that the sequence of succesive approximations (u_k) satisfies (1.3.2).

(B) Observe that to control the difference between any two approximations, it suffices to control the difference between two successive approximations. For the first two successive approximations,

$$\begin{aligned}
|u_2(t) - u_1(t)| &\leq \int_0^t |f(u_1(s)) - f(u_0(s))|\,ds \\
&\leq K \int_0^t |u_1(s) - u_0(s)|\,ds && \Big[f \text{ is locally Lipschitz}\Big] \\
&\leq Ka \max_{t\in[-a,a]} |u_1(t) - x_0| && \Big[u_0 = x_0\Big] \\
&\leq Kab && \Big[\text{from } (1.3.2)\Big]
\end{aligned}$$

We generalise this estimate by induction. Assuming that

$$\max_{t\in[-a,a]} |u_j(t) - u_{j-1}(t)| \leq (Ka)^{j-1}b, \tag{1.3.3}$$

for some $j \geq 1$. From (1.3.1), for $t \in [-a, a]$ we have that

$$\begin{aligned}
|u_{j+1}(t) - u_j(t)| &\leq \int_0^t |f(u_j(s)) - f(u_{j-1}(s))|\, ds \\
&\leq K \int_0^t |u_j(s) - u_{j-1}(s)|\, ds && \Big[f \text{ is locally Lipschitz}\Big] \\
&\leq Ka \max_{t \in [-a,a]} |u_j(t) - u_{j-1}(t)| \\
&\leq (Ka)^j b && \Big[\text{from (1.3.3)}\Big]
\end{aligned}$$

(C) We are ready to show that (u_k) is a Cauchy sequence in $C^0[-a, a]$. By choosing $0 < a < \dfrac{1}{K}$, we see that for all $m > n \geq N$ and $t \in [-a, a]$ we have

$$\begin{aligned}
|u_m(t) - u_n(t)| \leq \sum_{j=n}^{m-1} |u_{j+1}(t) - u_j(t)| &\leq \sum_{j=N}^{\infty} |u_{j+1}(t) - u_j(t)| \\
&\leq \sum_{j=N}^{\infty} (Ka)^j b \\
&= \left[\frac{(Ka)^N}{1 - Ka}\right] b \longrightarrow 0 \text{ as } N \longrightarrow \infty.
\end{aligned}$$

Thus, for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\|u_m - u_n\| = \max_{t \in [-a,a]} |u_m(t) - u_k(t)| < \varepsilon \quad \text{for all } m, n \geq N,$$

i.e. (u_k) is a Cauchy sequence. Completeness of $C^0[-a, a]$ with respect to the supremum norm implies that $u_k(t) \longrightarrow u(t)$ uniformly for all $t \in [-a, a]$ as $k \longrightarrow \infty$, for some $u(t) \in C^0[-a, a]$. Taking the limit as $k \longrightarrow \infty$ in (1.3.1) yields

$$\begin{aligned}
u(t) &= x_0 + \int_0^t f(u(s))\, ds \\
u(t) &= \lim_{k \to \infty} u_k(t).
\end{aligned}$$

i.e. $u(t)$ is a solution to (IVP). Note that interchange between limits and integrals is allowed here due to uniform convergence of (u_k) in $C^0[-a, a]$.

(D) Finally, we prove uniqueness. Suppose $u(t)$, $v(t)$ are solutions to (IVP) on $[-a, a]$. **Extreme Value Theorem** states that continuous function $|u(t) - v(t)|$ achieves its maximum at some $t_1 \in [-a, a]$.

$$\begin{aligned}
\|u - v\| = \max_{t \in [-a,a]} |u(t) - v(t)| &= \left| \int_0^{t_1} \Big(f(u(s)) - f(v(s)) \Big)\, ds \right| \\
&\leq \int_0^{|t_1|} |f(u(s)) - f(v(s))|\, ds \\
&\leq K \int_0^{|t_1|} |u(s) - v(s)|\, ds
\end{aligned}$$

$$\leq Ka \max_{t\in[-a,a]} |u(t) - v(t)|$$
$$= Ka\|u - v\|.$$

Since $0 < Ka < 1$, we must have $\|u - v\| = 0 \implies u(t) = v(t)$ for all $t \in [-a, a]$.

■

Remark 1.3.5.

1. Observe that $a > 0$ is chosen such that the sequence of successive approximations (u_k) remains in a neighbourhood where f is Lipschitz around x_0, and such that (u_k) is a Cauchy sequence in $C^0[-a, a]$. In this proof, $a > 0$ is chosen such that
$$0 < a < \min\left\{\frac{b}{M}, \frac{1}{K}\right\}.$$

2. One could also apply the **Contraction Mapping Principle**, which is a powerful machinery in existence and uniqueness problems. A similar result (and proof) holds for non-autonomous ODEs, where f is assumed to be C^1 with respect to x and C^0 with respect to t.

1.3.2 Dependence on Initial Conditions

We will now prove **Gronwall's inequality**, which is perhaps one of the most important tool in the theory of ODEs. In short, if a function satisfies an integral inequality implicitly, then Gronwall's inequality gives an explicit bound on the function itself.

Theorem 1.3.6 (Gronwall's Lemma)**.** *Suppose v, u, c are positive functions on $[0, t]$ and c is a differentiable function. Suppose $v(t)$ satisfies*

$$v(t) \leq c(t) + \int_0^t u(s)v(s)\, ds.$$

The following inequality holds

$$v(t) \leq c(0)\exp\left\{\int_0^t u(s)\, ds\right\} + \int_0^t c'(s)\exp\left\{\int_s^t u(\tau)\, d\tau\right\} ds.$$

Proof. The main idea is to derive estimates for the second term independent of $v(t)$. Let $R(t) = \displaystyle\int_0^t u(s)v(s)\, ds$, then

$$\dot{R}(t) = u(t)v(t) \leq u(t)\left[c(t) + \int_0^t u(s)v(s)\, ds\right] = u(t)[c(t) + R(t)].$$

Rearranging the above inequality and using the **integrating factor** technique gives

$$\dot{R}(t) - u(t)R(t) \leq c(t)u(t)$$

$$\frac{d}{dt}\left[\exp\left\{-\int_0^t u(s)\,ds\right\} R(t)\right] \leq \exp\left\{-\int_0^t u(s)\,ds\right\} c(t)u(t).$$

Integrating both sides with respect to t yields

$$\begin{aligned}\exp\left\{-\int_0^t u(\tau)\,d\tau\right\} R(t) \leq R(0) + \int_0^t \exp\left\{-\int_0^s u(\tau)\,d\tau\right\} c(s)u(s)\,ds \\ R(t) \leq \int_0^t \exp\left\{\int_s^t u(\tau)\,d\tau\right\} c(s)u(s)\,ds\end{aligned}$$

Hence,

$$\begin{aligned} v(t) \leq c(t) + R(t) &\leq c(t) + \int_0^t \exp\left\{\int_s^t u(\tau)\,d\tau\right\} c(s)u(s)\,ds \\ &= c(t) + \int_0^t c(s)\left[-\frac{d}{ds}\exp\left\{\int_s^t u(\tau)\,d\tau\right\}\right]\,ds\end{aligned}$$

where the negative sign is due to differentiating the lower limit of the integral. Finally, integrating by parts yields

$$\begin{aligned} v(t) &\leq c(t) - \left[c(s)\exp\left\{\int_s^t u(\tau)\,d\tau\right\}\right]\Bigg|_{s=0}^{s=t} + \int_0^t c'(s)\exp\left\{\int_s^t u(\tau)\,d\tau\right\} ds \\ &= c(0)\exp\left\{\int_0^t u(s)\,ds\right\} + \int_0^t c'(s)\exp\left\{\int_s^t u(\tau)\,d\tau\right\} ds.\end{aligned}$$

■

Theorem 1.3.7 (Dependence on Initial Conditions)**.** *Consider the following IVPs*

$$\begin{cases}\dot{x} = f(x),\ x(0) = y, \\ \text{with solution } x_0(t), t \in I.\end{cases} \qquad \begin{cases}\dot{x} = f(x),\ x(0) = y + h, \\ \text{with solution } x_\varepsilon(t), t \in I.\end{cases}$$

where f is Lipschitz continuous in x with Lipschitz constant L, $|h| \leq \varepsilon$, $\varepsilon > 0$, and I is the maximum time interval in which solution exists. The following holds for all $t \in I$

$$|x_\varepsilon(t) - x_0(t)| \leq \varepsilon e^{Lt}.$$

Proof. These IVPs are equivalent to the integral equation

$$\begin{cases} x_0(t) &= y + \displaystyle\int_0^t f(x_0(s))\,ds \\ x_\varepsilon(t) &= y + h + \displaystyle\int_0^t f(x_\varepsilon(s))\,ds.\end{cases}$$

Taking the difference yields

$$\begin{aligned}|x_\varepsilon(t) - x_0(t)| &\leq |h| + \int_0^t |f(x_\varepsilon(s)) - f(x_0(s))|\,ds \\ &\leq \varepsilon + L\int_0^t |x_\varepsilon(s) - x_0(s)|\,ds.\end{aligned}$$

Applying Gronwall's Lemma with $v(s) = |x_\varepsilon(s) - x_0(s)|, u(s) = L$ and $c(s) = \varepsilon$ yields the desired result.

■

1.4 Manifolds

1.5 Limit Cycle and Poincaré-Bendixson Theorem

Given a nonlinear dynamical systems, one usually locates fixed points of the system and analyses the behaviour of solutions in the neighbourhood of each fixed point. However, one could also look for periodic solutions, or solutions that form a closed curve eventually. The latter motivates the notion of limit cycles, which has been widely used in modelling behaviour of oscillatory systems, such as the well-known Van der Pol oscillator.

Definition 1.5.1. A **limit cycle** is an isolated periodic solution of an autonomous system represented in the phase plane by an isolated closed path.

- Autonomous linear systems cannot exhibit limit cycles.
- Cannot usually establish existence of a limit cycle for a nonlinear system.
  - Reduce to 2D (Poincaré-Bendixson), or
  - Finite dimension (Hopf bifurcation).

Remark 1.5.2. It turns out that the continuity of the vector field f imposes strong restriction on the possible arrangements of fixed points and periodic orbits. One can define the **Poincaré index** $I(\Gamma)$ of a closed curve Γ to be the number of times f rotates anti-clockwise as we go around Γ in the anti-clockwise direction. A limit cycle has index $I(\Gamma) = +1$ since the vector $f(x)$ is tangential to Γ at every point on it. It can be shown that the sum of indices of the fixed points enclosed by a limit cycle is $+1$. More examples:

- Closed curve without fixed points: $I(\Gamma) = 0$.
- Saddle: $I(\Gamma) = -1$.
- Sinks/sources: $I(\Gamma) = 1$.

Consequently, a periodic orbit cannot encloses a saddle and a sink/source because then $I(\Gamma) = 0 \neq +1$.

This next result provides a method of establishing non-existence of limit cycles.

Lemma 1.5.3 (Bendixson's Negative Criterion)**.** *There are no closed paths in a simply connected region of the phase plane, on which* $\dfrac{\partial X}{\partial x} + \dfrac{\partial Y}{\partial y}$ *is of one sign.* $(\nabla \cdot f \neq 0)$

Proof. Assume by contradiction that there exists a closed path C in a region D, where $\nabla \cdot f$ has one sign. By the **Divergence Theorem**,

$$\iint_M \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}\right) dxdy = \oint_C (X, Y) \cdot \mathbf{n}\, ds.$$

Since C is always tangent to the vector field (X, Y), the normal vector $\mathbf{n}$ to C is always perpendicular to the vector field (X, Y). Consequently, the integrand $(X, Y) \cdot \mathbf{n} \equiv 0$ and this contradicts the assumption that the integral on LHS cannot vanish due to $\nabla \cdot f \neq 0$. ■

Example 1.5.4 (Dampled nonlinear oscillator)**.**
Consider $\ddot{x} + p(x)\dot{x} + q(x) = 0$, where p, q are smooth functions and $p(x) > 0$. Rewrite this as a system of first order ODEs:

$$\begin{cases} \dot{x} &= y = X(x, y). \\ \dot{y} &= -q(x) - p(x)y = Y(x, y). \end{cases}$$

Thus, $\dfrac{\partial X}{\partial x} + \dfrac{\partial Y}{\partial y} = -p(x) < 0$, which implies that there are no contractible orbits.

1.5.1 Orbits, Invariant Sets and ω-Limit Sets

In order to discuss about limit cycles, one must first understand the long term behaviour of a dynamical system. This motivates the idea of limit sets, which can be described using flow map, a notion we first seen in Definition 1.1.3.

Definition 1.5.5.

(a) Consider the ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$. Solution to this equation defines a **flow**, $\phi(x, t)$, which satisfies

$$\begin{cases} \dfrac{d}{dt}\Big(\phi(x, t)\Big) = f(\phi(x, t)) \\ \phi(x, t_0) = x_0. \end{cases}$$

(b) A point x is **periodic of minimal period T** if and only if

(i) $\phi(x, t + T) = \phi(x, t)$ for all $t \in \mathbb{R}$, and

(ii) $\phi(x, t + s) \neq \phi(x, t)$ for all $s \in (0, T)$.

The curve $\Gamma = \{\phi(x, t) \colon 0 \leq t < T\}$ is called a **periodic orbit**, and it is a closed curve.

(c) A set M is **invariant under the flow $\boldsymbol{\phi}$** if and only if for all $x \in M$, $\phi(x, t) \in M$ for all $t \in \mathbb{R}$.

- Forward (backward) invariant if this holds for all $t > 0$ ($t < 0$).

(d) Suppose that the flow is defined for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The **orbit/trajectory** through x in $\mathbb{R}^n$ is the set $\gamma(x) = \bigcup\limits_{t \in \mathbb{R}} \phi(x, t)$.

- Positive semi-trajectory through x is the set $\gamma^+(x) = \bigcup\limits_{t > 0} \phi(x, t)$.
- Negative semi-trajectory through x is the set $\gamma^-(x) = \bigcup\limits_{t < 0} \phi(x, t)$.
- A set M is invariant (under ϕ) if and only if $\gamma(x) \in M$ for all $x \in M$.

Definition 1.5.6.

(a) The **ω-limit set** of x is

$$\omega(x) = \{y \in \mathbb{R}^n : \exists (t_n)_{n=1}^{\infty} \longrightarrow \infty \text{ s.t. } \phi(x, t_n) \longrightarrow y \text{ as } n \longrightarrow \infty\}.$$

Note that y is a limit point of $\gamma^+(x)$.

(b) The **α-limit set** of x is

$$\alpha(x) = \{y \in \mathbb{R}^n : \exists (\tau_n)_{n=1}^{\infty} \longrightarrow -\infty \text{ s.t. } \phi(x, \tau_n) \longrightarrow y \text{ as } n \longrightarrow \infty\}.$$

Note that y is a limit point of $\gamma^-(x)$.

Definition 1.5.7.

(a) An invariant set $M \subset \mathbb{R}^n$ is an **attracting set** of $\dot{x} = f(x)$ if there exists some neighbourhood N of M such that for all $x \in N$, $\phi(x, t) \longrightarrow M$ as $t \longrightarrow \infty$, and $\phi(x, t) \in N$ for all $t \geq 0$.

(b) An **attractor** is an attracting set which contains a dense orbit.

(c) If M is an attracting set, then the **basin of attraction** of M, $B(M)$ is

$$B(M) = \{x \in \mathbb{R}^n : \phi(x, t) \longrightarrow M \text{ as } t \longrightarrow \infty\}.$$

- Consider a stable limit cycle, Γ. Then $\omega(x) = \Gamma$ if x lies in the basin of attraction of Γ.

Theorem 1.5.8 (Properties of $\omega(x)$)**.**

(a) $\omega(x)$ is closed and invariant.

(b) If the positive orbit $\gamma^+(x)$ is bounded, then $\omega(x)$ is non-empty and compact. [Similar result holds for $\alpha(x)$, with $\gamma^-(x)$ bounded.]

Proof. **To show that $\omega(x)$ is closed**, consider a sequence of points in $\omega(x)$. Suppose that $y_k \in \omega(x)$ and $y_k \longrightarrow \bar{y} \in \mathbb{R}^n$ as $k \longrightarrow \infty$. For each $p \in \mathbb{N}$, there exists $k_p > 0$ such that $d(y_{k_p}, \bar{y}) < \dfrac{1}{p}$. Since $y_{k_p} \in \omega(x)$, for each $k_p > 0$, there exists $t_p > t_{p-1} + 1$ such that $d(\phi(x, t_p), y_{k_p}) < \dfrac{1}{p}$. By triangle-inequality,

$$\begin{aligned} d(\phi(x, t_p), \bar{y}) &\leq d(\phi(x, t_p), y_{k_p}) + d(y_{k_p}, \bar{y}) \\ &< \frac{2}{p} \longrightarrow 0 \quad \text{as } p \longrightarrow \infty. \end{aligned}$$

with (t_p) an increasing sequence to ∞. Thus, $\bar{y} \in \omega(x)$.

To show that $\omega(x)$ is invariant, suppose that $p \in \omega(x)$, then there exists a sequence $(t_n) \longrightarrow \infty$ such that $\phi(x, t_n) \longrightarrow p$ as $n \longrightarrow \infty$. We need to show that $\phi(p, t) \in \omega(x)$ for any $t \in \mathbb{R}$. Setting $\tilde{t}_n = t + t_n$ and applying Theorem 1.1.4 yields (for a fixed $t > 0$)

$$\phi(\tilde{t}_n, x) = \phi(t + t_n, x) = \phi(t, \phi(t_n, x)) \longrightarrow \phi(t, p) \quad \text{as } n \longrightarrow \infty.$$

Hence, the positive orbit containing p lies in $\omega(x)$, and thus $\omega(x)$ is positive-invariant.

If $\gamma^+(x)$ is bounded, then $\omega(x)$ is bounded. Recall that a bounded set in $\mathbb{R}^n$ with infinitely number of points has at least one accumulation point. This implies that $\omega(x)$ is non-empty. Since $\omega(x)$ is a closed, bounded subset of $\mathbb{R}^n$, **Heine-Borel theorem** implies that $\omega(x)$ is compact.

■

1.5.2 Local Transversals

Definition 1.5.9. A **local transversal**, L is a line segment such that all trajectories of the ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^2$ cross from the same side.

- If x_0 is not a fixed point, then one can always construct a local transversal in a neighbourhood of x_0 by continuity.
- It is a C^1 arc on which $f \cdot \underline{\mathbf{n}} \neq 0$, where $\underline{\mathbf{n}}$ is the outward unit normal to L. Thus, f is never tangent to L near x_0 and $f \cdot \mathbf{n}$ has constant sign, since otherwise f must be tangent to L at some point.

Lemma 1.5.10. *If a trajectory $\gamma(x)$ intersects a local transversal L several times, then the successive crossing points move monotonically along L.*

Proof. This is a consequence of the **Jordan-Curve lemma**: A closed curve in the plane separates the plane into 2 connected components: exterior (unbounded) and interior (bounded). The closed curve from P_1 to P_2 (union of shaded region and the line connecting P_1 and P_2) defines an interior, within which orbit cannot re-enter. This implies that P_3 must be beyond P_2.

■

Corollary 1.5.11. *If $x \in \omega(x_0)$ is not a fixed point, and $x \in \gamma(x_0)$, then $\gamma(x)$ is a closed curve.*

Proof. Since $x \in \gamma(x_0)$, it follows that $\omega(x) = \omega(x_0)$. Choose L to be a local transversal through x. Since $x \in \omega(x)$, there exists an increasing sequence $(t_n) \longrightarrow \infty$ such that $\phi(x, t_n) \longrightarrow x$ as $n \longrightarrow \infty$, with $\phi(x, t_n) \in L$ and $\phi(x, 0) = x$. We see immediately that we must have $\phi(x, t_n) = x$ for all $n \geq 1$, otherwise Lemma 1.5.10 implies that the succesive points $\phi(x, t_n)$ on L move monotonically away from x, contradicting $x \in \omega(x)$.

■

Remark 1.5.12. If $x \in \omega(x_0)$, then by definition, $\gamma^+(x_0)$ comes arbitrarily close to x as $t \longrightarrow \infty$, so it makes intersections with a local transversal L through x, arbitrary close to x.

1.5.3 Poincaré-Bendixson Theorem

Theorem 1.5.13 (Poincaré-Bendixson Theorem)**.** *Suppose that a trajectory $\gamma(x_0)$*

(a) enters and does not leave some compact region D, and

(b) there are no fixed points in D.

Then there is at least 1 periodic orbit in D, and this orbit lies in $\omega(x_0)$.

- If we have a compact, positively-invariant region without any fixed points, then the theorem gives that $\omega(x_0)$ is a periodic orbit for all x_0 in the region.
- Doesn't rule out existence of several periodic orbits.
- Different $\omega(x_0)$ could have different periodic orbits.

Proof. Since $\gamma(x_0)$ enters and does not leave the compact domain D, $\omega(x_0)$ is non-empty and is contained in D. Choose $x \in \omega(x_0)$, and note that x is not a fixed point by assumption, so we can define a local transversal L, through x. There are two possible cases:

1. $\boldsymbol{x \in \gamma(x_0)}$, so $\gamma(x)$ is a periodic orbit from Corollary 1.5.11
2. $\boldsymbol{x \notin \gamma(x_0)}$. Since $x \in \omega(x_0)$, we have $\gamma^+(x) \subset \omega(x_0)$ since $\omega(x_0)$ is positive-invariant, so $\gamma^+(x) \subset \omega(x_0) \subset D$. As D is compact, $\gamma^+(x)$ has a limit point $x^* \in D$ such that $x^* \in \omega(x) \subset \omega(x_0)$. There are two possible cases:
   (a) $\boldsymbol{x^* \in \gamma^+(x)}$, so $\gamma(x^*)$ is a periodic orbit from Corollary 1.5.11.
   (b) $\boldsymbol{x^* \notin \gamma^+(x)}$, which leads to a contradiction as we will show now. Indeed, choose a local transversal L through x^*. Since $x^* \in \omega(x) \subset \omega(x_0)$, the trajectory $\gamma^+(x)$ must intersect L at points $P_1, P_2, \ldots$ that accumulate monotonically on x^*. However, these points $(P_j)_{j=1}^{\infty} \in \omega(x_0) \bigcap L$ since $\gamma^+(x) \subset \omega(x_0)$. Hence, $\gamma(x_0)$ passes arbitrarily close to P_j, then P_{j+1}, and so on, infinitely number of times. This implies that $\gamma(x_0)$ intersections with L are not monotonic, which contradicts Lemma 1.5.10.

■

Remark 1.5.14.

1. The last sentence of the proof might be confusing, but the contradiction arises from the following corollary of Lemma 1.5.10: If L is a local transversal, then for any $z \in \mathbb{R}^2$, $\omega(z) \cap L$ contains at most one point.
2. $\boldsymbol{\omega(x_0)}$ **is actually a periodic orbit.** To prove this, it suffices to show that $\phi(x_0, t) \longrightarrow \gamma^+(x)$. Take a local transversal L' through x. There are two possible cases:
   (a) $x_0 \in \gamma^+(x)$, which trivially means $\omega(x_0) = \omega^+(x)$.
   (b) $x_0 \notin \gamma^+(x)$. Since $x \in \omega(x_0)$, $\gamma^+(x_0)$ intersects L' arbitrarily close to x as $(t_j) \longrightarrow \infty$. For any neighbourhood N of $\gamma^+(x)$, since $\phi(x_0, t_j) \longrightarrow x$ as $j \longrightarrow \infty$, there exists $j \in \mathbb{N}$ such that $\phi(x_0, t) \in N$ for all $t \in [t_j, t_{j+1}]$. This is also true for all $t \geq t_j$. Hence, $\phi(x_0, t) \longrightarrow \gamma^+(x)$ as $t \longrightarrow \infty$ and $\omega(x_0)$ contains no points outside $\gamma^+(x)$.
3. Poincaré-Bendixson theorem applies to sphere or cylinder, but not on torus since the Jordan Curve Theorem fails for simple closed curves on a torus!
4. In practice, one typically looks for an annular region $\mathcal{D}$ with

(a) a source in the hole (so trajectory enters $\mathcal{D}$ across the inner boundary),

(b) and the outer boundary is chosen so that trajectory are inward on this boundary.

More precisely, we choose $\mathcal{D}$ to be

$$\mathcal{D} = \{(r, \theta) \colon R_1 - \varepsilon \leq r \leq R_2 + \varepsilon\},$$

such that

$$\dot{r} > 0 \quad \text{for } 0 < r < R_1, \quad \dot{r} < 0 \text{ for } r > R_2, \quad \dot{\theta} \neq 0 \text{ in } \mathcal{D}.$$

Remark 1.5.15. One usually converts from Cartesian (x, y) to polar coordinates (r, θ) in analysing planar systems. The following expression for $\dot{r}$ and $\dot{\theta}$ are useful to keep in mind:

$$\dot{r} = x\dot{x} + y\dot{y}, \quad r^2\dot{\theta} = x\dot{y} - y\dot{x}.$$

Example 1.5.16. Consider the following ODE

$$\begin{cases} \dot{x} &= y + \dfrac{1}{4}x(1 - 2r^2) \\ \dot{y} &= -x + \dfrac{1}{2}y(1 - r^2) \\ r^2 &= x^2 + y^2. \end{cases}$$

- At a fixed point,

$$\begin{aligned} y &= -\frac{1}{4}x(1 - 2r^2) \\ x &= \frac{1}{2}y(1 - r^2) \\ \implies yx &= -\frac{1}{8}xy(1 - 2r^2)(1 - r^2) \\ \implies 0 &= xy\left[1 + \frac{1}{8}(1 - 2r^2)(1 - r^2)\right]. \end{aligned}$$

  Either $x = 0$, $y = 0$ or $(1 - 2r^2)(1 - r^2) = -8 \implies 2r^4 - 3r^2 + 9 = 0$; this equation has no real solution for $r \neq 0$. Therefore, the origin $(0, 0)$ is the only fixed point.

- Computing $\dot{r}$ using $r\dot{r} = x\dot{x} + y\dot{y}$ yields:

$$\begin{aligned} r\dot{r} &= x\dot{x} + y\dot{y} \\ &= \frac{1}{4}x^2(1 - 2r^2) + \frac{1}{2}y^2(1 - r^2) \\ &= \frac{1}{4}x^2 - \frac{1}{2}r^2x^2 + \frac{1}{2}y^2 - \frac{1}{2}r^2y^2 \\ &= \frac{1}{4}r^2 - \frac{1}{2}r^2 + \frac{1}{4}r^2\sin^2(\theta) \\ &= \frac{1}{4}r^2[1 + \sin^2(\theta)] - \frac{1}{2}r^4. \end{aligned}$$

- Thus,

$$\begin{aligned} r\dot{r} = \frac{1}{4}r^2[1+\sin^2(\theta)] - \frac{1}{2}r^4 \geq \frac{1}{4}r^2 - \frac{1}{2}r^4 > 0 &\iff \frac{1}{4}r^2 - \frac{1}{2}r^4 > 0 \\ &\iff \frac{1}{4}r^2(1-2r^2) > 0 \\ &\iff r^2 < \frac{1}{2},\ r < \frac{1}{\sqrt{2}} \\ r\dot{r} = \frac{1}{4}r^2[1+\sin^2(\theta)] - \frac{1}{2}r^4 \leq \frac{1}{2}r^2 - \frac{1}{2}r^4 < 0 &\iff r^2(1-r^2) < 0 \\ &\iff r^2 > 1,\ r > 1. \end{aligned}$$

We then choose $D = \left\{(r,\theta)\colon R_1 \leq r \leq R_2, R_1 < \frac{1}{\sqrt{2}}, R_2 > 1\right\}$.

1.6 Problems

1. The simple pendulum consists of a point mass m suspended from a fixed point by a massless rod L, which is allowed to swing in a vertical plane. If friction is ignored, then the equation of motion is

$$\ddot{x} + \omega^2\sin(x) = 0, \quad \omega^2 = \frac{g}{L}, \tag{1.6.1}$$

where x is the angle of inclination of the rod with respect to the downward vertical and g is the gravitational constant.

(a) Using the conservation of energy, show that the angular velocity of the pendulum satisfies

$$\dot{x} = \pm\sqrt{2}(C + \omega^2\cos(x))^{1/2},$$

where C is an arbitrary constant. Express C in terms of the total energy system.

Solution: Multiplying (1.6.1) by $\dot{x}$ and simplifying yields

$$\ddot{x}\dot{x} + \omega^2\sin(x)\dot{x} = 0$$

(b) Plot or sketch the phase diagram of the pendulum equation. That is, set up Cartesian axes x, y called the *phase plane* with $y = \dot{x}$ and illustrate the one parameter family of curves given by part (a) for different values of C. Take $-3\pi \leq x \leq 3\pi$ and indicate the fixed points of the system and the separatrices, curves linking the fixed points. Give a physical interpretation of the underlying trajectories of the two distinct dynamical regimes $|C| < \omega^2$ and $|C| > \omega^2$.

Solution:

(c) Show that in the regime where $|C| < \omega^2$, the period of oscillations is

$$T = 4\sqrt{\frac{L}{g}}K(\sin(x_0/2)),$$

where $\dot{x} = 0$ when $x = x_0$ and K is the complete elliptic integral of the first kind, defined by

$$K(\alpha) := \int_0^{\pi/2} \frac{1}{\sqrt{1 - \alpha^2 \sin^2(u)}} \, du.$$

Hint: Derive an integral expression for T and then perform the change of variables

$$\sin(u) = \frac{\sin(x/2)}{\sin(x_0/2)}.$$

Solution:

(d) For small amplitude oscillations, the pendulum equation can be approximated by the linear equation

$$\ddot{x} + \omega^2 x = 0.$$

Solve this equation for the initial conditions $x(0) = A, \dot{x}(0) = 0$ and sketch the phase-plane for different values of A. Compare with the phase plane for the full-nonlinear equation in part (b).

Solution:

(e) Write down Hamilton's equations for the pendulum and show that they are equivalent to the second order pendulum equation.

Solution:

2. The displacement x of a spring-mounted mass under the action of dry friction is assumed to satisfy

$$m\ddot{x} + kx = F_0 \text{sgn}(v_0 - \dot{x}). \tag{1.6.2}$$

An example would be a mass m connected to a fixed support by a spring with stiffness k and resting on a conveyor belt moving with speed v_0. F_0 is the frictional force between the mass and the belt. Set $m = k = 1$ for convenience and let $y = \dot{x}$.

(a) Calculate the phase paths in the (x, y) plane and draw the phase diagram. *Hint*: any trajectory that hits the line $y = v_0$ and $|x| < F_0$ moves horizontally at a rate v_0 to the point $x = F_0$ and $y = v_0$. Deduce that the system ultimately converges into a limit cycle oscillation. What happens if $v_0 = 0$?

Solution:

(b) Suppose $v_0 = 0$ and the initial conditions are $x = x_0 > 0, \dot{x} = 0$. Show that the phase path will spiral exactly n times before entering an equilibrium if

$$(4n - 1)F_0 < x_0 < (4n + 1)F_0.$$

Solution:

(c) Suppose $v_0 = 0$ and the initial conditions at $t = 0$ are $x = x_0 > 3F_0$ and $\dot{x} = 0$. Subsequently, whenever $x = -\alpha$ where $2F_0 = -x_0 < -\alpha < 0$ and $\dot{x} > 0$, a trigger operates to increase suddenly the forward velocity so that the kinetic energy increases by a constant amount E. Show that if $E > 8F_0^2$ then a periodic motion is approached and show that the largest value of x in the periodic motion is $F_0 + E/(4F_0)$.

Solution:

(d) In part (b), suppose that the energy is increase by E at $x = -\alpha$ for both $\dot{x} < 0$ and $\dot{x} > 0$; that is, there are two injections of energy per cycle. Show that periodic motion is possible if $E > 6F_0^2$, and find the amplitude of the oscillation.

Solution:

3. The interaction between two species is governed by the deterministic model

$$\begin{cases} \dot{H} & = (a_1 - b_1 H - c_1 P)H, \\ \dot{P} & = (-a_2 + c_2 H)P, \end{cases}$$

where $H \geq 0$ is the population of the host or prey and $P \geq 0$ is the population of the parasite or predator. All constants are positive. Find the fixed points of the system, identify nullclines and sketch the phase diagram. *Hint*: There can be either 2 or 3 fixed points.

Solution:

4. The response of a certain biological oscillator to a stimulus given by a constant b is described by

$$\begin{cases} \dot{x} & = x - ay + b, \\ \dot{y} & = x - cy, \end{cases}$$

where $x, y \geq 0$. Note if $x = 0$ and $y > b/a$ then we simply set $\dot{x} = 0$. Show that when $c < 1$ and $4a > (1 + c)^2$, then there exists a limit cycle, part of which lies on the y-axis, whose period is independent of b. Sketch the corresponding solution.

Solution:

5. Show that the initial value problem

$$\dot{x} = |x|^{1/2}, \quad x(0) = 0,$$

has four different solutions through the origin. Sketch these solutions in the (t, x)-plane.

Solution:

6. Consider the initial value problem $\dot{x} = x^2, x(0) = 1$.

   (a) Find the first three successive approximations $u_1(t), u_2(t), u_3(t)$. Use mathematical induction to show that for $n \geq 1$,

   $$u_n(t) = 1 + t + \ldots + t^n + \mathcal{O}(t^{n+1}) \quad \text{as } t \longrightarrow 0.$$

   Solution:

   (b) Solve the IVP and show that the function $x(t) = \dfrac{1}{1-t}$ is a solution to the IVP on the interval $(-\infty, 1)$. Also show that the first $n+1$ terms in $u_n(t)$ agree with the first $n+1$ terms in the Taylor series for $x(t) = \dfrac{1}{1-t}$ about $x = 0$.

   Solution:

7. Let $f \in C^1(U; \mathbb{R}^n)$, where $U \subset \mathbb{R}^n$ and $x_0 \in U$. Given the Banach space $X = C([0,T]; \mathbb{R}^n)$ with norm $\|x| = \max\limits_{t \in [0,T]} |x(t)|$, let

   $$K(x)(t) = x_0 + \int_0^t f(x(s))\, ds,$$

   for $x \in X$. Define $V = \{x \in X \colon \|x - x_0\| \leq \varepsilon\}$ for fixed $\varepsilon > 0$ and suppose $K(x) \in V$ (which holds for sufficiently small $T > 0$, so that $K \colon V \longrightarrow V$ with V a closed subset of X.

   (a) Using the fact that f is locally Lipschitz in U with Lipschitz constant L_0, and taking $x, y \in V$, show that

   $$|K(x(t)) - K(y(t))| \leq L_0 t \|x - y\|.$$

   Hence, show that

   $$\|Kx - Ky\| \leq L_0 T \|x - y\|, \quad x, y \in V.$$

   Solution:

   (b) Choosing $T < 1/L_0$, apply the contraction mapping principle to show that the integral equation has a unique continuous solution $x(t)$ for all $t \in [0, T]$ and sufficiently small T. Hence establish the existence and uniqueness of the initial value problem

   $$\frac{dx}{dt} = f(x), \quad x(0) = x_0.$$

Solution:

8. Consider the dynamical system described by

$$\begin{cases} \dot{x} & = -y + x(1 - z^2 - x^2 - y^2) \\ \dot{y} & = x + y(1 - z^2 - x^2 - y^2) \\ \dot{z} & = 0. \end{cases}$$

Determine the invariant sets and attracting set of the system. Determine the ω-limit set of any trajectory for which $|z(0)| < 1$. Sketch the flow.

Solution:

9. Consider the dynamical system described by

$$\begin{cases} \dot{x} = -y + x(1 - x^2 - y^2) \\ \dot{y} = x + y(1 - x^2 - y^2) \\ \dot{z} = \alpha > 0. \end{cases}$$

(a) Determine the invariant sets and attracting set of the system. Sketch the flow.

Solution:

(b) Describe what happens to the flow if we identify the points $(x, y, 0)$ and $(x, y, 2\pi)$ in the planes $z = 0$ and $z = 2\pi$. *Hint*: One of the invariant sets becomes a torus with $x^2 + y^2 = 1$.

Solution:

(c) By explicitly constructing solutions of the invariant torus $x^2 + y^2 = 1, 0 \leq z < 2\pi$, show that the torus is only an attractor if α is irrational.

Solution:

10. Use the Poincaré-Bendixson Theorem and the fact that the planar system

$$\begin{cases} \dot{x} & = x - y - x^3 \\ \dot{y} & = x + y - y^3, \end{cases}$$

has only one critical point at the origin to show that this system has a periodic orbit in the annular region

$$A = \{x \in \mathbb{R}^2 \colon 1 < |x| < \sqrt{2}\}.$$

Hint: Convert to polar coordinates and show that for all $\varepsilon > 0$, we have $\dot{r} < 0$ on the circle $r = \sqrt{2} + \varepsilon$ and $\dot{r} > 0$ on the circle $r = 1 - \varepsilon$. Then use the Poincaré-Bendixson Theorem to show that this implies there is a limit cycle in the closure of A, and then show that no limit cycle can have a point in common with either one of the circles $r = 1$ or $r = \sqrt{2}$.

Solution:

11. Show that the system

$$\begin{cases} \dot{x} &= x - rx - ry + xy, \\ \dot{y} &= y - ry + rx - x^2, \end{cases}$$

can be written in polar coordinates as $\dot{r} = r(1-r)$ and $\dot{\theta} = r(1-\cos\theta)$. Show that it has an unstable node at the origin and a saddle at $(1,0)$. Use this information and the Poincaré-Bendixson Theorem to sketch the phase portrait and then deduce that for all $(x,y) \neq (0,0)$, the flow $\phi_t(x,y) \longrightarrow (1,0)$ as $t \longrightarrow \infty$ but that $(1,0)$ is not linearly stable.

Solution:

Chapter 2

Linear Systems and Stability of Nonlinear Systems

2.1 Autonomous Linear Systems

We begin with the study of autonomous linear first order system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0 \in \mathbb{R}^n, \tag{2.1.1}$$

where $A \in \mathbb{R}^{n\times n}$. If $A \in \mathbb{R}$, it follows from **separation of variables** that the solution to (2.1.1) is given by $x(t) = x_0 e^{At}$. This result generalises to $A \in \mathbb{R}^{n\times n}$, but it requires some understanding of the term e^{At}.

2.1.1 Matrix Exponential

Definition 2.1.1 (Matrix Exponential)**.** Let $A \in \mathbb{R}^{n\times n}$. The **exponential of A** is defined by the power series

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{2.1.2}$$

Theorem 2.1.2. *For any $A \in \mathbb{R}^{n\times n}$, the power series (2.1.2) is absolutely convergent. That is, e^A is well-defined.*

Proof. Recall that the operator norm/induced matrix norm of a matrix $A \in \mathbb{R}^{n\times n}$ is defined to be

$$\|A\| = \sup_{x\in\mathbb{R}^n, x\neq \mathbf{0}} \frac{\|Ax\|}{\|x\|} = \sup_{x\in\mathbb{R}^n, \|x\|=1} \|Ax\|.$$

Observe that $\|A^k\| \leq \|A\|^k$ for every $k \geq 1$. Indeed,

$$\|AB\| = \sup_{x\in\mathbb{R}^n, \|x\|=1} \|ABx\| \leq \sup_{x\in\mathbb{R}^n, \|x\|=1} \|A\|\|Bx\| = \|A\|\|B\|.$$

This immediately implies

$$\sum_{k=0}^{\infty} \frac{\|A^k\|}{k!} \leq \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = e^{\|A\|} < \infty.$$

■

Theorem 2.1.3. *Let $A \in \mathbb{R}^{n\times n}$ be a constant coefficient matrix. The unique solution of* (2.1.1) *is $x(t) = e^{tA}x_0$.*

Proof. Let $E_m(t) = \displaystyle\sum_{k=0}^{m} \frac{t^k A^k}{k!}$, then $E_m(t) \in C^1(\mathbb{R}; \mathbb{R}^{n\times n})$ and

$$\dot{E_m}(t) = \frac{d}{dt}\left(\sum_{k=0}^{m} \frac{t^k A^k}{k!}\right) = \sum_{k=1}^{m} \frac{t^{k-1}A^k}{(k-1)!} = A\sum_{k=1}^{m} \frac{t^{k-1}A^{k-1}}{(k-1)!} = AE_{m-1}(t).$$

Theorem 2.1.2 together with the **Weierstrass M-test** gives uniform convergence of $E_m(t)$. It follows that $\dot{E_m}(t)$ converges uniformly and $\lim_{m\to\infty} E_m(t) = E(t)$ is differentiable with derivative $\dot{E}(t) = AE(t)$. Hence,

$$\frac{d}{dt}(e^{tA}x_0) = \frac{d}{dt}(e^{tA})x_0 = Ae^{tA}x_0 = Ax(t).$$

To prove uniqueness, suppose $y(t)$ is another solution with $y(0) = x_0$. Set $z(t) = e^{-tA}y(t)$, then $\dot{z} = -Ae^{-tA}y + e^{-tA}\dot{y} = \mathbf{0}$. This implies that $z(t) = \text{constant} = x_0$, since $z(0) = y(0) = x_0$. ■

Remark 2.1.4. We are able to deduce inductively from the relation $\dot{E}(t) = AE(t)$ that $E(t) \in C^\infty(\mathbb{R}; \mathbb{R}^{n\times n})$.

2.1.2 Normal Forms

Even though we show existence and uniqueness of solution to (2.1.1), in order to understand the dynamics of (2.1.1), we need to explicitly compute the matrix exponential e^{tA}. This is closedly related to the eigenvalues of the matrix A, which should not be surprising at all!

Real Distinct Eigenvalues

Suppose A has n distinct eigenvalues $(\lambda_j)_{j=1}^n$, with corresponding eigenvectors $(\underline{\mathbf{e}}_j)_{j=1}^n$. They satisfy $A\lambda_j = \lambda_j\underline{\mathbf{e}}_j$ for each $j = 1, \dots, n$. Introducing the matrix $P = [\underline{\mathbf{e}}_1, \dots, \underline{\mathbf{e}}_n]$, with eigenvectors as columns. Since we have distinct eigenvalues, the set of eigenvectors is linearly independent and $\det(P) \neq 0$. Also,

$$\begin{aligned} AP = [A\underline{\mathbf{e}}_1, \dots, A\underline{\mathbf{e}}_n] &= [\lambda_1\underline{\mathbf{e}}_1, \dots, \lambda_n\underline{\mathbf{e}}_n] \\ &= [\underline{\mathbf{e}}_1, \dots, \underline{\mathbf{e}}_n]\text{diag}(\lambda_1, \dots, \lambda_n) \\ &= P\Lambda, \text{ or } \Lambda = P^{-1}AP. \end{aligned}$$

Now, performing a change of variable $x = Py$ and using $\dot{x} = Ax$ yields

$$\dot{y} = P^{-1}\dot{x} = P^{-1}Ax = P^{-1}APy = \Lambda y.$$

Now that the system is decoupled, we can solve each of them separately, *i.e.*

$$\dot{y_j} = \lambda_j y_j \implies y_j(t) = e^{\lambda_j t}y_j(0) \quad \text{for each } j = 1, \dots, n.$$

In vector form, we have

$$\begin{cases} y(t) &= e^{t\Lambda}y(0), \\ e^{t\Lambda} &= \text{diag }(e^{\lambda_1 t}, \dots, e^{\lambda_n t}). \end{cases}$$

Hence, $x(t) = Py(t) = Pe^{t\Lambda}y(0) = Pe^{t\Lambda}P^{-1}x(0)$ which gives

$$\boxed{e^{tA} = Pe^{t\Lambda}P^{-1}.}$$

Remark 2.1.5. Alternatively, we can write down the general solution of (2.1.1) as

$$\begin{cases} x(t) = \displaystyle\sum_{j=1}^{n} c_j e^{\lambda_j t}\underline{\mathbf{e}}_j \\ x(0) = \displaystyle\sum_{j=1}^{n} c_j \underline{\mathbf{e}}_j = P\underline{\mathbf{c}}. \end{cases}$$

Define a fundamental matrix $\Psi(t) = Pe^{t\Lambda} = [e^{\lambda_1 t}\underline{\mathbf{e}}_1, \dots, e^{\lambda_n t}\underline{\mathbf{e}}_n]$, then

$$\begin{aligned} \underline{\mathbf{c}} &= P^{-1}x(0) = \Psi^{-1}(0)x(0) \\ \implies x(t) &= \Psi(t)\underline{\mathbf{c}} = \Psi(t)\Psi^{-1}(0)x(0). \end{aligned}$$

Conjugate Pair of Complex Eigenvalues

Suppose that A is a 2×2 matrix with a pair of complex conjugate eigenvalues $\rho \pm i\omega$. There exists a complex eigenvector $\underline{\mathbf{e}}_1 \in \mathbb{C}^2$ such that

$$A\underline{\mathbf{e}}_1 = (\rho + i\omega)\underline{\mathbf{e}}_1 \quad \text{and} \quad A^*\underline{\mathbf{e}}_1^* = (\rho - i\omega)\underline{\mathbf{e}}_1^*.$$

Looking at both real part and imaginary part of $A\underline{\mathbf{e}}_1$:

$$\begin{aligned} A\underline{\mathbf{e}}_1 &= A[\text{Re}(\underline{\mathbf{e}}_1) + i\text{Im}(\underline{\mathbf{e}}_1)] \\ A\underline{\mathbf{e}}_1 &= (\rho + i\omega)\underline{\mathbf{e}}_1 = \text{Re}[(\rho + i\omega)\underline{\mathbf{e}}_1] + i\text{Im}[(\rho + i\omega)\underline{\mathbf{e}}_1], \end{aligned}$$

which gives $A[\text{Re}(\underline{\mathbf{e}}_1)] = \text{Re}[(\rho + i\omega)\underline{\mathbf{e}}_1]$ and $A[\text{Im}(\underline{\mathbf{e}}_1)] = \text{Im}[(\rho + i\omega)\underline{\mathbf{e}}_1]$. Introducing $P = [\text{Im}(\underline{\mathbf{e}}_1), \text{Re}(\underline{\mathbf{e}}_1)]$, we see that

$$\begin{aligned} AP &= \Big[A[\text{Im}(\underline{\mathbf{e}}_1)], A[\text{Re}(\underline{\mathbf{e}}_1)]\Big] \\ &= \Big[\text{Im}[(\rho + i\omega)\underline{\mathbf{e}}_1], \text{Re}[(\rho + i\omega)\underline{\mathbf{e}}_1]\Big] \\ &= \Big[\rho\text{Im}(\underline{\mathbf{e}}_1) + \omega\text{Re}(\underline{\mathbf{e}}_1), \rho\text{Re}(\underline{\mathbf{e}}_1) - \omega\text{Im}(\underline{\mathbf{e}}_1)\Big] \\ &= P\begin{bmatrix} \rho & -\omega \\ \omega & \rho \end{bmatrix}, \text{ or } \Lambda = \begin{bmatrix} \rho & -\omega \\ \omega & \rho \end{bmatrix} = P^{-1}AP. \end{aligned}$$

As before, performing a change of variable $x = Py$ and using $\dot{x} = Ax$ gives

$$\dot{y} = \Lambda y \implies y(t) = e^{t\Lambda} y(0).$$

Decompose $\Lambda = D + C$, where

$$D = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}.$$

Since $DC = CD$, it follows that

$$e^{t\Lambda} = e^{tD}e^{tC} = \begin{bmatrix} e^{\rho t} & 0 \\ 0 & e^{\rho t} \end{bmatrix} \sum_{k=0}^{\infty} \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}^k \frac{1}{k!}.$$

Note that we have the following recurrence relation for C:

$$C^{2n} = (-1)^n \begin{bmatrix} \omega^{2n} & 0 \\ 0 & \omega^{2n} \end{bmatrix}, \quad C^{2n+1} = (-1)^n \begin{bmatrix} 0 & -\omega^{2n+1} \\ \omega^{2n+1} & 0 \end{bmatrix}$$

$$\implies e^{tC} = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix}.$$

Hence, $x(t) = Py(t) = Pe^{t\Lambda}y(0) = Pe^{t\Lambda}P^{-1}x(0)$ which gives

$$\boxed{e^{tA} = Pe^{\rho t} \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix} P^{-1}.}$$

Theorem 2.1.6. *Let $A \in \mathbb{R}^{n\times n}$ with*

(a) k distinct real eigenvalues $\lambda_1, \dots, \lambda_k$, and

(b) $m = \dfrac{n-k}{2}$ distinct complex conjugate eigenvalues $\rho_1 \pm i\omega_1, \dots, \rho_m \pm i\omega_m$.

There exists an invertible matrix P such that

$$\begin{cases} P^{-1}AP = \Lambda = \mathit{diag}\,(\lambda_1, \dots, \lambda_k, B_1, \dots, B_m) \\ B_j = \begin{bmatrix} \rho_j & -\omega_j \\ \omega_j & \rho_j \end{bmatrix}. \end{cases}$$

Moreover, $e^{tA} = Pe^{t\Lambda}P^{-1}$, with

$$\begin{cases} e^{t\Lambda} & = \mathit{diag}\,\left(e^{\lambda_1 t}, \dots, e^{\lambda_k t}, e^{B_1 t}, \dots, e^{B_m t}\right) \\ e^{B_j t} & = e^{\rho_j t} \begin{bmatrix} \cos(\omega_j(t)) & -\sin(\omega_j(t)) \\ \sin(\omega_j(t)) & \cos(\omega_j(t)) \end{bmatrix}. \end{cases}$$

Degenerate Eigenvalues

Suppose that $A \in \mathbb{R}^{n\times n}$ has p distinct eigenvalues $\lambda_1, \dots, \lambda_p$, with $p < n$. The corresponding characteristics polynomial of A is

$$\det\{(A - sI)\} = \prod_{k=1}^{p} (\lambda_k - s)^{n_k}.$$

where n_k is the **algebraic multiplicity** of λ_k. Another related quantity is the so called **geometric multiplicity**, which is defined to be $\dim(\mathcal{N}(A-\lambda_j I))$. The generalised eigenspace of λ_k is defined to be

$$E_k = \Big\{x \in \mathbb{R}^n \colon (A - \lambda_k I)^{n_k} x = \mathbf{0}\Big\}.$$

Associated with each degenerate eigenvalue are n_k linearly independent solutions of the form $P_1(t)e^{\lambda_k t}, \dots, P_{n_k}(t)e^{\lambda_k t}$, where $P_j(t)$ are vector polynomials of degree less than n_k.

Example 2.1.7. Suppose $A \in \mathbb{R}^{2\times 2}$ with a degenerate eigenvalue λ. Then $(A - \lambda I)^2 x = \mathbf{0}$ for all $x \in \mathbb{R}^2$ since A satisfies its own characteristics equation (Cayley-Hamilton theorem).

- Either $(A - \lambda I)x = \mathbf{0}$ for all $x \in \mathbb{R}^2 \implies A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$,
- or there exists a non-trivial $\underline{\mathbf{e}}_2 \neq \mathbf{0}$ such that $(A - \lambda I)\underline{\mathbf{e}}_2 \neq \mathbf{0}$.
  Define $\underline{\mathbf{e}}_1 = (A - \lambda I)\underline{\mathbf{e}}_2$, then $(A - \lambda I)\underline{\mathbf{e}}_1 = (A - \lambda I)^2\underline{\mathbf{e}}_2 = \mathbf{0}$. This implies that $\underline{\mathbf{e}}_1$ is an eigenvector of A with respect to λ.

  $$\begin{cases} A\underline{\mathbf{e}}_1 = \lambda\underline{\mathbf{e}}_1. \\ A\underline{\mathbf{e}}_2 = \underline{\mathbf{e}}_1 + \lambda\underline{\mathbf{e}}_2. \end{cases} \implies A\Big[\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2\Big] = \Big[\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2\Big] \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

  Thus, set $P = [\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2]$, we see that $P^{-1}AP = \Lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.
- In the transformed linear system for $y = P^{-1}x$, we have that $y = \Lambda y$, *i.e.*

  $$\begin{cases} \dot{y_1} &= \lambda y_1 + y_2 \\ \dot{y_2} &= \lambda y_2 \end{cases} \implies \begin{cases} y_1(t) = e^{\lambda t}[y_1(0) + y_2(0)t] \\ y_2(t) = e^{\lambda t} y_2(0) \end{cases}$$

2.2 Non-Autonomous Linear Systems

2.2.1 Homogeneous Equation

We study the homogeneous non-autonomous linear first order system

$$\dot{x}(t) = A(t)x(t), \quad x(t_0) = x_0 \in \mathbb{R}^n, \tag{2.2.1}$$

where coefficients of the matrix A are functions of t. Appealing to the 1D case, where $A(t)$ is a real-valued function of time, it follows from **method of integrating factors** that the solution of (2.2.1) is given by $x(t) = x_0 \exp\left(\int_0^t A(s)\,ds\right)$. Unfortunately, this result does not generalise to higher dimension; the problem lies on the fact that matrices do not commute in general, which implies that the property $e^{A(t)+B(t)} = e^{A(t)}e^{B(t)}$ fails to hold in this case. We begin by proving that the solution of (2.2.1) is unique, via an energy argument.

Theorem 2.2.1. *There exists at most one solution $x \in C^1([t_0, T]; \mathbb{R}^n)$ of (2.2.1).*

Proof. Suppose (2.2.1) has two solutions, and let $z := x - y$, then z satisfies (2.2.1) with homogeneous initial condition $z(t_0) = \mathbf{0}$. Taking the inner product of $\dot{z} = Az$ against z yields

$$z \cdot \dot{z} = z \cdot (A(t)z) \implies \frac{d}{dt}|z|^2 = 2z \cdot (A(t)z) \leq 2\|A(t)\|_F \|z\|^2.$$

Let $v(t) := |z(t)|^2 \in C^1([t_0, T]; \mathbb{R})$ and $a(t) := 2\|A(t)\|_F \in C^0([t_0, T]; \mathbb{R})$. The **method of integrating factors** then gives

$$\dot{v}(t) \leq a(t)v(t) \implies \frac{d}{dt}\left\{v(t)\exp\left(-\int_{t_0}^t a(s)\,ds\right)\right\} \leq 0.$$

Since $v(t) \geq 0$ and $v(t_0) = 0$, it follows that $v \equiv 0$ in $[t_0, T]$. ■

Definition 2.2.2. Let $(\psi_j(t))_{j=1}^n \in C^0(\mathbb{R}; \mathbb{R}^n)$ be a set of vector-valued functions, none of which are identically zero, *i.e.* each $\psi_j(t)$ has at least one non-trivial component. If there exists a set of scalars $(\alpha_j)_{j=1}^n$, not all zero, such that $\sum_{j=1}^n \alpha_j\psi_j(t) = \mathbf{0}$ for all $t \in \mathbb{R}$, then the set of vector-valued functions $(\psi_j)_{j=1}^n$ is said to be **linearly dependent**.

Theorem 2.2.3. *Any set of $(n+1)$ non-zero solutions of the system $\dot{x} = A(t)x$ is linearly dependent in $\mathbb{R}^n$.*

Proof. This is a non-trivial result which we will prove by exploiting the uniqueness property of solutions to the system $\dot{x} = A(t)x$. Consider any set of $(n+1)$ non-zero solutions $\psi_1(t), \dots, \psi_{n+1}(t)$. For a fixed time $t_0 \in \mathbb{R}$, the $(n+1)$ constant vectors $\psi_1(t_0), \dots \psi_{n+1}(t_0)$ are linearly dependent in $\mathbb{R}^n$, *i.e.* there exists constants $(\alpha_j)_{j=1}^{n+1}$ such that $\sum_{j=1}^{n+1} \alpha_j\psi_j(t_0) = \mathbf{0}$. Let $x(t) = \sum_{j=1}^{n+1} \alpha_j\psi_j(t)$, then $x(t_0) = \mathbf{0}$ and $\dot{x}(t) = A(t)x(t)$. But the trivial solution $x \equiv \mathbf{0}$ is a solution with initial condition $x(t_0) = \mathbf{0}$. It follows from Theorem 2.2.1 that $x(t) \equiv \mathbf{0}$ for all t, which implies that the set of (n+1) solutions $\{\psi_1(t), \dots, \psi_{n+1}(t)\}$ is linearly dependent. ■

Theorem 2.2.4. *There exists a set of n linearly independent solutions to the system $\dot{x} = A(t)x, x \in \mathbb{R}^n$.*

Proof. By existence theorem of ODEs, there exists a set of n solutions $\{\psi_1(t), \dots, \psi_n(t)\}$ corresponding to n initial conditions, $\psi_j(0) = \mathbf{e}_j$, $j = 1, \dots, n$, where $\mathbf{e}_j$'s are the canonical basis vectors in $\mathbb{R}^n$. We claim that since $\{\psi_1(0), \ldots, \psi_n(0)\}$ is a linearly independent set, so is the set $\{\psi_1(t), \ldots, \psi_n(t)\}$. Suppose not, by definition there exists scalars $(\alpha_j)_{j=1}^n$, not all zero, such that $\displaystyle\sum_{j=1}^n \alpha_j\psi_j(t) = \mathbf{0}$ for all t. In particular, we have that $\displaystyle\sum_{j=1}^n \alpha_j\psi_j(0) = \mathbf{0}$, which is a contradiction.

■

Corollary 2.2.5. *Let $\{\psi_1(t), \ldots, \psi_n(t)\}$ be any set of n linearly independent solutions of $\dot{x} = A(t)x, x \in \mathbb{R}^n$. Then every solution is a linear combination of $\{\psi_1(t), \ldots, \psi_n(t)\}$.*

Proof. For any non-trivial solution $\psi(t)$ of $\dot{x} = A(t)x$, the set $\{\psi(t), \psi_1(t), \ldots, \psi_n(t)\}$ is linearly dependent. Theorem 2.2.3 implies that $\psi(t) = \displaystyle\sum_{j=1}^n \alpha_j\psi_j(t)$.

■

Remark 2.2.6. Alternatively, we know that $\{\psi_j(t_0)\}_{j=1}^n$ is a basis for $\mathbb{R}^n$, so there exists scalars $(\alpha_j)_{j=1}^n$ such that $\psi(t_0) = \displaystyle\sum_{j=1}^n \alpha_j\psi_j(t_0)$. Observe that $\psi(t)$ and $\displaystyle\sum_{j=1}^n \alpha_j\psi_j(t)$ are both solutions with the same initial conditions. It follows from Theorem 2.2.1 that $\psi(t) = \displaystyle\sum_{j=1}^n \alpha_j\psi_j(t)$.

Definition 2.2.7. Let $\{\psi_1(t), \ldots, \psi_n(t)\}$ be n linearly independent solutions of $\dot{x} = A(t)x$, $x \in \mathbb{R}^n$. A **fundamental matrix** is defined to be

$$\Psi(t) = \Big[\psi_1(t), \ldots, \psi_n(t)\Big].$$

Remark 2.2.8. Fundamental matrix is not unique. Indeed, it follows from Corollary 2.2.5 that any 2 fundamental matrices Ψ_1, Ψ_2 are related by a non-singular constant matrix C, satisfying $\Psi_2(t) = \Psi_1(t)C$.

Theorem 2.2.9. *The solution of* (2.2.1) *is given by:*

$$x(t) = \Psi(t)\Psi^{-1}(t_0)x_0.$$

Proof. Choose a fundamental matrix $\Psi(t)$, it follows from Corollary 2.2.5 that the solution must be of the form $x(t) = \Psi(t)\underline{\mathbf{a}}$ for some constant vector $\underline{\mathbf{a}} \in \mathbb{R}^n$. Since $x_0 = \Psi(t_0)\underline{\mathbf{a}}$, it follows that $\underline{\mathbf{a}} = \Psi^{-1}(t_0)x_0$, where $\Psi(t_0)$ is invertible by linear independence of its columns. Hence, $x(t) = \Psi(t)\underline{\mathbf{a}} = \Psi(t)\Psi^{-1}(t_0)x_0$.

■

2.2.2 Inhomogeneous equation.

Having constructed a solution to the homogeneous problem, we consider the **inhomogeneous**, non-autonomous linear system

$$\dot{x}(t) = A(t)x(t) + f(t), \quad x(t_0) = x_0 \in \mathbb{R}^n. \tag{2.2.2}$$

Inspired by Theorem 2.2.9, we make an ansatz of the form

$$x(t) = \Psi(t)\Psi^{-1}(t_0)[x_0 + \phi(t)], \tag{2.2.3}$$

where $\phi(t)$ is a function to be determined. Note that $\phi(t_0) = \mathbf{0}$. A direct computation yields

$$\dot{x} = \dot{\Psi}(t)\Psi^{-1}(t_0)[x_0 + \phi(t)] + \Psi(t)\Psi^{-1}(t_0)[\dot{\phi}(t)]. \tag{2.2.4}$$

On the other hand, substituting the ansatz (2.2.3) into (2.2.2) yields

$$\dot{x} = A(t)\Big[\Psi(t)\Psi^{-1}(t_0)[x_0 + \phi(t)]\Big] + f(t) = \dot{\Psi}(t)\Psi^{-1}(t_0)[x_0 + \phi(t)] + f(t), \tag{2.2.5}$$

since $\dot{\Psi}(t) = A(t)\Psi(t)$. Comparing (2.2.4) and (2.2.5), we see that

$$\begin{aligned} \Psi(t)\Psi^{-1}(t_0)\dot{\phi}(t) = f(t) &\implies \dot{\phi}(t) = \Psi(t_0)\Psi^{-1}(t)f(t) \\ &\implies \phi(t) = \Psi(t_0)\int_{t_0}^{t} \Psi^{-1}(s)f(s)\,ds. \end{aligned}$$

Hence, we have the following formula for a solution of (2.2.4)

$$\boxed{x(t) = \Psi(t)\Psi^{-1}(t_0)x_0 + \Psi(t)\int_{t_0}^{t} \Psi^{-1}(s)f(s)\,ds.}$$

Example 2.2.10. Consider the following inhomogeneous, non-autonomous linear system

$$\begin{cases} \dot{x_1} &= x_2 + e^t \\ \dot{x_2} &= x_1 \\ \dot{x_3} &= te^{-t}(x_1 + x_2) + x_3 + 1 \end{cases}, \quad \text{with } x_0 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

We see that

$$A(t) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ te^{-t} & te^{-t} & 1 \end{bmatrix}, \quad f(t) = \begin{bmatrix} e^t \\ 0 \\ 1 \end{bmatrix}$$

- Consider the homogeneous equation:

$$\begin{cases} \dot{x_1} &= x_2 \\ \dot{x_2} &= x_1 \\ \dot{x_3} &= te^{-t}(x_1 + x_2) + x_3 \end{cases}$$

Differentiating $\dot{x_1}$ with respect to t gives $\ddot{x_1} = \dot{x_2} = x_1 \implies x_1(t) = e^t, e^{-t}, 0$.

(i) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \dot{x_3} - x_3 = 2t$, so that $x_3(t) = Ce^t - 2(1+t)$.

(ii) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \implies \dot{x_3} - x_3 = 0$, so that $x_3(t) = De^t$.

(iii) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \dot{x_3} - x_3 = 0$, so that $x_3(t) = Ee^t$.

- A fundamental solution matrix is constructed by choosing constants C, D, E in each cases; we want to avoid trivial solutions too. Remember, it doesn't matter how we choose these constants because any two fundamental matrices can be related by a non-singular constant matrix.

$$\Psi(t) = \begin{bmatrix} e^t & e^{-t} & 0 \\ e^t & -e^{-t} & 0 \\ -2(1+t) & 0 & e^t \end{bmatrix}$$

Computing its inverse:

$$\begin{aligned} \Psi^{-1}(t) &= \frac{1}{-2e^t} \begin{bmatrix} -1 & -1 & 0 \\ -e^{2t} & e^{2t} & 0 \\ -2(1+t)e^{-t} & -2(1+t)e^{-t} & -2 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} e^{-t} & e^{-t} & 0 \\ e^t & -e^t & 0 \\ 2(1+t)e^{-2t} & 2(1+t)e^{-2t} & 2e^{-t} \end{bmatrix} \end{aligned}$$

Evaluating $\Psi^{-1}(t_0)$ at $t_0 = 0$:

$$\Psi^{-1}(t_0)\Big|_{t_0=0} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & 2 & 2 \end{bmatrix}$$

- Hence, substituting everything into the formula gives us the following solution:

$$\begin{cases} x_1(t) &= \left(\dfrac{3}{4} + \dfrac{1}{2}t\right) e^t - \dfrac{3}{4}e^{-t} \\ x_2(t) &= \left(\dfrac{1}{4} + \dfrac{1}{2}t\right) e^t - \dfrac{3}{4}e^{-t} \\ x_3(t) &= 3e^t - t^2 - 3t - 4 \end{cases}$$

2.2.3 Stability and Bounded Sets

As already mentioned, one of the key questions regarding the long term behaviour of a dynamical system is stability of its solutions. **There is no single concept of stability, in fact**

various different definitions are possible.

Definition 2.2.11. Consider the ODE $\dot{x} = f(x,t)$, $x \in \mathbb{R}^n$. Let $\phi(x,t)$ be the flow of f, with $\phi(x,t_0) = x_0$.

(a) $\phi(x,t)$ is said to be **Lyapunov stable** for $t \geq t_0$ if and only if for all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon, t_0) > 0$ such that

$$\|y_0 - x_0\| < \delta \implies \|\phi(y,t) - \phi(x,t)\| < \varepsilon \quad \text{for all } t \geq t_0,$$

where $\phi(y,t)$ represents any other neighbouring flow.

- For an autonomous system, stability is independent of t_0. Thus a solution is either Lyapunov stable or unstable for all t_0.

(b) $\phi(x,t)$ is said to be **asymptotically stable** for $t \geq t_0$ if $\phi(x,t)$ is Lyapunov stable for $t \geq t_0$ and in addition, there exists $\eta(t_0) > 0$ such that

$$\|y_0 - x_0\| < \eta \implies \lim_{t\to\infty} \|\phi(y,t) - \phi(x,t)\| = 0.$$

- η might be smaller than δ.

Theorem 2.2.12. *For the regular linear system $\dot{x} = A(t)x$, the zero solution $x^*(t) \equiv \mathbf{0}$ is Lyapunov stable on $t \geq t_0$ (t_0 arbitrary) if and only if every solution is bounded as $t \longrightarrow \infty$.*

Proof. Suppose that the zero solution $x^*(t) \equiv \mathbf{0}$ is Lyapunov stable. By definition, there exists $\delta > 0$ such that $\|x(t_0)\| < \delta \implies \|x(t)\| < \varepsilon$ for all $t \geq t_0$. Consider the fundamental matrix $\Psi(t) = \Big[\psi_1(t), \dots, \psi_n(t)\Big]$ satisfying the initial condition $\Psi(t_0) = \dfrac{\delta}{2}I$. From Corollary 2.2.5, we know that any solution (with arbitrary initial conditions) can be written as $x(t) = \Psi(t)\underline{\mathbf{c}}$. For any $j = 1, \dots, n$, since $\|\psi_j(t_0)\| = \delta/2 < \delta$, Lyapunov stability implies that $\|\psi_j(t)\| < \varepsilon$. Thus,

$$\|x(t)\| = \|\Psi(t)\underline{\mathbf{c}}\| = \left\|\sum_{j=1}^n c_j\psi_j(t)\right\| \leq \sum_{j=1}^n |c_j|\|\psi_j(t)\| \leq \varepsilon\sum_{j=1}^n |c_j| < \infty.$$

Conversely, suppose that every solution of $\dot{x} = A(t)x$ is bounded. Let $\Psi(t)$ be any fundamental matrix. Boundedness of solutions implies that there exists $M > 0$ such that $\|\Psi(t)\| < M$ for all $t \geq t_0$. Given any $\varepsilon > 0$, choose $\delta = \dfrac{\varepsilon}{M\|\Psi^{-1}(x_0)\|} > 0$. Theorem 2.2.9 states that any solution has the form $x(t) = \Psi(t)\Psi^{-1}(t_0)x_0$. Thus, for $\|x(t_0)\| < \delta$ we have

$$\|x(t)\| \leq \|\Psi(t)\|\|\Psi^{-1}(t_0)\|\|x(t_0)\| \leq M\|\Psi^{-1}(t_0)\|\delta = \varepsilon.$$

■

Theorem 2.2.13. *All solutions of the inhomogeneous linear system $\dot{x} = A(t)x(t) + f(t)$ have the same Lyapunov stability property as the zero solution of homogeneous linear system $\dot{y} = A(t)y(t)$.*

Proof. Let $x^*(t)$ be a solution of the inhomogeneous equation, whose stability we wish to determine. Let $x(t)$ be any another solution, and set $y(t) = x(t) - x^*(t)$. It follows that

$$\begin{cases} \dot{y}(t) & = A(t)y(t) \\ y(t_0) & = x(t_0) - x^*(t_0). \end{cases}$$

Lyapunov stability of $x^*(t)$ means that for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|x(t_0) - x^*(t_0)\| < \delta \implies \|x(t) - x^*(t)\| < \varepsilon \quad \text{for all } t \geq t_0.$$

In terms of y, this is equivalent to $\|y(t_0)\| < \delta \implies \|y(t)\| < \varepsilon$, which is the condition for Lyapunov stability of the zero solution.

■

2.2.4 Equations With Coefficients That Have A Limit

Consider the equation

$$\dot{x} = Ax + B(t)x, \quad x \in \mathbb{R}^n, \tag{2.2.6}$$

where $A \in \mathbb{R}^{n\times n}$ is non-singular and $B(t)$ is continuous as a function of time. If $\lim_{t\to\infty} \|B(t)\| = 0$, then we might expect that solutions of (2.2.6) converges to solutions of $\dot{x} = Ax$, but this is not true even in the one-dimensional case! This is a glimpse that non-autonomous linear systems are dangerous!

Example 2.2.14. Consider $\ddot{x} - \dfrac{2}{t}\dot{x} + x = 0$, $t \geq 1$. There are 2 linearly independent solutions of the form:

$$\begin{cases} \sin(t) - t\cos(t) \\ \cos(t) + t\sin(t). \end{cases}$$

These are unbounded as $t \longrightarrow \infty$, whereas solutions of $\ddot{x} + x = 0$ are bounded.

If we view (2.2.6) as a special case of $\dot{x} = C(t)x$, where $C(t) = A + B(t)$ is in some sense a small perturbation away from the constant matrix A, then we have the following positive result.

Theorem 2.2.15. *Consider the non-autonomous linear system (2.2.6) with $B(t)$ continuous for $t \geq t_0$ and*

(a) the eigenvalues of A satisfy $Re(\lambda_j) \leq 0$, $j = 1, \ldots, n$,

(b) the eigenvalues of A for which $Re(\lambda_j) = 0$ are distinct, i.e. there are no degenerate pairs of complex eigenvalue satisfying $Re(\lambda_j) = 0$,

(c) $\displaystyle\int_{t_0}^{\infty} \|B(t)\|\, dt$ is bounded.

Then solutions of (2.2.6) are bounded and the zero solution $x(t) \equiv \mathbf{0}$ is Lyapunov stable.

Proof. We use the **variation of parameters** method. Consider the ansatz $x(t) := \Psi(t)z(t)$, where $\Psi(t)$ is the fundamental matrix of $\dot{x} = Ax$ with $\Psi(t_0) = I$, *i.e.* "formally speaking" $\Psi(t) = e^{(t-t_0)A}$. A direct computation using Chain rule yields

$$\dot{x} = \dot{\Psi}(t)z(t) + \Psi(t)\dot{z}(t) = A\Psi(t)z(t) + \Psi(t)\dot{z}(t), \tag{2.2.7}$$

since $\dot{\Psi}(t) = A\Psi(t)$. On the other hand, substituting the ansatz into (2.2.6) yields

$$\dot{x} = Ax + B(t)x = A\Psi(t)z(t) + B(t)\Psi(t)z(t). \tag{2.2.8}$$

Comparing (2.2.7) and (2.2.8), we see that

$$\begin{aligned}\Psi(t)\dot{z}(t) = B(t)\Psi(t)z(t) &\implies \dot{z}(t) = \Psi^{-1}(t)B(t)\Psi(t)z(t)\\ &\implies z(t) = z(t_0) + \int_{t_0}^{t} \Psi^{-1}(s)B(s)\Psi(s)z(s)\,ds\end{aligned}$$

Therefore,

$$\begin{aligned}x(t) = \Psi(t)z(t) &= \Psi(t)\left(z(t_0) + \int_{t_0}^{t} \Psi^{-1}(s)B(s)\Psi(s)z(s)\,ds\right)\\ &= \Psi(t)z(t_0) + \int_{t_0}^{t} \Big[\Psi(t)\Psi^{-1}(s)\Big]B(s)x(s)\,ds\end{aligned}$$

Note the following:

$$\begin{aligned}x(t_0) = \Psi(t_0)z(t_0) &= z(t_0) && \Big[\text{since } \Psi(t_0) = I \text{ by construction}\Big]\\ \Psi(t)\Psi^{-1}(s) &= e^{tA}\Psi(t_0)e^{-sA}\Psi(t_0) && \Big[\text{from Theorem 2.2.9}\Big]\\ &= e^{A(t-s)} && \Big[\text{since } A \text{ commutes with itself}\Big]\\ &= \Psi(t-s).\end{aligned}$$

Thus,

$$\begin{aligned}\|x(t)\| &= \left\|\Psi(t)x_0 + \int_{t_0}^{t} \Psi(t-s)B(s)x(s)\,ds\right\|\\ &\leq \|\Psi(t)\|\|x_0\| + \int_{t_0}^{t} \|\Psi(t-s)\|\|B(s)\|\|x(s)\|\,ds\\ &\leq C\|x_0\| + \int_{t_0}^{t} C\|B(s)\|\|x(s)\|\,ds\end{aligned}$$

where $\|\Psi(t)\|$ is bounded for all $t \geq t_0$, since $\mathrm{Re}(\lambda_j) \leq 0$ for all $j = 1, \ldots, n$ by assumption.

Referring to Theorem 1.3.6, applying Gronwall's inequality with $v(t) = \|x(t)\|$, $c(t) = C\|x_0\|$, and $u(t) = C\|B(t)\|$ gives

$$\|x(t)\| \leq C\|x_0\| \exp\left(C\int_{t_0}^{t} \|B(s)\|\,ds\right) < +\infty,$$

since $\displaystyle\int_{t_0}^{t} \|B(s)\|\,ds < \infty$ by assumption (c). Hence, $x(t)$ is bounded and by Theorem 2.2.12, $x \equiv \mathbf{0}$ is Lyapunov stable.

∎

Remark 2.2.16. Note that $\text{Re}(\lambda_j) \leq 0$ for all $j = 1, \ldots, n$ is not sufficient to establish boundedness of all solutions of (2.2.6).

Theorem 2.2.17. *Consider the non-autonomous linear system (2.2.6) with $B(t)$ continuous for all $t \geq t_0$ and*

(a) the eigenvalues of A satisfy $Re(\lambda_j) < 0$ for all $j = 1, \ldots, n$,

(b) $\lim_{t\to\infty} \|B(t)\| = 0$.

Then we have that $\lim_{t\to\infty} x(t) = \mathbf{0}$, and the zero solution $x(t) \equiv \mathbf{0}$ is asymptotically stable.

2.3 Floquet Theory

In this section, we consider a special case of the non-autonomous linear system

$$\dot{x} = P(t)x, \; x \in \mathbb{R}^n, \text{ where } P(t+T) = P(t) \text{ for all } t \in \mathbb{R}, \tag{2.3.1}$$

i.e. $P(\cdot)$ is a T-periodic continuous matrix-valued function. Such equation arises when linearising about a limit cycle solution on $\dot{x} = f(x)$. Observe that if $x(t)$ is a solution of (2.3.1), then periodicity of P implies that $x(t+T)$ is again a solution of (2.3.1). However, it does not say that $x(t)$ is periodic!

2.3.1 Floquet Multipliers

Theorem 2.3.1 (Floquet)**.** *The regular system (2.3.1) where $P(t)$ is an $n \times n$ time-dependent matrix with period $T > 0$ has at least 1 non-trivial solution $\chi(t)$, satisfying*

$$\chi(t+T) = \mu\chi(t), \quad t \in (-\infty, \infty),$$

*where μ is called a **Floquet multiplier**.*

Proof. Let $\Psi(t) = \Big[\psi_1(t), \ldots, \psi_n(t)\Big]$ be a fundamental matrix of (2.3.1). It satisfies

$$\dot{\Psi}(t) = P(t)\Psi(t) \text{ and } \dot{\Psi}(t+T) = P(t)\Psi(t+T), \text{ since } P(\cdot) \text{ is } T-\text{periodic}.$$

Thus, $\Phi(t) \colon = \Psi(t+T)$ is also a fundamental matrix of (2.3.1). Theorem 2.2.5 states that $\phi_j(t)$ is a linear combination of $\{\psi_1(t), \ldots, \psi_n(t)\}$, *i.e.* there exists constant vectors $e_j \in \mathbb{R}^n$ such that $\phi_j(t) = \Psi(t)e_j$ for every $j = 1, \ldots, n$. We can now rewrite $\Phi(t)$ as

$$\Phi(t) = \Psi(t+T) = \Psi(t)E, \tag{2.3.2}$$

where $E = \Big[e_1, \ldots, e_n\Big] \in \mathbb{R}^{n\times n}$ is non-singular. Consider any eigenvalue μ of E with its corresponding eigenvector $v \neq \mathbf{0}$, *i.e.* $Ev = \mu v$. Setting $\chi(t) = \Psi(t)v$, we see that $\chi(t) \neq \mathbf{0}$ since $\Psi(t)$ is non-singular and $v \neq \mathbf{0}$. Moreover,

$$\chi(t+T) = \Psi(t+T)v = \Psi(t)Ev = \mu\Psi(t)v = \mu\chi(t).$$

∎

Remark 2.3.2. In the proof, we exploit the intrinsic periodic structure of $\Psi(t)$ due to periodicity of $P(t)$. This theorem also suggest a positive result: there exists a periodic solution to (2.3.1) if the Floquet multiplier $\mu = 1$.

A natural question stems out from the previous theorem: Does μ affected by the choice of fundamental matrix of (2.3.1)? This is answered in the next theorem, which states that Floquet multipliers are intrinsic property of the periodic system (2.3.1).

Theorem 2.3.3. *The Floquet multipliers μ are independent of the choice of fundamental matrix of* (2.3.1)*.*

Proof. Let $\Psi(t)$ and $\Psi^*(t)$ be any 2 fundamental matrices of (2.3.1). Theorem 2.2.5 gives the relation $\Psi^*(t) = \Psi(t)A$, where $A \in \mathbb{R}^{n\times n}$ is some non-singular matrix. Using (2.3.2) from Theorem 2.3.1, we have that

$$\Psi^*(t+T) = \Psi(t+T)A = \Psi(t)EA = \Psi^*(t)A^{-1}EA = \Psi^*(t)E^*.$$

Since E and E^* are related by a similarity transformation, they have the same eigenvalues. More precisely,

$$\begin{aligned}
\det(E^* - \mu I) &= \det\big(A^{-1}EA - \mu A^{-1}A\big) \\
&= \det\big(A^{-1}(E-\mu I)A\big) \\
&= \det\big(A^{-1}\big)\det(E-\mu I)\det(A) \\
&= \det(E-\mu I).
\end{aligned}$$

∎

Remark 2.3.4. One often choose $\Psi(t)$ satisfying $\Psi(0) = I$ if possible, so that

$$\Psi(0+T) = \Psi(0)E = E, \quad i.e. \quad \Psi(T) = E.$$

Definition 2.3.5. Let μ be a Floquet multiplier of (2.3.1). A **Floquet exponent** $\rho \in \mathbb{C}$ is a complex number such that $\mu = e^{\rho T}$.

- Note that Floquet exponents are not unique, since ρ is defined up to an added multiple of $\dfrac{2\pi i m}{T}, m \in \mathbb{Z}$ due to periodicity of the complex exponential function e^{iz}. Thus, we usually make the restriction $-\pi < \operatorname{Im}(\rho T) < \pi$.

Theorem 2.3.6. *Suppose that the matrix E, for which $\Psi(t+T) = \Psi(t)E$ has n distinct eigenvalues $(\mu_j)_{j=1}^n$, not necessarily real. The periodic system* (2.3.1) *has n linearly independent solutions of the form*

$$\chi_j(t) = p_j(t)e^{\rho_j t}, \ \text{with } p_j(t) \text{ being } T-\text{periodic}. \tag{2.3.3}$$

Proof. From Floquet theorem 2.3.1, for every $\mu_j = e^{\rho_j T}$, $j = 1, \dots, n$, there exists a non-trivial solution $\chi_j(t)$ such that

$$\begin{aligned}\chi_j(t+T) = \mu_j \chi_j(t) = e^{\rho_j T}\chi_j(t) &\implies \chi_j(t+T)e^{-\rho_j T} = \chi_j(t)\\ &\implies \chi_j(t+T)e^{-\rho_j(t+T)} = e^{-\rho_j t}\chi_j(t).\end{aligned}$$

Setting $p_j(t) = \chi_j(t)e^{-\rho_j t}$, we see that $p_j(t)$ is T-periodic.

Recall that in the proof of Floquet theorem 2.3.1, we write $\chi_j(t) = \Psi(t)v_j$, where $v_j \neq \mathbf{0}$ is an eigenvector corresponding to eigenvalue μ_j. Since distinct eigenvalues implies linearly independent set of eigenvectors and $\Psi(t)$ is by definition a non-singular matrix, we conclude that the set of solutions $\{\chi_1(t), \dots, \chi_n(t)\}$ is linearly independent in $\mathbb{R}^n$.

■

Remark 2.3.7. Note that existence of solutions of the form (2.3.3) continues to hold without the assumption that E has distinct eigenvalues.

Example 2.3.8. Consider the periodic system

$$\begin{bmatrix}\dot{x_1}\\ \dot{x_2}\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & h(t)\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}, \quad h(t) = \frac{\cos(t)+\sin(t)}{2+\sin(t)-\cos(t)},$$

where $h(\cdot)$ has period $T = 2\pi$.

- x_2 can be solved explicitly. Setting $f(t) := 2 + \sin(t) - \cos(t)$, we see that

$$\dot{x_2}(t) = \frac{\dot{f}(t)}{f(t)}x_2(t) \implies f(t)\dot{x_2}(t) - x_2(t)\dot{f}(t) = 0 \implies \frac{d}{dt}\left(\frac{x_2(t)}{f(t)}\right) = 0.$$

  The solution is given by $x_2(t) = bf(t) = b[2+\sin(t)-\cos(t)]$.

- We can now solve for x_1. Using **method of integrating factors**,

$$\begin{aligned}\dot{x_1} - x_1 &= b[2+\sin(t)-\cos(t)]\\ \frac{d}{dt}\left[e^{-t}x_1\right] &= be^{-t}[2+\sin(t)-\cos(t)]\\ \frac{d}{dt}\left[e^{-t}x_1\right] &= -\frac{d}{dt}\left[be^{-t}(2+\sin(t))\right]\\ e^{-t}x_1 &= a - be^{-t}[2+\sin(t)]\\ x_1(t) &= ae^t - b[2+\sin(t)].\end{aligned}$$

- Next, we construct a fundamental matrix $\Psi(t)$ by choosing constants a, b. Setting $(a,b) = (0,1)$ and $(a,b) = (1,0)$, we obtain two linearly independent solutions and so

$$\Psi(t) = \begin{bmatrix}-2-\sin(t) & e^t\\ 2+\sin(t)-\cos(t) & 0\end{bmatrix}$$

Now, introduce the non-singular matrix E such that $\Psi(t+2\pi) = \Psi(t)E$ for any $t \in \mathbb{R}$. In particular,

$$E = \Psi^{-1}(0)\Psi(2\pi) = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} -2 & e^{2\pi} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi} \end{bmatrix}$$

which gives $\mu_1 = 1 = e^{2\pi\rho_1}$ and $\mu_2 = e^{2\pi} = e^{2\pi\rho_2}$, or equivalently $\rho_1 = 0$ and $\rho_2 = 1$.

- Hence, the general solution is given by

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= b \begin{bmatrix} -2-\sin(t) \\ 2+\sin(t)-\cos(t) \end{bmatrix} + a \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t \\ &= bp_1(t)e^{\rho_1 t} + ap_2(t)e^{\rho_2 t} \\ &= bp_1(t) + ap_2(t)e^t. \end{aligned}$$

Note that p_1, p_2 are 2π-periodic, but $x(t)$ is not even periodic.

Remark 2.3.9. Observe that fixing a and b is the same as fixing initial conditions. In general, the linear combination $(a,b) = (1,0)$ and $(a,b) = (0,1)$ almost always gives two linearly independent solutions of the system.

2.3.2 Stability of Limit Cycles

As already mentioned, the periodic system (2.3.1) arises when one linearises about a limit cycle on the nonlinear autonomous system $\dot{x} = f(x), x \in \mathbb{R}^n$. Suppose $u(t)$ is such a limit cycle solution with period T. It follows that $\dot{u}(t) = f(u(t))$. Differentiating with respect to t gives

$$\ddot{u} = Df(u(t))\dot{u},$$

where $Df(u(t))$ is the Jacobian evaluated at the limit cycle solution. Observe that this equation implies that $Df(u(t))$ is T-periodic, since $u(t)$ is T-periodic (which means both $\dot{u}$ and $\ddot{u}$ are both T-periodic too).

On the other hand, setting $v(t) \coloneqq x(t) - u(t)$ and linearising yields

$$\begin{aligned} \dot{v}(t) &= \dot{x}(t) - \dot{u}(t) \\ &= f(x(t)) - f(u(t)) \\ &\approx Df(u(t))[x(t) - u(t)] \\ &= Df(u(t))v(t) \\ &= P(t)v(t). \end{aligned}$$

Hence, one possible solution of $\dot{v} = P(t)v$ is $v(t) = \dot{u}(t)$. Since u and $\dot{u}$ are both T-periodic, it follows that there exists a Floquet multiplier $\mu = 1$. To see this, recall from Theorem 2.2.9 that the solution of $\dot{v} = P(t)v(t)$ is given by

$$v(t) = \Psi(t)\Psi^{-1}(0)v(0), \tag{2.3.4}$$

where $\Psi(t)$ is any fundamental matrix of the periodic system. In particular, we choose a fundamental matrix $\Psi(t)$ such that $\Psi(0) = I$, it satisfies $\Psi(T) = E$ from Remark 2.3.4. Substituting $v(t) = \dot{u}(t)$ into (2.3.4) yields

$$\dot{u}(t) = \Psi(t)\Psi^{-1}(0)\dot{u}(0) \implies \dot{u}(T) = \dot{u}(0) = \Psi(T)\dot{u}(0).$$

It follows that E has an eigenvalue $\mu = 1$, which is the desired unit Floquet multiplier.

Remark 2.3.10.

1. We need to look at the linearised system to relate to the concept we discussed before.
2. $\dot{u}(t)$ is the vector tangential to the limit cycle at time t. Write $\dot{u} = \Psi(t)v$, where v is an eigenvector corresponding to $\mu = 1$. Assuming that $\Psi(0) = I$, we see that $\underline{\mathbf{a}}$ is tangential to the limit cycle at time $t = 0$, which is consistent with what we showed above.
3. The existence of an unit Floquet multiplier $\mu = 1$ reflects the time or phase-shift invariance of an autonomous system, *i.e.* phase shift around the limit cycle.
4. The limit cycle is linearly stable (up to small perturbations) provided that the other $n-1$ Floquet multipliers lie inside the unit circle in the complex plane $\mathbb{C}$.

Floquet theory is such a powerful tool in periodic dynamical system that computing Floquet multipliers becomes an important subject itself. Theoretically, one has to first construct a fundamental matrix of the system, which is a daunting task itself since this amounts to solving the periodic system. Fortunately, the periodic matrix $P(t)$ allows us to extract some information about Floquet multipliers, as we shall see in Theorem 2.3.13.

Definition 2.3.11. Let $\Psi(t) = \left[\psi_1(t), \dots, \psi_n(t)\right]$ be a fundamental matrix satisfying $\dot{\psi}_j(t) = A(t)\psi_j(t)$. The **Wronskian** of $\Psi(t)$ is defined as $W(t) = \det(\Psi(t))$.

Theorem 2.3.12 (Liouville/Abel Identity). *Let $\Psi(t)$ be any fundamental matrix of the non-autonomous linear system $\dot{x} = A(t)x$. For all initial time t_0, we have the following expression for the Wronskian of $\Psi(t)$:*

$$\boxed{W(t) = W(t_0)\exp\left(\int_{t_0}^{t} tr(A(s))\,ds\right).}$$

Proof. Using **Leibniz rule of calculus**,

$$\frac{d}{dt}[W(t)] = \sum_{k=1}^{n} \Delta_k(t)$$

where $\Delta_k(t)$ is $W(t)$ with k-th row replaced by $\dot{\psi}_{jk}$ instead of ψ_{jk}.

$$\Delta_1 = \det\begin{bmatrix} \dot{\psi}_{11} & \cdots & \dot{\psi}_{n1} \\ \psi_{12} & \cdots & \psi_{n2} \\ \vdots & \ddots & \vdots \\ \psi_{1n} & \cdots & \psi_{nn} \end{bmatrix}$$

$$= \det \begin{bmatrix} \sum_{k=1}^{n} A_{1k}\psi_{1k} & \cdots & \sum_{k=1}^{n} A_{1k}\psi_{nk} \\ \psi_{12} & \cdots & \psi_{n2} \\ \vdots & \vdots & \vdots \\ \psi_{1n} & \cdots & \psi_{nn} \end{bmatrix}$$

$$= \sum_{k=1}^{n} A_{1k} \det \begin{bmatrix} \psi_{1k} & \cdots & \psi_{nk} \\ \psi_{12} & \cdots & \psi_{n2} \\ \vdots & \vdots & \vdots \\ \psi_{1n} & \cdots & \psi_{nn} \end{bmatrix}$$

$$= A_{11} W(t)$$

since we have 2 identical rows for every $k \neq 1$ which results in zero determinant. Therefore,

$$\frac{d}{dt}[W(t)] = \sum_{k=1}^{n} A_{kk} W(t) = \mathrm{Tr}(A(t)) W(t)$$

$$\implies W(t) = W(t_0) \exp\left(\int_{t_0}^{t} \mathrm{Tr}(A(s))\, ds \right).$$

■

Theorem 2.3.13. *The Floquet multipliers of a periodic system of the form* (2.3.1) *satisfy*

$$\boxed{\prod_{j=1}^{n} \mu_j = \exp\left(\int_0^T tr(P(s))\, ds \right).}$$

Proof. Let $\Psi(t)$ be the fundamental matrix of (2.3.1) satisfying $\Psi(0) = I$, so that $E = \Psi(T)$. We know from Floquet Theorem 2.3.1 that the Floquet multipliers μ_j satisfy the characteristics equation:

$$0 = \det(E - \mu I) = \prod_{j=1}^{n} (\mu_j - \mu).$$

By Theorem 2.3.12, we have that (using $W(0) = 1$)

$$\prod_{j=1}^{n} \mu_j = \det(E) = \det(\Psi(T)) = W(T) = \exp\left(\int_0^T \mathrm{tr}(P(s))\, ds \right).$$

■

Example 2.3.14. Consider a nonlinear damping oscillator given by $\ddot{x} + f(x)\dot{x} + g(x) = 0$. Suppose that it has a T-periodic solution $\phi(t)$. Rewrite equation as

$$\begin{cases} \dot{x} & = y \\ \dot{y} & = -f(x)y - g(x). \end{cases}$$

Linearising about $\phi(t)$ gives

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -Df(\phi(t))y - Dg(\phi(t)) & -f(\phi(t)) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = P(t) \begin{bmatrix} x \\ y \end{bmatrix}$$

where x, y are now some small perturbations about ϕ. Since we must have $\mu_1 = 1$,

$$\mu_2 = \exp\left(\int_0^T \operatorname{tr}(P(s))\, ds\right) = \exp\left(-\int_0^T f(\phi(s))\, ds\right).$$

Thus, $\phi(t)$ is linearly stable provided that $\displaystyle\int_0^T f(\phi(s))\, ds \geq 0$.

2.3.3 Mathieu Equation

Consider the Mathieu equation

$$\ddot{x} + [\alpha + \beta\cos(t)]x = 0.$$

This can be written as $\dot{x} = P(t)x, x \in \mathbb{R}^2$, where $P(t) = \begin{bmatrix} 0 & 1 \\ -\alpha - \beta\cos(t) & 0 \end{bmatrix}$. There is no explicit formula for the Floquet multipliers for all $\alpha, \beta \in \mathbb{R}$. However, since $\operatorname{tr}(P(t)) = 0$, it follows that the Floquet multipliers satisfy $\mu_1\mu_2 = 1$. Hence, μ_1, μ_2 are solutions of the quadratic characteristic equation

$$\mu^2 - \phi(\alpha, \beta)\mu + 1 = 0, \quad \text{with roots } \mu_{1,2} = \frac{1}{2}\left(\phi \pm \sqrt{\phi^2 - 4}\right).$$

In principle, $\phi(\alpha, \beta)$ can be determined for particular α and β. Although we have no explicit formula for $\phi(\alpha, \beta)$, we can still deduce the behaviour of solutions based on values of ϕ.

1. **$\boldsymbol{\phi > 2}$, so that $\boldsymbol{\mu_{1,2}}$ are real, positive and distinct.**

   - Setting $\mu_{1,2} = e^{\pm\sigma 2\pi}$, $\sigma > 0$, we can write the general solution as:

     $$x(t) = C_1 p_1(t) e^{\sigma t} + C_2 p_2(t) e^{-\sigma t},$$

     where $p_j(t)$ are 2π-periodic.
   - Solution is unbounded.

2. **$\boldsymbol{\phi = 2}$ (degenerate case), so that $\boldsymbol{\mu_1 = \mu_2 = 1}$.**

   - We have that $\rho_1 = \rho_2 = 0$, so the general solution is given by

     $$x(t) = C_1 p_1(t) + C_2 p_2(t),$$

     where $p_j(t)$ are 2π-periodic.

- There is one solution of period 2π that is stable. It can be shown that the other solution is unstable.

3. **$-2 < \phi < 2$, so that $\mu_{1,2}$ form a complex conjugate pair.**

- Since $\mu_1\mu_2 = 1$, we must have $\rho_{1,2} = \pm i\nu$, $\nu \in \mathbb{R}$, otherwise $e^{a+i\nu}e^{a-i\nu} = e^{2a} \neq 1$. Thus, the general solution is given by

$$x(t) = C_1p_1(t)e^{i\nu t} + C_2p_2(t)e^{-i\nu t}.$$

where $p_j(t)$ are 2π-periodic.

- Solution is bounded and oscillatory, but is not periodic in general as there are 2 frequencies 2π and ν. Solution is called **quasiperiodic** if $\nu/2\pi$ is irrational.
- Circle map: $\theta_{n+1} = \theta_n + \tau$.
  - If τ is rational, then eventually the trajectory becomes a periodic orbit.
  - If τ is irrational, we have a dense orbit instead.

4. **$\phi = -2$, so that $\mu_1 = \mu_2 = -1$.**

- $\mu_{1,2} = -1 = e^{(\rho_{1,2})2\pi} \implies \rho_{1,2} = i/2$, so the general solution as

$$x(t) = C_1p_1(t)e^{it/2} + C_2p_2(t)e^{it/2},$$

where $p_j(t)$ are 2π-periodic.

- There exists a 4π-periodic solution, and the other solution is unbounded.

5. **$\phi < -2$, so that $\mu_{1,2}$ are real, negative and distinct.**

- Observe that $\mu_{1,2} = 1$ implies that $\rho_1 = -\rho_2$. To enforce $\mu_{1,2}$ to take negative value, $\mu_{1,2}$ must have the form

$$\mu_{1,2} = \exp\left\{\left(\pm\sigma + \frac{i}{2}\right)2\pi\right\}, \quad \text{since } \exp\left\{\left(\frac{i}{2}\right)(2\pi)\right\} = -1.$$

Thus, the general solution is given by

$$x(t) = C_1p_1(t)e^{(\sigma+i/2)t} + C_2p_2(t)e^{(-\sigma+i/2)t},$$

where $p_j(t)$ are 2π-periodic.

- Solution is unbounded.
- The general solution can be rewritten in a more compact form

$$x(t) = C_1q_1(t)e^{\sigma t} + C_2q_2(t)e^{-\sigma t},$$

where $q_j(t) = p_j(t)e^{(i/2)t}$ is now 4π-periodic.

2.3.4 Transition Curves

The curves $\phi(\alpha, \beta) = \pm 2$ separates region in (α, β) parameter space where all solutions are bounded ($|\phi| < 2$) from region where unbounded solutions exist ($|\phi| > 2$). Although $\phi(\alpha, \beta)$ is not known explicitly, we do know that along the transition curve $\phi = \pm 2$, there are solutions of period 2π or of period 4π. We illustrate how one can determine the transition curves using Fourier series.

The region $\phi = 2$, corresponding to 2π-periodic solutions

The 2π-periodic solution can be represented as

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{int},$$

where $c_{-n} = \bar{c_n}$ since $x(t)$ is real. Substituting this representation into Mathieu equation yields

$$-\sum_{n=-\infty}^{\infty} n^2 c_n e^{int} + \left[\alpha + \frac{\beta}{2}(e^{it} + e^{-it})\right] \sum_{n=-\infty}^{\infty} c_n e^{int} = 0.$$

By comparing coefficients of the term e^{int}, we obtain the following recurrence relation

$$\frac{1}{2}\beta c_{n+1} + (\alpha - n^2)c_n + \frac{1}{2}\beta c_{n-1} = 0.$$

Assuming $\alpha \neq n^2$, the recurrence relation becomes

$$\gamma_n c_{n+1} + c_n + \gamma_n c_{n-1} = 0, \quad \text{where } \gamma_n = \frac{1}{2}\left(\frac{\beta}{\alpha - n^2}\right), n = 0, \pm 1, \pm 2, \ldots.$$

Note that $\gamma_{-n} = \gamma_n$. This can be written as an infinite matrix equation $\Gamma c = \mathbf{0}$, where

$$\Gamma(\alpha, \beta) = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \cdots & \cdots \\ \ddots & \gamma_1 & 1 & \gamma_1 & 0 & 0 & \cdots \\ \ddots & 0 & \gamma_0 & 1 & \gamma_0 & 0 & \ddots \\ \cdots & 0 & 0 & \gamma_1 & 1 & \gamma_1 & \ddots \\ \cdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

and non-trivial solutions exist if $\det(\Gamma(\alpha, \beta)) = 0$. Hence, the curve $\phi(\alpha, \beta) = 2$ is equivalent to the curve along which $\det(\Gamma(\alpha, \beta)) = 0$.

The region $\phi = -2$, corresponding to 4π-periodic solutions

The 4π-periodic solution can be represented as

$$x(t) = \sum_{n=-\infty}^{\infty} d_n e^{int/2}.$$

Substituting this representation into Mathieu equation yields

$$-\frac{1}{4}\sum_{n=-\infty}^{\infty} n^2 d_n e^{int/2} + \left[\alpha + \frac{\beta}{2}(e^{it} + e^{-it})\right] \sum_{n=-\infty}^{\infty} d_n e^{int/2} = 0.$$

By comparing coefficients of $e^{int/2}$, we obtain the following recurrence relation

$$\frac{1}{2}\beta d_{n+2} + \left(\alpha - \frac{n^2}{4}\right) d_n + \frac{1}{2}\beta d_{n-2} = 0.$$

We have two branches of solutions, since this set of equations splits into 2 independent sets, namely odd and even n.

1. Even $n = 2m$, which leads to the condition $\det(\Gamma(\alpha, \beta)) = 0$. This is consistent since 2π-periodic solutions are also 4π-periodic.

2. Odd $n = 2m - 1$, which leads to 4π-periodic solutions. Assuming that $\alpha \neq \frac{1}{4}(2m-1)^2$, the recurrence relation can be written as an infinite matrix equation $\hat{\Gamma} d = \mathbf{0}$:

$$\hat{\Gamma} = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots & \cdots \\ \ddots & \delta_2 & 1 & \delta_2 & 0 & 0 & 0 & \cdots \\ \ddots & 0 & \delta_1 & 1 & \delta_1 & 0 & 0 & \ddots \\ \ddots & 0 & 0 & \delta_1 & 1 & \delta_1 & 0 & \ddots \\ \cdots & 0 & 0 & 0 & \delta_2 & 1 & \delta_2 & \ddots \\ \cdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

with $\delta_m = \frac{1}{2}\left[\frac{\beta}{\alpha - \frac{1}{4}(2m-1)^2}\right]$, $m = \pm 1, \pm 2, \ldots$.

2.3.5 Perturbation Analysis of Transition Curves

Suppose $|\beta|$ is small, we can use **regular perturbation method** to find the transition curves, *i.e.* we take transition curves to be

$$\alpha = \alpha(\beta) = \alpha_0 + \beta\alpha_1 + \beta^2\alpha_2 + \ldots, \tag{2.3.5}$$

with corresponding 2π or 4π periodic solutions

$$x(t) = x_0(t) + \beta x_1(t) + \beta^2 x_2(t) + \ldots. \tag{2.3.6}$$

Substituting both (2.3.5) and (2.3.6) into Mathieu equation, and expanding in powers of β yields

$$(\ddot{x}_0 + \beta\ddot{x}_1 + \ldots) + \Big[(\alpha_0 + \beta\alpha_1 + \ldots) + \beta\cos(t)\Big](x_0 + \beta x_1 + \ldots) = 0,$$

which then gives the following set of ODEs, in order of $\mathcal{O}(1), \mathcal{O}(\beta), \mathcal{O}(\beta^2)$ respectively

$$\ddot{x}_0 + \alpha_0 x_0 = 0 \tag{2.3.7a}$$

$$\ddot{x}_1 + \alpha_0 x_1 = -[\alpha_1 + \cos(t)]x_0 \tag{2.3.7b}$$
$$\ddot{x}_2 + \alpha_0 x_2 = -\alpha_2 x_0 - [\alpha_1 + \cos(t)]x_1. \tag{2.3.7c}$$

From previous subsection, we know that solutions have minimal period of 2π if

$$\alpha_0 = n^2, \quad n = 0, 1, 2, \dots,$$

where different α_0 leads to different transition curves. Similarly, solutions have minimal period of 4π if

$$\alpha_0 = \frac{1}{4}(2m-1)^2 = \left(m - \frac{1}{2}\right)^2, \quad m = 1, 2, \dots,$$

or equivalently

$$\alpha_0 = \left(n + \frac{1}{2}\right)^2, \quad n = 0, 1, 2, \dots.$$

We include both cases by looking for solutions with $\alpha_0 = \dfrac{n^2}{4}$, $n = 0, 1, 2, \cdots$.

The case $n = 0$, *i.e.* $\alpha_0 = 0$

- (2.3.7a) becomes $\ddot{x}_0 = 0$, which has solution $x_1(t) = At + B$. This has periodic solution provided $A = 0$, so periodic solution of (2.3.7a) is given by

$$x_0(t) = C_0. \tag{2.3.8}$$

- (2.3.7b) becomes $\ddot{x}_1 = -[\alpha_1 + \cos(t)]C_0$. This has periodic solutions provided $\alpha_1 = 0$, so periodic solution of (2.3.7b) has the form

$$x_1(t) = C_0 \cos(t) + C_1. \tag{2.3.9}$$

- Substituting (2.3.8) and (2.3.9) into (2.3.7c) yields

$$\begin{aligned}
\ddot{x}_2 &= -\alpha_2 x_0 - [\alpha_1 + \cos(t)]x_1 \\
&= -C_0\alpha_2 - \cos(t)(C_0\cos(t) + C_1) \\
&= -C_0\alpha_2 - C_0\cos^2(t) - C_1\cos(t) \\
&= -C_0\alpha_2 - \frac{C_0}{2}[1 + \cos(2t)] - C_1\cos(t).
\end{aligned}$$

  This has periodic solutions provided $C_0\alpha_2 + \dfrac{C_0}{2} = 0 \iff \alpha_2 = -\dfrac{1}{2}$, so periodic solution of (2.3.7c) has the form

$$x_2(t) = \frac{C_0}{8}\cos(2t) + C_1\cos(t). \tag{2.3.10}$$

Hence, for small $|\beta|$, the transition curve takes the form $\alpha(\beta) = -\dfrac{1}{2}\beta^2 + O(\beta^2)$, which is approximately quadractic. The corresponding 2π periodic solution has the form

$$x(t) = C_0 + \beta\Big[C_0\cos(t) + C_1\Big] + \beta^2\left[\frac{C_0}{8}\cos(2t) + C_1\cos(t)\right] + \mathcal{O}(\beta^3).$$

The case $n = 1$, *i.e.* $\alpha_0 = 1/4$

- (2.3.7a) becomes $\ddot{x}_0 = -\dfrac{1}{4}x_0 = 0$, which has solution

$$x_0(t) = C_0 \cos\left(\frac{t}{2}\right) + D_0 \sin\left(\frac{t}{2}\right). \tag{2.3.11}$$

- Substituting (2.3.11) into (2.3.7b) yields

$$\begin{aligned}
\ddot{x}_1 + \frac{1}{4}x_1 &= -\Big[\alpha_1 + \cos(t)\Big]\left[C_0\cos\left(\frac{t}{2}\right) + D_0\sin\left(\frac{t}{2}\right)\right] \\
&= -C_0\alpha_1\cos\left(\frac{t}{2}\right) - D_0\alpha_1\sin\left(\frac{t}{2}\right) \\
&\qquad - C_0\cos(t)\cos\left(\frac{t}{2}\right) - D_0\cos(t)\sin\left(\frac{t}{2}\right) \\
&= -C_0\alpha_1\cos\left(\frac{t}{2}\right) - D_0\alpha_1\sin\left(\frac{t}{2}\right) \\
&\qquad - \frac{C_0}{2}\left[2\cos(t)\cos\left(\frac{t}{2}\right)\right] - \frac{D_0}{2}\left[2\cos(t)\sin\left(\frac{t}{2}\right)\right] \\
&= -C_0\alpha_1\cos\left(\frac{t}{2}\right) - D_0\alpha_1\sin\left(\frac{t}{2}\right) \\
&\quad - \frac{C_0}{2}\left[\cos\left(t+\frac{t}{2}\right) + \sin(t)\sin\left(\frac{t}{2}\right) + \cos\left(t-\frac{t}{2}\right) - \sin(t)\sin\left(\frac{t}{2}\right)\right] \\
&\quad - \frac{D_0}{2}\left[\sin\left(t+\frac{t}{2}\right) - \sin(t)\cos\left(\frac{t}{2}\right) - \sin\left(t-\frac{t}{2}\right) + \sin(t)\cos\left(\frac{t}{2}\right)\right] \\
&= -C_0\left[\alpha_1 + \frac{1}{2}\right]\cos\left(\frac{t}{2}\right) - D_0\left[\alpha_1 - \frac{1}{2}\right]\sin\left(\frac{t}{2}\right) \\
&\qquad - \frac{1}{2}C_0\cos\left(\frac{3t}{2}\right) - \frac{1}{2}D_0\sin\left(\frac{3t}{2}\right).
\end{aligned}$$

  In order to have a 4π-periodic solution, we must eliminate the **secular/resonant** terms. There are two possible cases:

  1. $C_0 = 0, \alpha_1 = 1/2$, which gives the 4π-periodic (particular) solution
  $$x_1(t) = \frac{1}{4}D_0\sin\left(\frac{3}{2}t\right).$$
  2. $D_0 = 0, \alpha_1 = -1/2$, which gives the 4π-periodic (particular) solution
  $$x_2(t) = \frac{1}{4}C_0\cos\left(\frac{3}{2}t\right).$$

Hence, for small $|\beta|$, the transition curve takes the form $\alpha(\beta) = \dfrac{1}{4} \pm \dfrac{1}{2}\beta + O(\beta^2)$, which is approximately linear.

2.4 Stability of Periodic Solutions

Stability of a fixed point is a **local problem**, since its stability can be determined by analysing the behaviour of an arbitrary neighbourhood around the fixed point. On the contrary, stability of a periodic orbit is loosely speaking a **global problem**, since one needs to analyse the behaviour of the corresponding vector field in a neighbourhood of the entire periodic orbit. Consequently, such problem becomes extremely difficult, and new mathematical tools need to be developed. Periodic solutions are generally not stable in the sense of Lyapunov. That is, orbits starting out in a neighbourhood of the periodic orbit will tend to have slightly different frequencies, and so they develop a phase shift as time evolves. This motivates the notion of a **Poincaré section**.

Consider an autonomous system of the form $\dot{x} = f(x), x \in \mathbb{R}^n$, with periodic solution $\Gamma(t)$. Choose any point a on $\Gamma(t)$ and construct an $(n-1)$-dimensional submanifold $V \subset \mathbb{R}^n$ such that V is a transversal section of f through a, *i.e.* for any $x_0 \in V$, $(x_0, f(x_0)) \notin T_{x_0}(V)$, the tangent space of V at x_0. Consider an orbit $\gamma(x_0)$ starting at $x_0 \in V$, we follow the orbit until it returns to V for the first time. By continuity of f, we can always choose x_0 sufficiently close to a for this to hold; all such points constitute a Poincaré section. More precisely, an open subset $\Sigma \subset V$ containing a is called a **Poincaré section** if each point of Σ returns to V. Viewing the flow as a map, we can define the **Poincaré return map**

$$P\colon \Sigma \longrightarrow V\colon x_0 \mapsto \phi_{T(x_0)}(x_0),$$

where $T(x_0) > 0$ is the first time where the orbit $\gamma(x_0)$ returns to V. Observe that a is a fixed point of P by construction. We now apply the idea of Lyapunov stability on the Poincaré section.

Definition 2.4.1. Given an autonomous system, with periodic solution $\phi(t)$, a transversal V and a Poincare return map P, whose fixed point is a.

(a) We say that $\phi(t)$ is **stable** if for all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$|x_0 - a| \leq \delta,\, x_0 \in \Sigma \implies |P^n(x_0) - a| \leq \varepsilon.$$

(b) We say that $\phi(t)$ is **asymptotically stable** if it is stable, and there exists $\delta_1 > 0$ such that

$$|x_0 - a| \leq \delta_1,\, x_0 \in \Sigma \implies \lim_{n\to\infty} P^n(x_0) = a.$$

Stability of a periodic orbit is closely related to the stability of fixed points of its corresponding Poincaré map. Indeed, choose an arbitrary $a \in \Sigma \cap \Gamma(t)$ and let $\delta x_0 = x_0 - a, \delta P(x_0) = P(x_0) - a$. Linearising the Poincaré map P about the point a yields the linearised system

$$\delta P(x_0) = DP(a)\delta x_0.$$

One can show that if eigenvalues of the Jacobian $DP(a) \in \mathbb{R}^{(n-1)\times(n-1)}$ lies inside the unit circle in $\mathbb{C}$, then the periodic solution $\Gamma(t)$ is asymptotically stable. Moreover, these eigenvalues are independent of the choice of $a \in \Gamma(t)$ and its transversal section V. While Poincaré map is a powerful concept in the study of periodic orbits, it is difficult to construct a Poincaré map analytically in general.

Non-Autonomous Case

Consider the non-autonomous system

$$\dot{x} = f(x,t), x \in \mathbb{R}^n, t \in \mathbb{R},$$

where $f(x,\cdot)$ is T-periodic. One can apply the previous definition of stability by rewriting this system as the $(n+1)$-dimensional autonomous system given by

$$\begin{cases} \dot{x} & = f(x,\theta) \\ \dot{\theta} & = 1, \quad \theta(0) = 0, \end{cases}$$

where θ is the phase variable and $(x,\theta) \in \mathbb{R}^n \times S^1$ with $S^1 \simeq \mathbb{R}/[0,T]$, due to T-periodicity of $f(x,\cdot)$. The transversal V is now n-dimensional and a natural choice is to consider the mapping $P\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, obtained by strobing the solution at times $0, T, 2T, \ldots$.

Example 2.4.2. Consider a linear dissipative oscillator with natural frequency ω_0

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = h\cos(\omega t), \quad 0 < \mu < \omega_0,\ h, \omega > 0.$$

This is an inhomogeneous second-order constant coefficient ODEs, which can be solve explicitly. The general solution is given by

$$\begin{aligned} x(t) = C_1 e^{-\mu t}\cos\left(t\sqrt{\omega_0^2-\mu^2}\right) + C_2 e^{-\mu t}\sin\left(t\sqrt{\omega_0^2-\mu^2}\right) \\ + \alpha\cos(\omega t) + \beta\sin(\omega t), \end{aligned} \tag{2.4.1}$$

where

$$\alpha = \frac{(\omega_0^2-\omega^2)h}{4\mu^2\omega^2 + (\omega_0^2-\omega^2)}, \quad \beta = \frac{2\mu\omega h}{4\mu^2\omega^2 + (\omega_0^2-\omega^2)}.$$

Observe that the solution has its amplitude at $\omega = \omega_0$. If $\mu = 0$, *i.e.* no dissipation, then α blows up as ω approaches the natural frequency ω_0 and we have resonance.

The constant C_1, C_2 are determined by the initial condition $x(0), \dot{x}(0)$. Solving for C_1, C_2 yields

$$C_1 = x(0) - \alpha, \quad C_2 = \frac{\dot{x}(0) + \mu x(0) - \mu\alpha - \omega\beta}{(\omega_0^2-\mu^2)^{1/2}}.$$

Note that (2.4.1) is a periodic solution with period $2\pi/\omega$ if $C_1 = C_2 = 0$, *i.e.* $x(0) = \alpha, \dot{x}(0) = \omega\beta$. This suggests that we choose a Poincaré section containing the point $(\alpha, \omega\beta)$. Rewriting the equation as follows

$$\begin{cases} \dot{x} & = y \\ \dot{y} & = -\omega_0^2 x - 2\mu y + h\cos(\omega\theta) \\ \dot{\theta} & = 1, \quad \theta(0) = 0. \end{cases}$$

We construct Poincaré return map P by strobing at times $t = 0, 2\pi/\omega, 4\pi/\omega, \ldots$. The point $(\alpha, \omega\beta)$ is the fixed point of P. Substitution of $t = \dfrac{2\pi}{\omega}$ with $\gamma = \left(\dfrac{2\pi}{\omega}\right)\sqrt{(\omega_0^2-\mu^2)}$ yields

$$P\begin{bmatrix} x(0) \\ \dot{x}(0) \end{bmatrix} = P\begin{bmatrix} C_1 + \alpha \\ C_2(\omega_0^2-\mu^2)^{1/2} + \omega\beta - \mu C_1 \end{bmatrix} = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix}\Bigg|_{t=2\pi/\omega}.$$

Moreover, it can be shown that $\lim_{n\to\infty} P^n \begin{bmatrix} x(0) \\ \dot{x}(0) \end{bmatrix} = \begin{bmatrix} \alpha \\ \omega\beta \end{bmatrix}$.

2.5 Stable Manifold Theorem

As opposed to nonlinear system, we have a far more satisfactory understanding towards the behaviour of fixed points of a linear system. A common approach in understanding the local behaviour of nonlinear systems near their fixed points is to study the linearised dynamics about fixed points. Indeed, we have a far more satisfactory theoretical understanding on linear systems. However, a major issue needs to be addressed: does linearisation technique provide a reasonable prediction about the local behaviour of the original nonlinear system around a fixed point? On a different perspective, which features of the linearised dynamic survive the addition of nonlinear effect? This motivates the definition of hyperbolic fixed point.

Definition 2.5.1. A fixed point x_0 of an ODE $\dot{x} = f(x), x \in \mathbb{R}^n$ is said to be **hyperbolic** if the Jacobian evaluated at x_0, $Df(x_0)$ has all eigenvalues with non-zero real part.

- Hyperbolic fixed points are structurally robust/stable, *i.e.* small changes do not alter the nature of the problem (up to topological equivalence).

Definition 2.5.2. Let x_0 be a hyperbolic fixed point of an ODE $\dot{x} = f(x), x \in \mathbb{R}^n$.

(a) The **forward contracting/stable subspace** for x_0 of the linearised system, denoted by E^s is the subset spanned by eigenvectors or generalised eigenvectors of $Df(x_0)$ with eigenvalues having negative real part.

$$E^s = \bigoplus_{\mathrm{Re}(\lambda_j)<0} \mathcal{N}(Df(x_0) - \lambda_j)^{n_j}.$$

Thus, E^s is the set of vectors whose forward orbits go to the origin under the linearised dynamics.

(b) The **backward contracting/unstable subspace** for x_0 of the linearised system, denoted by E^u is the subset spanned by eigenvectors or generalised eigenvectors of $Df(x_0)$ with eigenvalues having positive real part.

$$E^u = \bigoplus_{\mathrm{Re}(\lambda_j)>0} \mathcal{N}(Df(x_0) - \lambda_j)^{n_j}.$$

Thus, E^u is the set of vectors whose backward orbits go to the origin under the linearised dynamics.

Theorem 2.5.3 (Stable Manifold Theorem)**.** *Consider the first order ODE $\dot{x} = f(x), x \in \mathbb{R}^n$ and let ϕ_t be the flow of the system. Near a hyperbolic fixed point x_0,* *(a) there exists a stable differentiable manifold S tangent at x_0 to the stable subspace E^s of the linearised system $\dot{x} = Df(x_0)x$, having the same dimension as E^s. Moreover, S is forward/positively invariant under ϕ_t and for all $x \in S$,*

$$\lim_{t\to\infty} \phi_t(x) = x_0.$$

(b) there exists an unstable differentiable manifold U tangent at x_0 to the unstable subspace E^u of the linearised system $\dot{x} = Df(x_0)x$, having the same dimension as E^u. Moreover, S is backward/negatively invariant under ϕ_t and for all $x \in U$,

$$\lim_{t\to-\infty} \phi_t(x) = x_0.$$

Example 2.5.4. Consider the following 3×3 system

$$\begin{cases} \dot{x_1} &= -x_1 \\ \dot{x_2} &= -x_2 + x_1^2 \\ \dot{x_3} &= x_3 + x_1^2. \end{cases}$$

The only fixed point is the origin $(0,0,0)$. Since $Df((0,0,0)) = \text{diag}(-1,-1,1)$, we see that E^s is the (x_1, x_2)-plane and E^u is the x_3 axis. Writing $x(0) = c = [c_1, c_2, c_3]^T$, the general solution is given by

$$\begin{aligned} x_1(t) &= c_1 e^{-t} \\ x_2(t) &= c_2 e^{-t} + c_1^2(e^t - e^{-2t}) \\ x_3(t) &= c_3 e^t + \frac{c_1^2}{3}(e^t - e^{-2t}). \end{aligned}$$

Clearly, $\lim_{t\to\infty} \phi_t(c) = \mathbf{0} \iff c_3 + c_1^2/3 = 0$. Thus,

$$S = \left\{ c \in \mathbb{R}^3 \colon c_3 = -\frac{c_1^2}{3} \right\}.$$

Similarly, $\lim_{t\to-\infty} \phi_t(c) = \mathbf{0} \iff c_1 = c_2 = 0$. Thus,

$$U = \left\{ c \in \mathbb{R}^3 \colon c_1 = c_2 = 0 \right\}.$$

Proof. (Sketch) WLOG, we assume that the hyperbolic fixed point is at $x = \mathbf{0}$.

(A) Rewrite ODE as $\dot{x} = f(x) = Ax + F(x)$, where $A = Df(\mathbf{0})$ and $F \in C^1(U)$. Note that $F(\mathbf{0}) = \mathbf{0}$ and $DF(\mathbf{0}) = \mathbf{0}$. Since $x = \mathbf{0}$ is a hyperbolic fixed point, there exists an $n \times n$ transformation via an invertible matrix C such that

$$B = C^{-1}AC = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix},$$

whose eigenvalues

- $\lambda_1, \dots, \lambda_k$ of $(k \times k)$ matrix P have negative real part.
- $\lambda_{k+1}, \dots, \lambda_n$ of $(n-k) \times (n-k)$ matrix Q have positive real part.

(B) Performing a change of variable $y := C^{-1}x$ gives

$$\dot{y} = C^{-1}\dot{x} = C^{-1}\Big[Ax + F(x)\Big] = C^{-1}ACy + C^{-1}F(Cy) = By + G(y),$$

where $G(y) := C^{-1}F(Cy)$. Define the following matrices

$$U(t) = \begin{bmatrix} e^{Pt} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad V(t) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & e^{Qt} \end{bmatrix}.$$

The derivatives of $U(t)$ and $V(t)$ are

$$\begin{aligned} \dot{U} &= BU \quad \text{with} \quad \|U(t)\| \longrightarrow 0 \text{ as } t \longrightarrow \infty. \\ \dot{V} &= BV \quad \text{with} \quad \|V(t)\| \longrightarrow 0 \text{ as } t \longrightarrow -\infty. \end{aligned}$$

(C) Let a be a parameter and consider the following integral equation

$$u(t,a) = U(t)a + \int_0^t U(t-s)G(u(s,a))\,ds - \int_t^\infty V(t-s)G(u(s,a))\,ds. \tag{2.5.1}$$

Differentiating (2.5.1) with respect to t using **Leibniz's rule** shows that $\dot{u}$ satisfies the following ODE

$$\begin{aligned} \dot{u} = \dot{U}a \quad &+ \int_0^t \dot{U}(t-s)G(u(s,a))\,ds \quad + U(\mathbf{0})G(u(t,a)) \\ &- \int_t^\infty \dot{V}(t-s)G(u(s,a))\,ds \quad + V(\mathbf{0})G(u(t,a)) \\ = BUa + B &\int_0^t U(t-s)G(u(s,a))\,ds \; + U(\mathbf{0})G(u(t,a)) \\ &- B\int_t^\infty V(t-s)G(u(s,a))\,ds + V(\mathbf{0})G(u(t,a)) \\ = Bu \quad &+ G(u). \end{aligned}$$

since $U(\mathbf{0}) + V(\mathbf{0}) = I$. Choose a such that $a_j = 0, j = k+1, \dots, n$. Then,

$$u_j(0,a) = \begin{cases} a_j & \text{for } j = 1, \dots, k, \\ -\displaystyle\int_0^\infty V(-s)G(u(s, a_1, \dots, a_k, 0, \dots, 0))\,ds & \text{for } j = k+1, \dots, n. \end{cases}$$

Setting $y_j = u_j(0,a)$, we thus have the condition $y_j = \psi_j(y_1, \dots, y_k)$ for $j = k+1, \dots, n$. This defines a stable manifold $\tilde{S}$, *i.e.*

$$\tilde{S} = \Big\{(y_1, \dots, y_n) \colon y_j = \psi_j(y_1, \dots, y_k), \quad j = k+1, \dots, n\Big\}.$$

(D) It can be proven that this defines a stable manifold $\tilde{S}$, such that if $Y(0) \in \tilde{S}$, then $Y(t) \longrightarrow 0$ as $t \longrightarrow \infty$. The stable manifold S is obtained in the original x-space, under the linear transformation $x = Cy$. To find the unstable manifold U, take $t \mapsto -t$; we obtain $\dot{y} = -By - G(y)$ and write

$$y = (y_{k+1}, \dots, y_n, y_1, \dots, y_k).$$

∎

Remark 2.5.5. The integral equation (2.5.1) is chosen in such a way that it converges to the hyperbolic fixed point $x = \mathbf{0}$ as $t \longrightarrow \infty$. One can show that the general solution to the equation $\dot{y} = By + G(y)$ has the form

$$y(t) = e^{tB}y_0 + \int_0^t e^{(t-s)B}G(u(s))\,ds$$

With the form of B we have, the general solution has the form

$$y(t,a) = \begin{bmatrix} e^{Pt} & \mathbf{0} \\ \mathbf{0} & e^{Qt} \end{bmatrix} a + \int_0^t \begin{bmatrix} e^{P(t-s)} & \mathbf{0} \\ \mathbf{0} & e^{Q(t-s)} \end{bmatrix} G(u(s,a))\,ds. \tag{2.5.2}$$

(2.5.1) is then obtained from (2.5.2) by removing terms that are divergent as $t \longrightarrow \infty$. In practice, one can construct an approximation of the stable manifold using the method of successive approximation, given by the integral equation (2.5.1).

Example 2.5.6. Consider the system $\begin{cases} \dot{x_1} &= -x_1 - x_2^2 \\ \dot{x_2} &= x_2 + x_1^2. \end{cases}$. The iterative scheme for approximating $\tilde{S}$ is given by

$$\begin{cases} u^0(t,a) &= \mathbf{0} \\ u^{(j+1)}(t,a) &= U(t)a + \displaystyle\int_0^t U(t-s)G(u^j(s,a))\,ds - \int_t^\infty V(t-s)G(u^j(s,a))\,ds \end{cases}$$

- We see that $A = B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $F(x) = G(x) = \begin{bmatrix} -x_2^2 \\ x_1^2 \end{bmatrix}$. This implies that:

$$U(t) = \begin{bmatrix} e^{-t} & 0 \\ 0 & 0 \end{bmatrix}, \quad V(t) = \begin{bmatrix} 0 & 0 \\ 0 & e^t \end{bmatrix}, \quad a = \begin{bmatrix} a_1 \\ 0 \end{bmatrix}.$$

- Substituting into the iterative scheme yields

$$u^{(1)}(t,a) = U(t)a = \begin{bmatrix} e^{-t}a_1 \\ 0 \end{bmatrix}$$

$$\begin{aligned}
u^{(2)}(t,a) &= \begin{bmatrix} e^{-t}a_1 \\ 0 \end{bmatrix} + \int_0^t \begin{bmatrix} e^{-(t-s)} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ (e^{-s}a_1)^2 \end{bmatrix} ds - \int_t^\infty \begin{bmatrix} 0 & 0 \\ 0 & e^{t-s} \end{bmatrix} \begin{bmatrix} 0 \\ (e^{-s}a_1)^2 \end{bmatrix} ds \\
&= \begin{bmatrix} e^{-t}a_1 \\ 0 \end{bmatrix} - \int_t^\infty \begin{bmatrix} 0 \\ e^{t-s}e^{-2s}a_1^2 \end{bmatrix} ds \\
&= \begin{bmatrix} e^{-t}a_1 \\ -\dfrac{e^{-2t}}{3}a_1^2 \end{bmatrix} \\
u^{(3)}(t,a) &= \begin{bmatrix} e^{-t}a_1 \\ 0 \end{bmatrix} + \int_0^t \begin{bmatrix} e^{-(t-s)} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -e^{-4s}a_1^4/9 \\ e^{-2s}a_1^2 \end{bmatrix} ds - \int_t^\infty \begin{bmatrix} 0 & 0 \\ 0 & e^{t-s} \end{bmatrix} \begin{bmatrix} -e^{-4s}a_1^4/9 \\ e^{-2s}a_1^2 \end{bmatrix} ds \\
&= \begin{bmatrix} e^{-t}a_1 \\ 0 \end{bmatrix} - \frac{1}{9}\int_0^t \begin{bmatrix} e^{-(t-s)}e^{-4s}a_1^4 \\ 0 \end{bmatrix} ds - \int_t^\infty \begin{bmatrix} 0 \\ e^{t-s}e^{-2s}a_1^2 \end{bmatrix} ds \\
&= \begin{bmatrix} e^{-t}a_1 + \dfrac{1}{27}\left[e^{-4t} - e^{-t}\right]a_1^4 \\ -\dfrac{1}{3}e^{-2t}a_1^2 \end{bmatrix}
\end{aligned}$$

- Letting $t \longrightarrow 0$, we see that $u^{(3)}(t,a) = \begin{bmatrix} a_1 \\ -\frac{1}{3}a_1^2 \end{bmatrix} \implies x_2 = \psi_2(a_1) = -\frac{1}{3}a_1^2$. Thus, the stable manifold for small a_1 is given by $x_2 = -\frac{x_1^2}{3}$. Note that the linear stable manifold is $\begin{bmatrix} a_1 \\ 0 \end{bmatrix}$.

In the proof, the stable and unstable manifolds S and U are only defined in a small neighbourhood of the origin and are referred to as **local manifolds**. One can define **global manifolds** as follows.

Definition 2.5.7. Let ϕ_t be the flow of $\dot{x} = f(x)$, $x \in \mathbb{R}^n$. The **global stable/unstable manifolds** are

$$\begin{aligned}
W^s(0) &= \bigcup_{t\leq 0} \phi_t(S) \\
W^u(0) &= \bigcup_{t\geq 0} \phi_t(U)
\end{aligned}$$

2.6 Centre Manifolds

Bifurcation theory concerns the change in qualitative behaviour of a dynamical system as one or more parameters are changed. For a hyperbolic fixed point, the local behaviour of the flow is completely determined by the linearised flow up to homeomorphism [**Stable Manifold Theorem/ Hartman-Grobman Theorem**]. It follows that small perturbation of the equation will also have a fixed point of the same stability type, called **structural stability**. This suggests that bifurcations of fixed points can only occur at parameter values for which a fixed point is non-hyperbolic, and provides a criteria for detecting bifurcation:

Find parameter values for which the linearised flow near a fixed point has a zero or purely imaginary eigenvalues.

Definition 2.6.1. Let x_0 be a fixed point of an ODE $\dot{x} = f(x), x \in \mathbb{R}^n$. The centre subspace for x_0 of the linearised system, denoted by E^c is the subset spanned by eigenvectors or generalised eigenvectors of $Df(x_0)$ with eigenvalues having zero real part.

$$E^c = \bigoplus_{\mathrm{Re}(\lambda_j)=0} \mathcal{N}(Df(x_0) - \lambda_j)^{n_j}.$$

Theorem 2.6.2 (Centre Manifold Theorem)**.** *There exists a nonlinear mapping*

$$h\colon E^c \longrightarrow E^s \oplus E^u,$$

with $h(0) = 0$, $Dh(0) = 0$, and a neighbourhood U of $x = 0$ in $\mathbb{R}^n$ such that the center manifold

$$M = \Big\{(x, h(x))\colon x \in E^c\Big\}$$

has the following properties:

(a) ***Invariance***: *the center manifold M is locally invariant with respect to the given ODEs, i.e. if $x(0) = M \bigcap U$, then $x(t) \in M$ as long as $x(t) \in U$. That is, $x(t)$ can only leave M when it leaves the neighbourhood U.*

(b) ***Attracting***: *if $E^u = \{\mathbf{0}\}$, then M is locally attracting, i.e. all solutions staying in U converges exponentially to a solution on M.*

(c) M has the same dimension as E^c, contains $x = \mathbf{0}$ and is tangential to E^c at the origin.

Construction of Center Manifold

Suppose that $E^u = \{\mathbf{0}\}$ and the given ODE can be written in the form

$$\begin{cases} \dot{x} &= Ax + f_1(x, y) \\ \dot{y} &= -By + f_2(x, y), \end{cases}$$

where eigenvalues of A have zero real part, eigenvalues of B have strictly positive real part and for $j = 1, 2,\ f_j(0, 0) = 0,\ Df_j(0, 0) = 0$. The Centre Manifold Theorem implies that the flow on the nonlinear centre manifold M can be written as:

$$\dot{x} = Ax + f_1(x, h(x)) \equiv G(x).$$

To determine $h(\cdot)$, first note that we have $y = h(x)$ on M. Differentiating this with respect to t gives

$$\dot{y} = Dh(x)\dot{x} = Dh(x)\Big[Ax + f_1(x, h(x))\Big]. \tag{2.6.1}$$

On the other hand, we have that

$$\dot{y} = -Bh(x) + f_2(x, h(x)). \tag{2.6.2}$$

Comparing (2.6.1) and (2.6.2) yields

$$Dh(x)\Big[Ax + f_1(x, h(x))\Big] = -Bh(x) + f_2(x, h(x)) \tag{2.6.3}$$

together with the condition $h(\mathbf{0}) = \mathbf{0}, Dh(\mathbf{0}) = \mathbf{0}$; these two conditions tells us that $h(\cdot)$ is at least quadratic. We can solve for $h(x)$ by Taylor expansions in power of components of x.

Example 2.6.3. Consider the nonlinear 2×2 system

$$\begin{cases} \dot{x} &= xy \\ \dot{y} &= -y - x^2. \end{cases}$$

It should be clear that $x \in E^c$ and $y \in E^s$. We try a solution of the form:

$$y = h(x) = ax^2 + bx^3 + cx^4 + dx^5 + O(x^6).$$

From previous discussions, we get the following two equations:

$$\begin{aligned} \dot{y} = h'(x)\dot{x} &= h'(x)[xh(x)] \\ &= x\Big[ax^2 + bx^3 + \dots\Big]\Big[2ax + 3bx^2 + \dots\Big] \\ &= 2ax^4 + 5abx^5 + O(x^6). \\ \dot{y} = -y - x^2 &= -h(x) - x^2 \\ &= -(a+1)x^2 - bx^3 - cx^4 - dx^5 + O(x^6). \end{aligned}$$

Comparing coefficients of powers of x, we obtain $a = -1, b = 0, c = 2, d = 0$. Thus

$$y = h(x) = -x^2 - 2x^4 + O(x^6).$$

The dynamics on the center manifold is approximated by:

$$\dot{x} = xh(x) = -x\Big[x^2 + 2x^4 + O(x^6)\Big].$$

Since the big bracket term is positive definite, it appears from phase plane analysis that we are going to converge to the fixed point.

2.7 Problems

1. Let $x\colon [0,\infty) \longrightarrow \mathbb{R}^n$ such that $|x(t)| \le Me^{\alpha t}$ for some constant $M \ge 0$ and $\alpha \in \mathbb{R}$. Introduce the Laplace transform

$$\tilde{x}(s) = \int_0^\infty e^{-st} x(t)\, dt.$$

   (a) Laplace transform the first order ODE in $\mathbb{R}^n$

$$\dot{x} = Ax + f(t), \quad x(\mathbf{0}) = x_0,$$

   to obtain the following solution in Laplace space

$$\tilde{x}(s) = (sI - A)^{-1}[x_0 + \tilde{f}(s)].$$

   Solution:

   (b) In the scalar case ($n = 1$), invert the Laplace transform to obtain the solution

$$x(t) = e^{-At} + \int_0^t e^{-(t-t')} f(t')\, dt'.$$

   Solution:

   (c) Use Laplace transforms to solve the second order ODE (for a RLC electronic circuit)

$$L\ddot{I}(t) + R\dot{I}(t) + \frac{1}{C} I(t) = V_0 \cos(\omega t),$$

   where $\omega \neq \omega_0 := 1/\sqrt{LC}$. This is a model of a RLC circuit in electronics where R is resistance, C is the capacitance and L is the inductance.

   Solution:

2. Consider a linear chain of $2N$ atoms consisting of two different masses m, M, $M > m$, placed alternately. The atoms are equally spaced with lattice spacing a with nearest-neighbour interactions represented by Hookean springs with spring constant β. Label the light atoms by even integers $2n, n = 0, \ldots, N-1$ and the heavy atoms by odd integers $2n-1, n = 1, \ldots, N$. Denoting their displacements from equilibrium by the variables U_{2n} and V_{2n-1} respectively, Newton's law of motion gives

$$\begin{cases} m\ddot{U}_{2n} &= \beta[V_{2n-1} + V_{2n+1} - 2U_{2n}] \\ M\ddot{V}_{2n-1} &= \beta[U_{2n} + U_{2n-2} - 2V_{2n-1}]. \end{cases}$$

   Assume periodic boundary conditions $U_0 = U_{2N}$ and $V_1 = V_{2N+1}$.

   (a) Sketch the configuration of atoms and briefly explain how the dynamical equations arise f rom Newton's law of motion.

Solution:

(b) Assuming a solution of the form

$$U_{2n} = \Phi e^{2inka} e^{-i\omega t}, \quad V_{2n+1} = \Psi e^{i(2n+1)ka} e^{-i\omega t},$$

derive an eigenvalue equation for the amplitudes (Φ, Ψ) and determine the eigenvalues.

Solution:

(c) Using part (b), show that there are two branches of solution and determine the speed ω/k on the two branches for small k.

Solution:

3. Construct a fundamental matrix for the system

$$\begin{cases} \dot{x_1} &= x_2 \\ \dot{x_2} &= x_3 \\ \dot{x_3} &= -2x_1 + x_2 + 2x_3, \end{cases}$$

and deduce the solution of

$$\begin{cases} \dot{x_1} &= x_2 + e^t \\ \dot{x_2} &= x_3 \\ \dot{x_3} &= -2x_1 + x_2 + 2x_3, \end{cases}$$

with initial condition $x_1(0) = 1, x_2(0) = x_3(0) = 0$.

Solution:

4. Determine the stability of the solutions of

(a) $\dot{x_1} = x_2 \sin(t), \dot{x_2} = 0$.

Solution:

(b) $\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^t.$

Solution:

(c) $\ddot{x} + e^{-t}\dot{x} + x = e^t$.

Solution:

5. (a) Consider the equation $\dot{x} = f(t)A_0x, x \in \mathbb{R}^2$, with $f(t)$ a scalar T-periodic solution and A_0 a constant matrix with real distinct eigenvalues. Determine the corresponding Floquet multipliers.

Solution:

(b) Consider the Hill equation

$$\ddot{x} + q(t)x = 0, \quad q(t+T) = q(t).$$

Rewrite as a first order system. Use Liouville's formula to show that the characteristic multipliers have the form

$$\mu_{\pm} = \Delta \pm \sqrt{\Delta^2 - 1},$$

where $\Delta = \mathrm{tr}(\Psi(T))/2$ and $\Psi(t)$ is the fundamental matrix with $\Psi(0) = I$. Hence, show that all solutions are bounded if $|\Delta| < 1$.

Solution:

6. (a) Use a perturbation expansion in β to show that the transition curves for Mathieu's equation

$$\ddot{x} + (\alpha + \beta\cos(t))x = 0,$$

for $\alpha \approx 1, \beta \approx 0$, are given approximately by

$$\alpha = 1 - \frac{\beta^2}{12}, \quad \alpha = 1 + \frac{5}{12}\beta^2.$$

Solution:

(b) Now suppose that $\alpha \approx 1/4 + \alpha_1\beta, \beta \approx 0$. In the unstable region near $\alpha = 1/4$, solutions of Mathieu's equation are of the form

$$c_1e^{\sigma t}q_1(t) + c_2e^{-\sigma t}q_2(t),$$

where σ is real and positive, and q_1, q_2 are 4π-periodic. Derive the second order equation for q_1, q_2 and perform a power series expansion in β to show that $\sigma \approx \pm\beta\sqrt{1/4 - \alpha_1^2}$.

Solution:

(c) Use part (b) to deduce that solutions of the damped Mathieu equation

$$\ddot{x} + \kappa\dot{x} + (\alpha + \beta\cos(t))x = 0,$$

where $\kappa = \kappa_1\beta + \mathcal{O}(\beta^2)$, are stable if to first order in β,

$$\alpha < \frac{1}{4} - \frac{\beta}{2}\sqrt{1 - \kappa_1^2} \quad \text{or} \quad \alpha > \frac{1}{4} + \frac{\beta}{2}\sqrt{1 - \kappa_1^2}.$$

Chapter 3

Perturbation Theory

3.1 Basics of Perturbation Theory

3.1.1 Asymptotic Expansion

Suppose we want to evaluate the following integral:

$$f(\varepsilon) = \int_0^\infty \frac{e^{-t}}{1+\varepsilon t} dt \quad , \quad \varepsilon > 0.$$

We can develop an approximation of $f(\varepsilon)$ for small ε by repeating integration by parts. More precisely,

$$\begin{aligned} f(\varepsilon) &= 1 - \varepsilon \int_0^\infty \frac{e^{-t}}{(1+\varepsilon t)^2} dt \\ &= 1 - \varepsilon + 2\varepsilon^2 \int_0^\infty \frac{e^{-t}}{(1+\varepsilon t)^3} dt \\ &= \qquad \vdots \qquad \vdots \qquad \vdots \\ &= 1 - \varepsilon + 2\varepsilon^2 - 3!\varepsilon^3 + \cdots\cdots + (-1)^N N!\varepsilon^N + R_N(\varepsilon). \end{aligned}$$

where $R_N(\varepsilon)$ is the remainder term:

$$R_N(\varepsilon) = (-1)^{N+1}(N+1)!\varepsilon^{N+1} \int_0^\infty \frac{e^{-t}}{(1+\varepsilon t)^{N+2}} dt.$$

- Since $\displaystyle\int_0^\infty \frac{1}{(1+\varepsilon t)^{N+2}} e^{-t} dt \leq \int_0^\infty e^{-t} dt = 1$, we see that:

  $$R_N(\varepsilon) \leq (N+1)!\varepsilon^{N+1} <<< |(-1)^N N!\varepsilon^N|$$

  Thus, for fixed N,

  $$\lim_{\varepsilon \to 0} \left| \frac{f(\varepsilon) - \sum_{k=0}^N a_k \varepsilon^k}{\varepsilon^N} \right| = 0$$

  or

  $$f(\varepsilon) = \sum_{k=0}^N a_k \varepsilon^k + O(\varepsilon^{N+1}).$$

- The formal series $\sum_{k=0}^{N} a_k \varepsilon^k$ is said to be an ***asymptotic expansion*** of $f(\varepsilon)$ such that for fixed N, it provides a good approximation of $f(\varepsilon)$ as $\varepsilon \longrightarrow 0$.
- However, the expansion is not convergent for fixed ε since $(-1)^N N!\varepsilon^N \longrightarrow \infty$ as $N \longrightarrow \infty$.

Definition 3.1.1 (Big O and Little O notation)**.**

(a) $f(\varepsilon) = O(g(\varepsilon))$ as $\varepsilon \longrightarrow 0$ means that there exists $M > 0$ such that $|f| \leq M|g|$ as $\varepsilon \longrightarrow 0$.

(b) $f(\varepsilon) = o(g(\varepsilon))$ as $\varepsilon \longrightarrow 0$ means that $\lim\limits_{\varepsilon \to 0} \left| \dfrac{f(\varepsilon)}{g(\varepsilon)} \right| = 0$.

Definition 3.1.2.

(a) The ordered sequence of functions $\{\delta_j(\varepsilon)\}_{j=0}^{\infty}$ is called an **asymptotic sequence** as $\varepsilon \longrightarrow 0$ if $\delta_{j+1}(\varepsilon) = o(\delta_j(\varepsilon))$ as $\varepsilon \longrightarrow 0$.

- One example is the sequence $1, \varepsilon, \varepsilon^2, \cdots$ for small ε.

(b) Let $f(\varepsilon)$ be a continuous function of ε, and $\{\delta_j(\varepsilon)\}_{j=0}^{\infty}$ be an asymptotic sequence. The formal series $\sum_{j=0}^{N} a_j \delta_j(\varepsilon)$ is called an **asymptotic expansion** of $f(\varepsilon)$, valid to order $\delta_N(\varepsilon)$, if for all $N \geq 0$,

$$\lim_{\varepsilon \to 0} \left| \frac{f(\varepsilon) - \sum_{j=0}^{N} a_j \delta_j(\varepsilon)}{\delta_N(\varepsilon)} \right| = 0.$$

- Usually write $f(\varepsilon) \sim \sum_{j=0}^{N} a_j \delta_j(\varepsilon)$, $\varepsilon \longrightarrow 0$.

In the case of perturbation solution to an ODE, we will consider asymptotic expansions of the form $x(t,\varepsilon) \sim \sum_k a_k(t)\delta_k(\varepsilon)$, which are valid over some interval of time. It is often useful to characterise the time interval, i.e. we say that the estimate is valid on a time scale $1/\hat{\delta}(\varepsilon)$ if the following holds:

$$\lim_{\varepsilon \to 0} \left| \frac{x(t,\varepsilon) - \sum_{k=0}^{N} a_k(t)\delta_k(\varepsilon)}{\delta_N(\varepsilon)} \right| = 0.$$

for $0 \leq \hat{\delta}(\varepsilon) \leq C$, with C independent of ε.

Example 3.1.3. Suppose $x(t,\varepsilon) = \varepsilon t \sin(t)$, $x \in \mathbb{R}$, $t \geq 0$, then

$$x(t,\varepsilon) = O(\varepsilon) \text{ for } t = O(1).$$
$$x(t,\varepsilon) = O(1) \text{ for } t \sim \frac{1}{\varepsilon}.$$

3.1.2 Naive Expansions

Consider $\dot{x} = f(t, x; \varepsilon)$, $x \in \mathbb{R}^n$, $x(0) = x_0$. Suppose that we can expand f as a Taylor series in ε, i.e.

$$f(t, x; \varepsilon) = f_0(t, x) + \varepsilon f_1(t, x) + \cdots .$$

One might expect a similar expansion exists for the solution:

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots . \quad \left[\delta_k(\varepsilon) = \varepsilon^k.\right]$$

Substitute this expansion into ODE and equate terms in equal powers of ε, this gives us an asymptotic approximation for $t = O(1)$.

Example 3.1.4. Consider $\ddot{x} + 2\varepsilon\dot{x} + x = 0$, $x(0) = a$, $\dot{x}(0) = 0$.

- This has explicit solution of the form:

$$x(t) = ae^{-\varepsilon t}\cos\left(t\sqrt{(1-\varepsilon^2)}\right) + \frac{\varepsilon a}{\sqrt{(1-\varepsilon^2)}}e^{-\varepsilon t}\sin\left(t\sqrt{(1-\varepsilon^2)}\right).$$

  There is 2 time scales here:

  - The slowly decaying exponential which varies on time scale of $\dfrac{1}{\varepsilon}$.
  - The oscillatory term which varies on time scale of 1.

- If we apply naive perturbation, we will pick up secular terms. More precisely,

$$e^{-\varepsilon t}\cos\left(t\sqrt{(1-\varepsilon^2)}\right) \approx \left(1 - \varepsilon t + \frac{\varepsilon^2 t^2}{2!}\right)\cos(t).$$

- Alternatively, assume that an expansion exists for the solution and set

$$x(t) = x_0(t) + \varepsilon x_1(t) + \cdots .$$

  Equating terms in equal powers of ε yields:

$$\begin{aligned} \ddot{x_0} + x_0 &= 0 \\ \vdots \quad \vdots \quad &\vdots \\ \ddot{x_n} + x_n &= -2\dot{x}_{n-1} \quad , \quad n = 1, 2, \cdots . \end{aligned}$$

  Solving these equations with initial conditions $x_0(0) = a$, $\dot{x_0}(0) = 0$; $x_n(0) = 0$, $\dot{x_n}(0) = 0$ gives:

$$\begin{aligned} x_0(t) &= a\cos(t) \\ x_1(t) &= a\sin(t) - at\cos(t). \end{aligned}$$

  - We see that $x_1(t)$ contains a secular term that grows linearly in t, making the asymptotic expansion valid for only small values of t.
  - Another way to interpret this is that the solution breaks down when $t \sim \dfrac{1}{\varepsilon}$, since $\varepsilon x_1(t)$ will be of the same order as $x_0(t)$.

3.2 Method of Multiple Scales

One often finds that solution oscillates on a time-scale of order t with an amplitude or phase that drifts on a slower time-scale εt. This suggest we should look for solutions of the form

$$R(\varepsilon t)\sin(t+\mathcal{O}(\varepsilon t)).$$

or equivalently

$$A(\varepsilon t)\sin(t)+B(\varepsilon t)\cos(t).$$

- In order to separate these time scales, we treat t and $\tau=\varepsilon t$ as independent variables, and introduce the following asymptotic expansion:

  $$X(t)=x(t,\tau,\varepsilon)\sim x_0(t,\tau)+\varepsilon x_1(t,\tau)+\cdots.$$

  which is asymptotic up to order $1/\varepsilon$. In general, it might be $\sqrt{\varepsilon}$ instead of ε

- $\tau=\varepsilon t$ is called the slow time scale because it does not affect the asymptotic expansion until t is comparable with $1/\varepsilon$.

- Using the chain rule, we have that $\dfrac{d}{dt}=\dfrac{\partial}{\partial t}+\varepsilon\dfrac{\partial}{\partial\tau}$. Thus,

$$\begin{aligned}
\dot{x} &= \frac{\partial x_0}{\partial t}+\varepsilon\frac{\partial x_0}{\partial\tau}+\varepsilon\left[\frac{\partial x_1}{\partial t}+\varepsilon\frac{\partial x_1}{\partial\tau}\right]+\dots\\
&= \frac{\partial x_0}{\partial t}+\varepsilon\left[\frac{\partial x_0}{\partial\tau}+\frac{\partial x_1}{\partial t}\right]+\dots\\
\ddot{x} &= \left[\frac{\partial^2 x_0}{\partial t^2}+\varepsilon\frac{\partial^2 x_0}{\partial\tau\partial t}\right]+\varepsilon\left[\frac{\partial^2 x_0}{\partial t\partial\tau}+\varepsilon\frac{\partial^2 x_0}{\partial\tau^2}\right]+\varepsilon\left[\frac{\partial^2 x_1}{\partial t^2}+\varepsilon\frac{\partial^2 x_1}{\partial\tau\partial t}\right]+\dots\\
&= \frac{\partial^2 x_0}{\partial t^2}+\varepsilon\left[2\frac{\partial^2 x_0}{\partial t\partial\tau}+\frac{\partial^2 x_1}{\partial t^2}\right]+\dots
\end{aligned}$$

Example 3.2.1. Consider the following initial value problem:

$$\begin{cases}\ddot{x}+\varepsilon\dot{x}(x^2-1)+x=0.\\ x(0)=1,\ \dot{x}(0)=0.\end{cases}$$

Suppose that $X(t)$ has the asymptotic expansion of the form:

$$X(t)=x(t,\varepsilon,\tau)\sim x_0(t,\tau)+\varepsilon x_1(t,\tau)$$

for $t=\mathcal{O}(1/\varepsilon)$. Substituting this expansion into the ODE and equate terms in equal powers of ε gives the following:

- Equation for $\mathcal{O}(1)$ term is:

  $$\begin{cases}\dfrac{\partial^2 x_0}{\partial t^2}+x_0=0.\\[2ex] x_0(0,0)=1,\ \dfrac{\partial x_0}{\partial t}(0,0)=0.\end{cases}$$

  This has solution $x_0(t,\tau)=R(\tau)\cos\Big(t+\theta(\tau)\Big)$, with $R(0)=1$, $\theta(0)=0$. Note that the constant might depends on τ. Also, when $t=0$, so is τ.

- Equation for $\mathcal{O}(\varepsilon)$ term is:

$$\begin{cases} \dfrac{\partial^2 x_1}{\partial t^2} + x_1 = -2\dfrac{\partial^2 x_0}{\partial t \partial \tau} - \dfrac{\partial x_0}{\partial t}(x_0^2 - 1). \\ x_1(0,0) = 0, \ \left(\dfrac{\partial x_0}{\partial \tau} + \dfrac{\partial x_1}{\partial t}\right)\Big|_{(0,0)} = 0. \end{cases}$$

  - Substituting for x_0 into the RHS yields:

$$\begin{aligned} -2\frac{\partial^2 x_0}{\partial t \partial \tau} - \frac{\partial x_0}{\partial t}(x_0^2 - 1) &= -2\frac{\partial}{\partial \tau}\Big[R\sin(t+\theta)\Big] \\ &\quad + R\sin(t+\theta)\Big[R^2\cos^2(t+\theta) - 1\Big] \\ &= 2\Big[R\theta_\tau \cos(t+\theta) + R_\tau \sin(t+\theta)\Big] \\ &\qquad + R\sin(t+\theta)\Big[R^2\cos^2(t+\theta) - 1\Big] \\ &= 2\Big[R\theta_\tau \cos(t+\theta) + R_\tau \sin(t+\theta)\Big] \\ &\qquad + R\sin(t+\theta)\Big[R^2\Big(1 - \sin^2(t+\theta)\Big) - 1\Big] \\ &= \Big[2R_\tau + R(R^2-1)\Big]\sin(t+\theta) + 2R\theta_\tau\cos(t+\theta) \\ &\qquad - R^3\sin^3(t+\theta). \end{aligned}$$

  - The first two terms are secular/resonant terms, which we need to eliminate. From this, we obtain two initial value problems for $R(\tau)$ and $\theta(\tau)$:

$$\begin{cases} 2R_\tau = R(1-R^2), \ R(0) = 1 \\ \theta_\tau = 0, \ \theta(0) = 0. \end{cases}$$

  The second equation immediately implies that $\theta(\tau) = 0$ for all $\tau \geq 0$.

  - In particular, we can solve for $R(\tau)$ explicitly by first using the substitution $y = R^2$:

$$\begin{aligned} 1 &= \frac{2}{R(1-R^2)}\frac{dR}{d\tau} \\ &= \frac{2}{R(1-R^2)}\left(\frac{1}{2R}\frac{dy}{d\tau}\right) \\ &= \frac{1}{y(1-y)}\frac{dy}{d\tau} \\ &= \left[\frac{1}{y} + \frac{1}{1-y}\right]\frac{dy}{d\tau} \end{aligned}$$

  Integrating both sides with respect to τ yields:

$$\begin{aligned} \ln|y| - \ln|1-y| &= \tau + A \\ \tau + A &= \ln\left|\frac{y}{1-y}\right| \end{aligned}$$

3.3 Averaging Theorem: Periodic Case

3.4 Phase Oscillators and Isochrones

3.5 Problems

1. Consider the equation
$$\ddot{x} + \dot{x} = -\varepsilon(x^2 - x), \quad 0 < \varepsilon \ll 1.$$
Using the method of multiple scales, show that
$$x_0(t, \tau) = A(\tau) + B(\tau)e^{-t},$$
and identify any resonant terms at $\mathcal{O}(\varepsilon)$. Show that the non-resonance condition is
$$A_\tau = A - A^2,$$
and describe the asymptotic behaviour of solutions.

Solution:

2. Consider the Van der Pol equation
$$\ddot{x} + x + \varepsilon(x^2 - 1)\dot{x} = \Gamma\cos(\omega t), \quad 0 < \varepsilon \ll 1,$$
with $\Gamma = \mathcal{O}(1)$ and $\omega \neq \frac{1}{3}, 1, 3$. Using the method of multiple scales, show that the solution is attracted to
$$x(t) = \frac{\Gamma}{1 - \omega^2}\cos(\omega t) + \mathcal{O}(\varepsilon)$$
when $\Gamma^2 \geq 2(1 - \omega^2)^2$ and
$$x(t) = 2\left(1 - \frac{\Gamma^2}{2(1 - \omega^2)^2}\right)^{1/2}\cos(t) + \frac{\Gamma}{1 - \omega^2}\cos(\omega t) + \mathcal{O}(\varepsilon)$$
when $\Gamma^2 < 2(1 - \omega^2)^2$. Explain why this result breaks down when $\omega = \frac{1}{3}, 1, 3$.

Solution:

3. Consider the following differential equation
$$\ddot{x} + x = -\varepsilon f(x, \dot{x}),$$
with $|\varepsilon| \ll 1$. Let $y = \dot{x}$.
   (a) Show that if $E(x, y) = (x^2 + y^2)/2$, then $\dot{E} = -\varepsilon f(x, y)y$. Hence show that an approximate periodic solution of the form $x = A\cos(t) + \mathcal{O}(\varepsilon)$ exists if
$$\int_0^{2\pi} f(A\cos(t), -A\sin(t))\sin(t)\,dt = 0.$$

Solution:

(b) Let $E_n = E(x(2\pi n), y(2\pi n))$ and $E_0 = E(x(0), y(0))$. Show that to lowest order E_n satisfies a difference equation of the form

$$E_{n+1} - E_n + \varepsilon F(E_n)$$

and write down $F(E_n)$ explicitly as an integral. Hence deduce that a periodic orbit with approximate amplitude $A^* = \sqrt{2E^*}$ exists if $F(E^*) = 0$ and this orbit is stable if

$$\varepsilon \frac{dF}{dE}(E^*) < 0.$$

Hint: Spiralling orbits close to the periodic orbit $x = A^* \cos(t) + \mathcal{O}(\varepsilon)$ can be approximated by a solution of the form $x = A\cos(t) + \mathcal{O}(\varepsilon)$.

Solution:

(c) Using the above result, find the approximate amplitude of the periodic orbit of the Van der Pol equation

$$\ddot{x} + x + \varepsilon(x^2 - 1)\dot{x} = 0,$$

and verify that it is stable.

Solution:

4. Consider the forced Duffing equation

$$\ddot{x} + \omega^2 x = \varepsilon[\gamma \cos(\omega t) - \kappa \dot{x} - \beta x - x^3].$$

Using the method of averaging with

$$\begin{cases} x(t) &= a(t)\cos(\omega t) + \dfrac{b(t)}{\omega}\sin(\omega t) \\ \dot{x}(t) &= -a(t)\omega \sin(\omega t) + b(t)\cos(\omega t), \end{cases}$$

derive dynamical equations for the amplitudes a, b and show that they are identical to the amplitude equations obtained using the method of multiple scales.

Solution:

Chapter 4

Boundary Value Problems

4.1 Compact Symmetric Linear Operators

Definition 4.1.1. Let V be a complex vector space. An **inner product** is a mapping $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}$ with the following properties

(a) $\langle \alpha f_1 + \beta f_2, g\rangle = \alpha^* \langle f_1, g\rangle + \beta^* \langle f_2, g\rangle$ for all $\alpha, \beta \in \mathbb{C}$.

(b) $\langle f, g\rangle = \langle g, f\rangle^*$.

(c) $\langle f, f\rangle = 0 \Longleftrightarrow f = \mathbf{0}$.

Associated with every inner product is an associated **norm** $\|f\| = \sqrt{\langle f, f\rangle}$. If V is complete with respect to the above norm, then V is a ***Hilbert space***.

Theorem 4.1.2 (Cauchy-Schwarz)**.** $|\langle f, g\rangle| \leq \|f\|\|g\|$.

Proof. The inequality is trivial if $g = \mathbf{0}$, so assume that $g \neq \mathbf{0}$. Define $h = f - \dfrac{\langle f, g\rangle}{\|g\|} g$. Since $\langle h, g\rangle = 0$, **Pythagorean theorem** yields

$$\begin{aligned}
\|f\|^2 &= \left\| h + \frac{\langle f, g\rangle g}{\|g\|} \right\|^2 \\
&= \left\langle h + \frac{\langle f, g\rangle g}{\|g\|}, h + \frac{\langle f, g\rangle g}{\|g\|} \right\rangle \\
&= \|h\|^2 + \frac{|\langle f, g\rangle|^2}{\|g\|^2} \\
&\geq \frac{|\langle f, g\rangle|^2}{\|g\|^2}.
\end{aligned}$$

■

Lemma 4.1.3 (Bessel's inequality)**.** *Suppose (u_j) is an orthonormal sequence in a Hilbert space H. The following inequality holds for any $f \in H$*

$$\sum_{j=1}^{\infty} |\langle u_j, f\rangle|^2 \leq \|f\|^2.$$

Proof. Suppose (u_j) is an orthonormal sequence in H. Any $f \in H$ can be written as

$$f = f_\perp + \sum_{j=1}^{n} \langle u_j, f\rangle u_j.$$

Note that for any fixed $k = 1, \ldots, n$,

$$\begin{aligned} \langle u_k, f_\perp\rangle &= \left\langle u_k, f - \sum_{j=1}^{n} \langle u_j, f\rangle u_j \right\rangle = \langle u_k, f\rangle - \langle u_k, f\rangle = 0. \\ \implies \|f\|^2 &= \|f_\perp\|^2 + \sum_{j,k=1}^{n} \langle u_j, f\rangle^* \langle u_k, f\rangle \langle u_j, u_k\rangle \\ &= \|f_\perp\|^2 + \sum_{j=1}^{n} |\langle u_j, f\rangle|^2 \\ &\geq \sum_{j=1}^{n} |\langle u_j, f\rangle|^2. \end{aligned}$$

The result follows by taking limit as $n \longrightarrow \infty$.

■

Definition 4.1.4.

(a) A **linear operator** is a linear mapping $A \colon \mathcal{D}(A) \longrightarrow H$, where H is a Hilbert space and $\mathcal{D}(A)$ is a linear subspace of H, called the domain of A.

(b) A linear operator A is **symmetric** if

  (i) its domain is dense, that is, $\overline{\mathcal{D}(A)} = H$, and

  (ii) $\langle g, Af\rangle = \langle Ag, f\rangle$ for all $f, g \in \mathcal{D}(A)$.

(c) A number $\lambda \in \mathbb{C}$ is called an **eigenvalue** of a linear operator A if there exists a non-zero vector $u \in \mathcal{D}(A)$, called an **eigenvector** of λ such that $Au = \lambda u$.

(d) An **eigenspace** is $\mathcal{N}(A - \lambda I) = \left\{ u \in \mathcal{D}(A) \colon (A - \lambda I)u = \mathbf{0} \right\}$.

Theorem 4.1.5. *Let A be a symmetric linear operator. Then all eigenvalues of A (if they exists) are real and eigenvectors of distinct eigenvalues are orthogonal.*

Proof. Suppose λ is an eigenvalue with unit normalised eigenvector u.

$$\lambda = \langle u, Au\rangle = \langle Au, u\rangle = \lambda^*$$

This implies that λ is real. If $Au_j = \lambda_j u_j$, $j = 1, 2$, then

$$(\lambda_1 - \lambda_2)\langle u_1, u_2\rangle = \langle Au_1, u_2\rangle - \langle u_1, Au_2\rangle = 0.$$

since A is symmetric. Thus, if $\lambda_1 - \lambda_2 \neq 0$, we must have $\langle u_1, u_2\rangle = 0 \implies u_1 \perp u_2$.

■

Definition 4.1.6 (Boundedness and Compactness)**.**

(a) A linear operator A with $\mathcal{D}(A) = H$ is said to be **bounded** if

$$\|A\| = \sup_{f\in\mathcal{D}(A),\|f\|=1} \|Af\| < \infty$$

- By construction, $\|Af\| \leq \|A\|\|f\|$, *i.e.* a bounded linear operator is continuous.

(b) A linear operator A with $\mathcal{D}(A) = H$ is said to be **compact** if for any bounded sequence $(f_n) \subset \mathcal{D}(A)$, the sequence (Af_n) has a convergent subsequence in H.

- It can be shown that every compact linear operator is bounded.

Theorem 4.1.7. *A compact symmetric operator A has an eigenvalue λ_0 for which $|\lambda_0| = \|A\|$.*

Proof. The result is trivial is $A = \mathbf{0}$, so suppose not. Denote $a = \|A\|$, with $a \neq 0$. By definition of $\|A\|$, we have that

$$\|A\|^2 = \sup_{\|f\|=1} \|Af\|^2 = \sup_{\|f\|=1} \langle Af, Af\rangle = \sup_{\|f\|=1} \langle f, A^2 f\rangle.$$

This implies that there exists a normalised sequence (u_n) such that $\lim_{n\to\infty} \langle u_n, A^2u_n\rangle = a^2$. Since A is compact, A^2 is also compact, so there exists a subsequence (u_{n_k}) such that $A^2u_{n_k} \longrightarrow y = a^2u$ in H as $k \longrightarrow \infty$ for some $u \in H$. Since A is symmetric,

$$\begin{aligned}
\|(A^2 - a^2)u_{n_k}\|^2 &= \|A^2u_{n_k}\|^2 - \langle A^2u_{n_k}, a^2u_{n_k}\rangle - \langle a^2u_{n_k}, A^2u_{n_k}\rangle + \|a^2u_{n_k}\|^2 \\
&= \|A^2u_{n_k}\|^2 - 2a^2\langle u_{n_k}, A^2u_{n_k}\rangle + a^4\|u_{n_k}\|^2 \\
&\leq (a^2)^2\|u_{n_k}\|^2 - 2a^2\langle u_{n_k}, A^2u_{n_k}\rangle + a^4\|u_{n_k}\|^2 && \Big[A \text{ is bounded}\Big] \\
&= 2a^4 - 2a^2\langle u_{n_k}, A^2u_{n_k}\rangle && \Big[\|u_{n_k}\|^2 = 1\Big] \\
&= 2a^2\Big[a^2 - \langle u_{n_k}, A^2u_{n_k}\rangle\Big] \longrightarrow 0 \text{ as } k \longrightarrow \infty
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\|a^2u_{n_k} - a^2u\| &\leq \|a^2u_{n_k} - A^2u_n\| + \|A^2u_{n_k} - a^2u\| \\
&\longrightarrow 0 \text{ as } k \longrightarrow \infty.
\end{aligned}$$

which implies that $u_{n_k} \longrightarrow u$ in $\mathcal{D}(A) = H$ as $k \longrightarrow \infty$. It follows that

$$\mathbf{0} = \lim_{k\to\infty} (A^2 - a^2)u_{n_k} = (A^2 - a^2)u = (A + aI)(A - aI)u,$$

There are two possible cases

1. Either $(A - aI)u = \mathbf{0} \implies \lambda_0 = a$,
2. or $(A - aI)u = v \neq \mathbf{0}$ and $(A + aI)v = \mathbf{0}$, which gives $\lambda_0 = -a$.

■

Remark 4.1.8. A bounded operator cannot have an eigenvalue $|\lambda| > \|A\|$ since

$$|\lambda|\|u\| = \|\lambda u\| = \|Au\| \leq \|A\|\|u\|.$$

Theorem 4.1.9 (Spectral Theorem for Compact Symmetric Linear Operators)**.** *Suppose H is a Hilbert space and $A\colon H \longrightarrow H$ is a compact symmetric linear operator.*

(a) There exists a sequence of real eigenvalues $(\lambda_n)_{n=0}^{\infty} \longrightarrow 0$. The corresponding normalised eigenvectors $\{u_0, u_1, \dots\}$ form an orthornormal set in H.

(b) Every $f \in \mathcal{R}(A)$ can be written as

$$f = \sum_{j=0}^{N} \langle u_j, f\rangle u_j. \tag{4.1.1}$$

(c) If $\mathcal{R}(A)$ is dense in H, then the set of normalised eigenvectors form an orthonormal basis of H.

Proof. We first establish existence of an orthonormal set of eigenvectors (eigenfunctions) using Theorem 4.1.7.

- Let $H^{(0)} = H$ and $A^0 = A|_{H^{(0)}}\colon H^{(0)} \longrightarrow H^{(0)}$. From Theorem 4.1.7, there exists an eigenvalue λ_0 and a normalised eigenvector $u_0 \in H^{(0)}$ such that

$$Au_0 = \lambda_0 u_0, \quad \text{with } |\lambda_0| = \|A^0\| = \|A\|.$$

- Define $H^{(1)} := (\text{span}(u_0))^{\perp} = \{f \in H\colon \langle f, u_0\rangle = 0\}$. Note that $H^{(1)} \subset H^{(0)}$.
  - Since A is symmetric, for any $f \in H^{(1)}$ we have that

$$\langle Af, u_0\rangle = \langle f, Au_0\rangle = \lambda_0\langle f, u_0\rangle = 0.$$

    This means that $Af \in H^{(1)}$.
  - Let $A_1 = A|_{H^{(1)}}\colon H^{(1)} \longrightarrow H^{(1)}$ be the restriction of A to $H^{(1)}$, this is again a compact symmetric linear operator. From Theorem 4.1.7, there exists an eigenvalue λ_1 and a normalised eigenvector $u_1 \in H^{(1)}$ such that

$$A_1u_1 = \lambda_1 u_1, \quad \text{with } |\lambda_1| = \|A_1\|.$$

    Moreover, we have that $\langle u_1, u_0\rangle = 0$ since $u_1 \in H^{(1)}$.

By iterating this argument, we construct a family of spaces $H^{(n)}$ and compact symmetric linear operators $A^{(n)}$, together with a sequence of eigenvalues (λ_j) with corresponding normalised eigenvectors (u_j), satisfying

1. $\dots H^{(n)} \subset H^{(n-1)} \subset \dots \subset H^{(0)} = H$.
2. $H^{(n)} = \Big[\text{span}(u_0, \cdots, u_{n-1})\Big]^{\perp}$ for all $n \geq 1$.

3. $A_n \colon H^{(n)} \longrightarrow H^{(n)}$, with $A_n u_n = \lambda_n u_n$ and $|\lambda_n| = \|A_n\|$ for each $n \geq 0$.

Moreover, the nested property of $H^{(n)}$ implies that $\langle u_k, u_j \rangle = 0$ for all $0 \leq k < j$. This iterative procedure does not terminate unless $\mathcal{R}(A)$ is finite-dimensional. Indeed, if $\mathcal{R}(A)$ is of dimension $n < \infty$, then

$$A_n = A|_{H^{(n)}} = \mathbf{0} \implies \lambda_j = 0 \quad \text{for all } j \geq n.$$

Now, suppose $\lambda_j \not\to 0$. There exists a subsequence λ_{j_k} and $\varepsilon > 0$ such that $|\lambda_{j_k}| > \varepsilon$ for all $k \geq 1$. Consider the sequence (v_k) defined by $v_k = \dfrac{u_{j_k}}{\lambda_{j_k}}$, which is well-defined since $\lambda_{j_k} \neq 0$. Observe that (v_k) is a bounded sequence since (u_{j_k}) has unit norm and (λ_{j_k}) is bounded below. Moreover, for any $k \neq l$ we have that

$$\begin{aligned}
\|Av_k - Av_l\|^2 &= \left\| \frac{Au_{j_k}}{\lambda_{j_k}} - \frac{Au_{j_l}}{\lambda_{j_l}} \right\|^2 \\
&= \|u_{j_k} - u_{j_l}\|^2 && \Big[\text{since } Au_n = \lambda_n u_n\Big] \\
&= \|u_{j_k}\|^2 + \|u_{j_l}\|^2 && \Big[\text{since } \langle u_m, u_n \rangle = \delta_{mn}\Big] \\
&= 2 && \Big[\text{since } \|u_n\| = 1\Big]
\end{aligned}$$

Thus, any subsequence of (Av_k) is not a Cauchy sequence and does not converge. This contradicts the compactness of the linear operator A since (v_k) is a bounded sequence.

Next, let $f = Ag \in \mathcal{R}(A)$, and denote $f_n = \sum_{j=0}^{n} \langle u_j, f \rangle u_j$. Observe that

$$\begin{aligned}
f_n = \sum_{j=0}^{n} \langle u_j, f \rangle u_j &= \sum_{j=0}^{n} \langle u_j, Ag \rangle u_j \\
&= \sum_{j=0}^{n} \langle Au_j, g \rangle u_j && \Big[\text{since } A \text{ is symmetric}\Big] \\
&= \sum_{j=0}^{n} \lambda_j^* \langle u_j, g \rangle u_j \\
&= \sum_{j=0}^{n} \lambda_j \langle u_j, g \rangle u_j && \Big[\text{from Theorem 4.1.5}\Big] \\
&= \sum_{j=0}^{n} \langle u_j, g \rangle Au_j && \Big[\text{since } Au_j = \lambda_j u_j\Big] \\
&= A \left[\sum_{j=0}^{n} \langle u_j, g \rangle u_j \right] && \Big[\text{since } A \text{ is linear}\Big] \\
&= Ag_n.
\end{aligned}$$

For any $k = 0, \dots, n$, we also have that

$$\langle u_k, g - g_n \rangle = \langle u_k, g \rangle - \left\langle u_k, \sum_{j=0}^{n} \langle u_j, g \rangle u_j \right\rangle = \langle u_k, g \rangle - \langle u_k, g \rangle = 0,$$

i.e. $g - g_n \in \Big[\text{span}(u_0, \cdots, u_n)\Big]^{\perp} = H^{(n+1)}$. There are two possible cases:

1. $\underline{\dim(\mathcal{R}(A)) = n + 1 < \infty}$.
   This means that $A^{n+1} = \mathbf{0}$. Since $g - g_n \in H^{(n+1)}$, we have that
   $$f - f_n = Ag - Ag_n = A(g - g_n) = A_{n+1}(g - g_n) = \mathbf{0}.$$
2. $\underline{\dim(\mathcal{R}(A)) = \infty}$.
   Since $g - g_n \in H^{(n+1)}$, **Pythagorean theorem** yields
   $$\|g\|^2 = \|g - g_n + g_n\|^2 = \|g - g_n\|^2 + \|g_n\|^2 \geq \|g - g_n\|^2.$$
   Boundedness of A_{n+1} gives
   $$\|f - f_n\|^2 = \|A_{n+1}(g - g_n)\|^2 \leq \|A_{n+1}\|^2 \|g - g_n\|^2 \leq |\lambda_{n+1}|^2 \|g\|^2 \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

In both cases, we prove that $f_n \longrightarrow f$ and so (4.1.1) holds.

Finally, let $f \in H$ and suppose $\mathcal{R}(A)$ is dense in H. Given any $\varepsilon > 0$, there exists $f_\varepsilon \in \mathcal{R}(A)$ such that $\|f - f_\varepsilon\| < \varepsilon$. Moreover, from previous step, there exists $\hat{f}_\varepsilon \in \text{span}(u_0, \cdots, u_n)$ for sufficiently large n such that $\|f_\varepsilon - \hat{f}_\varepsilon\| < \varepsilon$. Thus,
$$\|f - \hat{f}_\varepsilon\| \leq \|f - f_\varepsilon\| + \|f_\varepsilon - \hat{f}_\varepsilon\| < 2\varepsilon$$
for sufficiently large n.

■

Remark 4.1.10. We would like to carry out a similar analysis to derive a spectral theorem for differential operator. The problem is that symmetric differential operator are not compact, or even bounded. Fortunately, one can construct an inverse of a differential operator that is compact (Green's functions and integral equations).

4.2 Linear Differential Operators

Consider the **formal linear differential operator** on $[a, b] \subset \mathbb{R}$ defined by
$$L = p_0(x)\frac{d^n}{dx^n} + p_1(x)\frac{d^{n-1}}{dx^{n-1}} + \ldots + p_n(x). \tag{LDO}$$
where we are not concerned about the space of functions it is applied to. A natural choice of function space is the Hilbert space $L^2[a, b]$, since we can work with inner product and orthogonality.

- First, note that L cannot act on all functions in $L^2[a, b]$ because not all of them are differentiable. Because of this, we require that our **domain** $\mathcal{D}(L)$ contains only functions that are sufficiently differentiable, but the image of these functions under L need only be in $L^2[a, b]$.
- Usually, we further restrict $\mathcal{D}(L)$ by imposing boundary conditions at the endpoints of the interval. The boundary conditions that we will impose will always be *linear* and *homogeneous*, this will make $\mathcal{D}(L)$ a vector space.

4.2.1 Formal Adjoint

With the above linear differential operator L, consider a **weight function** $\omega(x)$ which is defined to be a real and positive function on (a, b). We can construct another operator $L^\dagger$ such that the following holds for any smooth function $u(x)$, $v(x)$

$$\omega\Big[u^* L v - (L^\dagger u)^* v\Big] = \frac{d}{dx}\Big(Q[u, v]\Big), \tag{4.2.1}$$

for some function $Q[\cdot, \cdot]$ that depends bilinearly on its arguments and their first $n-1$ derivatives. (4.2.1) is called the **Lagrange's identity** and $L^\dagger$ is called the **formal adjoint** of L with respect to the weight function ω. Note that we are yet to specify $\mathcal{D}(L^\dagger)$.

Now, consider a **weighted inner product** defined by

$$\langle u, v\rangle_\omega = \int_a^b \omega u^* v\, dx.$$

If u and v have boundary conditions such that $Q[u, v]|_a^b = 0$, then

$$\langle u, Lv\rangle_\omega = \langle L^\dagger u, v\rangle_\omega$$

which looks similar to how we define an adjoint operator on a normed vector space! A common method for finding the formal adjoint is using **integration by parts**.

Example 4.2.1 (Sturm-Liouville operator)**.** Consider the following linear differential operator

$$L = p_0(x)\frac{d^2}{dx^2} + p_1(x)\frac{d}{dx} + p_2(x),$$

where $p_j(x)$ are all real and $\omega(x) \equiv 1$. Integrating by parts twice yields

$$\begin{aligned}
\langle u, Lv\rangle &= \int_a^b u^*(x)\left[p_0(x)\frac{d^2 v}{dx^2} + p_1(x)\frac{dv}{dx} + p_2(x)v(x)\right]\, dx \\
&= \left[u^*(x)p_0(x)\frac{dv}{dx}\right]\Bigg|_a^b - \int_a^b \frac{d}{dx}\Big(p_0(x)u^*(x)\Big)\frac{dv}{dx}\, dx \\
&\quad + \Big[u^*(x)p_1(v)v(x)\Big]\Bigg|_a^b - \int_a^b \frac{d}{dx}\Big(p_1(x)u^*(x)\Big)v(x)\, dx \\
&\quad + \int_a^b u^*(x)p_2(x)v(x)\, dx \\
&= \int_a^b \left[\frac{d^2}{dx^2}\Big(p_0(x)u(x)\Big) - \frac{d}{dx}\Big(p_1(x)u(x)\Big) + p_2(x)u(x)\right]^* v(x)\, dx \\
&\quad + \left[u^*(x)p_0(x)\frac{dv}{dx} - \frac{d}{dx}\Big(p_0(x)u^*(x)\Big)v(x) + u^*(x)p_1(x)v(x)\right]\Bigg|_a^b
\end{aligned}$$

Thus, we see that

$$L^\dagger = \frac{d^2}{dx^2}p_0 - \frac{d}{dx}p_1 + p_2$$

$$= p_0 \frac{d^2}{dx^2} + (2p_0' - p_1)\frac{d}{dx} + p_0'' - p_1' + p_2.$$

In order for L to be **formally self-adjoint**, we choose

$$\begin{cases} 2p_0' - p_1 = p_1 & \implies p_0' = p_1 \\ p_0'' - p_1' + p_2 = p_2 & \implies p_0'' = p_1' \implies p_0' = p_1 \end{cases}$$

Thus, by choosing $p_1 = p_0'$, we have the well-known **Sturm-Liouville operator** given by

$$L = \frac{d}{dx}\left(p_0(x)\frac{d}{dx}\right) + p_2(x).$$

Remark 4.2.2. What happens if $p_1 \neq p_0'$? We can still make L formally self-adjoint by choosing a suitable weight function $\omega(x)$. Suppose $p_0(x)$ is positive definite on (a, b), *i.e.* $p_0(x) > 0$ and $p_j(x)$ are all real. Define

$$\omega(x) = \frac{1}{p_0(x)} \exp\left[\int_a^x \frac{p_1(x')}{p_0(x')}\, dx'\right].$$

We can rewrite L in the following form

$$Lu = \frac{1}{\omega}\frac{d}{dx}\left(\omega p_0 \frac{du}{dx}\right) + p_2 u.$$

One can show that $L^\dagger = L$. Thus, **as long as $\boldsymbol{p_0(x) \neq 0}$**, we can always define a weighted inner product in which a real second-order differential operator is formally self-adjoint with respect to this weighted inner product.

4.2.2 A Simple Eigenvalue Problem

Recall that a finite self-adjoint (Hermitian) matrix has a complete set of orthonormal eigenvectors. Does the same property hold for self-adjoint differential operators?

Example 4.2.3. Consider the differential operator $L = -\frac{d^2}{dx^2}$, with domain

$$\mathcal{D}(L) = \left\{f, Lf \in L^2[0,1] \colon f(0) = f(1) = 0\right\}.$$

Integrating by parts yields

$$\begin{aligned} \langle f_1, Lf_2 \rangle &= \int_0^1 f_1^* \left(-\frac{d^2 f_2}{dx^2}\right) dx \\ &= \int_0^1 \left(-\frac{d^2 f_1}{dx^2}\right)^* f_2\, dx + \left[- f_1^* f_2' + (f_1^*)' f_2 \right]\Big|_0^1 \\ &= \langle Lf_1, f_2 \rangle \end{aligned}$$

We see that L is self-adjoint with the given boundary conditions.

The eigenvalue equation $L\psi = \lambda\psi$ or $\frac{d^2\psi}{dx^2} + \lambda\psi = 0$, with the boundary conditions $\psi(0) = \psi(1) = 0$ has solutions $\psi_n(x) = \sin(n\pi x)$ with eigenvalues $\lambda_n = n^2\pi^2$. Observe that

1. The eigenvalues are real.

2. The eigenfunctions are orthogonal, that is
$$2\int_0^1 \sin(m\pi x)\sin(n\pi x)\,dx = \delta_{mn}.$$

3. The normalised eigenfunctions are complete, *i.e.* any $f \in L^2[0,1]$ can be written as an L^2 convergent series of the following form
$$f(x) = \sum_{n=1}^{\infty} a_n\sqrt{2}\sin(n\pi x).$$

Example 4.2.4. Consider the differential operator $L = -i\dfrac{d}{dx}$, with domain
$$\mathcal{D}(L) = \Big\{f, Lf \in L^2[0,1]\colon f(0) = f(1) = 0\Big\}.$$

We claim that L is self-adjoint. Indeed, integrating by parts gives
$$\begin{aligned}\langle f_1, Lf_2\rangle &= \int_0^1 f_1^*\left(-i\frac{df_2}{dx}\right)\,dx \\ &= \Big[-if_1^*f_2\Big]\Big|_0^1 + i\int_0^1 \frac{df_1^*}{dx}f_2\,dx \\ &= \langle Lf_1, f_2\rangle.\end{aligned}$$

On the other hand, the eigenvalue equation $-i\dfrac{d\psi}{dx} = \lambda\psi$ or $\dfrac{d\psi}{dx} = i\lambda\psi$ has solution $\psi(x) \sim e^{i\lambda x}$. However, $\psi(x)$ does not satisfy the boundary conditions as exponential function is a positive function. Thus, $L = -i\dfrac{d}{dx}$ doesn't have an eigenvalue! This suggest that the problem lies in the boundary conditions.

4.2.3 Adjoint Boundary Conditions

To rectify the problem, we remove the requirement that $\mathcal{D}(L) = \mathcal{D}(L^\dagger)$, but still ensuring that $\mathcal{D}(L^\dagger)$ is defined such that $Q[u,v] = 0$ in (4.2.1). More precisely, given $v \in \mathcal{D}(L)$ satisfying certain boundary conditions, we impose any required boundary conditions on $u \in \mathcal{D}(L^\dagger)$ such that $Q[u,v] = 0$ in (4.2.1). Such boundary conditions on $u \in \mathcal{D}(L^\dagger)$ are called the **adjoint boundary conditions** and they define $\mathcal{D}(L^\dagger)$.

Example 4.2.5. Consider the differential operator $L = -i\dfrac{d}{dx}$, with domain
$$\mathcal{D}(L) = \Big\{f, Lf \in L^2[0,1]\colon f(1) = 0\Big\}.$$

From previous computation, we find that
$$\langle u, Lv\rangle - \langle L^\dagger u, v\rangle = -i\Big[u^*(1)v(1) - u^*(0)v(0)\Big]$$

The first term vanishes since $v(1) = 0$. Because $v(0)$ can take any value, we must choose $u(0) = 0$ in order for $Q[u, v]$ to vanish. Imposing this boundary condition on $u \in \mathcal{D}(L^\dagger)$ thus yields $L^\dagger = L = -i\dfrac{d}{dx}$ and

$$\mathcal{D}(L^\dagger) = \Big\{ f, Lf \in L^2[0,1] \colon f(0) = 0 \Big\}.$$

We remark that although L is formally self-adjoint, $L, L^\dagger$ are not the same operator since $\mathcal{D}(L) \neq \mathcal{D}(L^\dagger)$. We then say that L is not truly self-adjoint.

Example 4.2.6. Consider the same differential operator $L = -i\dfrac{d}{dx}$, but now with a different domain

$$\mathcal{D}(L) = \Big\{ f, Lf \in L^2[0,1] \colon f(0) = f(1) = 0 \Big\}.$$

It is clear that $Q[u, v] = 0$ irrespective of any boundary conditions on u. That is, there is no constraint on u. Thus, $L^\dagger = L$ and

$$\mathcal{D}(L^\dagger) = \Big\{ f, Lf \in L^2[0,1] \Big\}.$$

Again, L is not truly self-adjoint since $\mathcal{D}(L) \neq \mathcal{D}(L^\dagger)$.

4.2.4 Self-Adjoint Boundary Conditions

Definition 4.2.7. A formally self-adjoint operator L is **truly self-adjoint** if and only if $L = L^\dagger$ and $\mathcal{D}(L) = \mathcal{D}(L^\dagger)$. Thus, one needs to impose **self-adjoint boundary conditions** in order to obtain a self-adjoint operator.

Example 4.2.8. Consider the differential operator $L = -i\dfrac{d}{dx}$. We investigate what boundary conditions we need to impose so that L is truly self-adjoint. Previous calculation shows that

$$\langle u, Lv \rangle - \langle Lu, v \rangle = -i\Big[u^*(1)v(1) - u^*(0)v(0) \Big].$$

Demanding the right hand side vanishes gives us the condition

$$\frac{u^*(1)}{u^*(0)} = \frac{v(0)}{v(1)}.$$

We require this to be true for any u, v obeying the same boundary conditions.

Since u, v are not a-priori related, it follows that

$$\frac{u^*(1)}{u^*(0)} = \Lambda = \frac{v(0)}{v(1)}.$$

Since we require that both u, v belong to the same domain, we also have

$$\Lambda^* = \frac{u(1)}{u(0)} = \frac{v(1)}{v(0)} = \Lambda^{-1}.$$

i.e. $\Lambda^* = \Lambda^{-1} \implies \Lambda = e^{i\theta}$ for some real phase $\theta \in \mathbb{R}$. The domain is therefore

$$\mathcal{D}(L) = \Big\{ f, Lf \in L^2[0,1] \colon f(1) = e^{i\theta} f(0) \Big\}.$$

These are called **twisted periodic boundary conditions**. With this self-adjoint boundary conditions, everything is groovy:

1. The eigenfunctions are $\psi_n(x) = e^{(2\pi n+\theta)ix}$ with eigenvalues $\lambda_n = 2\pi n + \theta$.
2. The eigenvalues are real and the normalised eigenfunctions (ψ_n) form a complete orthonormal set.

Example 4.2.9. Consider the Sturm-Liouville differential operator defined by

$$L = \frac{d}{dx}\left(p(x)\frac{d}{dx} \right) + q(x), \quad x \in [a,b].$$

We shown previously that this is formally self-adjoint, satisfying

$$\langle u, Lv \rangle - \langle Lu, v \rangle = \Big[p\Big(u^* v' - (u^*)' v \Big) \Big] \Big|_a^b.$$

As before, demanding the right hand side vanishes gives us conditions at both ends:

$$\frac{(u^*)'(a)}{u^*(a)} = \frac{v'(a)}{v(a)}, \quad \frac{(u^*)'(b)}{u^*(b)} = \frac{v'(b)}{v(b)}.$$

To ensure $\mathcal{D}(L) = \mathcal{D}(L^\dagger)$, we also require that

$$\frac{v'(a)}{v(a)} = \frac{u'(a)}{u(a)} = \Lambda_a, \quad \frac{v'(b)}{v(b)} = \frac{u'(b)}{u(b)} = \Lambda_b.$$

Comparing these two conditions, we thus require that $\Lambda_a = \Lambda_a^*$ and $\Lambda_b = \Lambda_b^*$, *i.e.* $\Lambda_a, \Lambda_b \in \mathbb{R}$. Thus, we have the following self-adjoint boundary conditions:

$$\begin{aligned} \Lambda_a f(a) - f'(a) &= 0 \\ \Lambda_b f(b) - f'(b) &= 0. \end{aligned}$$

4.3 Eigenvalue Problems and Spectral Theory

It can be proven that a truly self-adjoint operator T with respect to $L^2[a,b]$ inner product posseses a complete set of mutually orthogonal eigenfunctions. The set of eigenvalues belong to the spectrum of T, denoted $\sigma(T)$. It forms the **point spectrum** provided their corresponding eigenfunctions belong to $L^2[a,b]$. Usually, the eigenvalues of the point spectrum form a discrete set, so the point spectrum is also known as the **discrete spectrum**. When a differential operator acts on functions on $\mathbb{R}$, the functions may fail to be normalisable; the associated eigenvalues then belong to the **continuous spectrum**. The spectrum maybe partially discrete and partially continuous. There can also be a **residual spectrum**, but this is empty for self-adjoint operators.

4.3.1 Discrete Spectrum

Let T be a self-adjoint linear operator on a bounded domain $[a, b]$. The eigenvalue equation is

$$T\phi_n(x) = \lambda_n \phi_n(x)$$

for integer $n \in \mathbb{Z}$. We also imposed the normalised condition

$$\int_a^b \phi_n^*(x)\phi_m(x)\,dx = \delta_{nm}.$$

Completeness of these eigenfunctions can be expressed by the following condition

$$\sum_{n\in\mathbb{Z}} \phi_n(x)\phi_n^*(y) = \delta(x-y), \tag{4.3.1}$$

where convergence is understood in the sense of distributions. More precisely, in order to make sense of this notion of convergence, we have to multiply (4.3.1) by a smooth **test function** $f(y)$ and integrate with respect to y, which gives

$$f(x) = \sum_{n\in\mathbb{Z}} \phi_n(x) \int_a^b \phi_n^*(y) f(y)\,dy = \sum_{n\in\mathbb{Z}} a_n\phi_n(x),$$

where $a_n = \displaystyle\int_a^b \phi_n^*(x')f(x')dx'$. The upshot is we can expand any L^2 functions if we can represent a delta function in terms of the eigenfunctions $\phi_n(x)$.

4.3.2 Continuous Spectrum

Consider the linear differential operator $H = -\dfrac{d^2}{dx^2}$ on $\left[-\dfrac{L}{2}, \dfrac{L}{2}\right]$. This operator has eigenfunctions $\phi_k(x) = e^{ikx}$ corresponding to eigenvalues $\lambda_k = k^2$. Suppose we impose periodic boundary conditions at $x = \pm L/2$. This means that

$$\phi_k(-L/2) = \phi_k(L/2) \implies k = k_n = \frac{2\pi n}{L},\ n \in \mathbb{Z}.$$

To find the normalised eigenfunctions,

$$\int_{-L/2}^{L/2} \phi_n^*(x)\phi_m(x)\,dx = \int_{-L/2}^{L/2} e^{i\frac{2\pi}{L}(m-n)x}\,dx = \begin{cases} L & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases} \implies \phi_n(x) = \frac{1}{\sqrt{L}}e^{ik_nx}.$$

Referring to (4.3.1), the completeness condition is

$$\sum_{n=-\infty}^{\infty} \frac{1}{L}e^{ik_nx}e^{-ik_ny} = \delta(x-y), \quad x, y \in \left[-\frac{L}{2}, \frac{L}{2}\right].$$

As $L \longrightarrow \infty$, the eigenvalues become so close together, they can no longer be distinguished and $\sigma(H) \longrightarrow \mathbb{R}$. Moreover, the sum over n becomes an integral

$$\sum_{n=-\infty}^{\infty} \frac{1}{L}e^{ik_nx}e^{-ik_ny} \longrightarrow \int \left(\frac{1}{L}e^{ikx}e^{-iky}\right)\,dn = \int \left(\frac{1}{L}e^{ikx}e^{-iky}\right)\left(\frac{dn}{dk}\right)\,dk,$$

where $\dfrac{dn}{dk} = \dfrac{L}{2\pi}$ is called the **density of states**. Note that $\dfrac{dn}{dk}$ is the **Radon-Nikodym** derivative, *i.e.* it is a ratio of measures. Thus, in the limit as $L \longrightarrow \infty$, the completeness condition becomes:

$$\int_{-\infty}^{\infty} e^{ik(x-y)}\,\frac{dk}{2\pi} = \delta(x-y).$$

When $L = \infty$, $\phi_k(x) \sim e^{ikx}$ is no longer normalisable on $L^2(\mathbb{R})$ since

$$\int_{-\infty}^{\infty} \phi_k^*(x)\phi_k(x)\,dx = \int_{-\infty}^{\infty} e^{-ikx}e^{ikx}\,dx = \int_{-\infty}^{\infty} dx = \infty!$$

Strictly speaking, $\phi_k(x)$ is not a true eigenfunction and points in the continuous spectrum are not actually eigenvalues. More rigorously, a point λ lies in the continuous spectrum if for any $\varepsilon > 0$, there exists an **approximate eigenfunction** $\phi_\varepsilon \in L^2(\mathbb{R})$ such that $\|\phi_\varepsilon\| = 1$ and $\|H\phi_\varepsilon - \lambda\phi_\varepsilon\| < \varepsilon$. In other words, the inverse operator (**resolvent**) $(H-\lambda I)^{-1}$ is unbounded.

4.3.3 Mixed Spectrum

Consider the following eigenvalue equation, known as the **Poschel-Teller equation**

$$H\psi = \left[-\frac{d^2}{dx^2} - 2\mathrm{sech}^2(x)\right]\psi = \lambda\psi, \quad x \in \mathbb{R}.$$

We try a solution of the form $\psi_k(x) = \Big[a + b\tanh(x)\Big]e^{ikx}$. This gives

$$\left[-\frac{d^2}{dx^2} - 2\mathrm{sech}^2(x)\right]\psi_k(x) = k^2\psi_k(x) - 2(ikb+a)\mathrm{sech}^2(x)e^{ikx}.$$

Thus, $H\psi_k = k^2\psi_k$ provided $ikb + a = 0 \implies \psi_k(x) = b\Big[-ik + \tanh(x)\Big]e^{ikx}$.

For $k \neq 0$, the formal normalised eigenfunctions are

$$\psi_k(x) = \frac{1}{\sqrt{1+k^2}}e^{ikx}\Big[-ik + \tanh(x)\Big].$$

It is constructed in such a way that

$$\begin{aligned}
\psi_k(x)\psi_k^*(x') &= \frac{1}{\sqrt{1+k^2}}e^{ikx}\Big[-ik+\tanh(x)\Big]\frac{1}{\sqrt{1+k^2}}e^{-ikx'}\Big[ik+\tanh(y)\Big] \\
&= \frac{1}{1+k^2}e^{ik(x-x')}\Big[-ik+\tanh(x)\Big]\Big[ik+\tanh(x')\Big] \\
&\longrightarrow \frac{1}{1+k^2}e^{ik(x-x')}(-ik+1)(ik+1) \text{ as } |x| \longrightarrow \infty \\
&= \frac{1+k^2}{1+k^2}e^{ik(x-x')} = e^{ik(x-x')}.
\end{aligned}$$

Thus, $\sigma(H)$ contains a continuous part with $\lambda_k = k^2$ and $\psi_k(x)$ as "eigenfunctions". For the special case $\lambda_k = 0$,

$$\frac{d^2\psi_0(x)}{dx^2} = -2\text{sech}^2(x)\psi_0(x) \implies \psi_0(x) = \frac{1}{\sqrt{2}}\text{sech}(x).$$

This is normalisable, *i.e.* $\lambda_0 = 0$ belongs to the discrete spectrum.

Let us compute the difference:

$$\begin{aligned} I = \delta(x-x') - \int_{-\infty}^{\infty} \psi_k(x)\psi_k^*(x')\,\frac{dk}{2\pi} &= \int_{-\infty}^{\infty} \left[e^{ik(x-x')} - \psi_k(x)\psi_k^*(x')\right]\frac{dk}{2\pi} \\ = \int_{-\infty}^{\infty} \frac{e^{ik(x-x')}}{1+k^2}\Big[(1+k^2) - k^2 + ik\Big(\tanh(x') - \tanh(x)\Big) &- \tanh(x)\tanh(x')\Big]\frac{dk}{2\pi}. \end{aligned}$$

It can be shown (by residue theorem) that

$$\begin{aligned} \int_{-\infty}^{\infty} \left(\frac{e^{ik(x-x')}}{1+k^2}\right)\frac{dk}{2\pi} &= \frac{1}{2}e^{-|x-x'|} \\ \int_{-\infty}^{\infty} \left(\frac{e^{ik(x-x')}ik}{1+k^2}\right)\frac{dk}{2\pi} &= -\frac{1}{2}e^{-|x-x'|}\text{sgn}(x-x'). \end{aligned}$$

Hence, $I = \dfrac{1}{2}\Big[1 - \text{sgn}(x-x')\Big(\tanh(x') - \tanh(x)\Big) - \tanh(x)\tanh(x')\Big]e^{-|x-x'|}$. WLOG, take $x > x'$. Then

$$\begin{aligned} I &= \frac{1}{2}\Big[1 - \Big(\tanh(x') - \tanh(x)\Big) - \tanh(x)\tanh(x')\Big]e^{-(x-x')} \\ &= \frac{1}{2}\Big[1 + \tanh(x) - \tanh(x') - \tanh(x)\tanh(x')\Big]e^{-(x-x')} \\ &= \frac{1}{2}\Big[1 + \tanh(x)\Big]\Big[1 - \tanh(x')\Big]e^{-(x-x')} \\ &= \frac{1}{2}\text{sech}(x)\text{sech}(x') \\ &= \psi_0(x)\psi_0(x'). \end{aligned}$$

Thus, we have the completeness condition

$$\psi_0(x)\psi_0(x') + \int_{-\infty}^{\infty} \psi_k^*(x)\psi_k(x')\,\frac{dk}{2\pi} = \delta(x-x').$$

4.3.4 Rayleigh-Ritz Variational Principle

Consider the Sturm-Liouville operator

$$L\phi = -\frac{d}{dx}\left(p\frac{d\phi}{dx}\right) + q\phi, \quad x \in [a,b], \tag{4.3.2}$$

together with self-adjoint boundary conditions

$$\alpha\phi(a) + \beta\phi'(a) = 0 \tag{4.3.3a}$$

$$\hat{\alpha}\phi(b) + \hat{\beta}\phi'(b) = 0. \tag{4.3.3b}$$

where p, q are real-valued functions. Suppose the following two statements hold:

1. There exists a countable infinite set of eigenvalues (λ_n) of (4.3.2) that are real-valued and non-negative such that

$$0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots\ldots$$

2. The corresponding eigenfunctions form a complete orthonormal set.

$$\langle \phi_n, \phi_m \rangle = \int_a^b \phi_n^*(x)\phi_m(x)\, dx = \delta_{mn}.$$

Using these, we derive a variational principle characterising eigenvalues of (4.3.2).

Consider a sufficiently smooth, bounded function u with the following properties:

$$u(x) = \sum_{k\in\mathbb{Z}^+} u_k\phi_k(x), \quad u_k = \langle u, \phi_k \rangle, \quad \|u\|^2 = 1. \tag{4.3.4}$$

Introducing the **energy integral** $E[u] = \langle u, Lu \rangle$. Substituting the Fourier series expansion (4.3.4) of u into $E[u]$ yields

$$\begin{aligned} E[u] = \langle u, Lu \rangle &= \int_a^b \left[\sum_{k\in\mathbb{Z}^+} u_k^*\phi_k^*(x)\right] L \left[\sum_{j\in\mathbb{Z}^+} u_j\phi_j(x)\right] dx \\ &= \sum_{k,j\in\mathbb{Z}^+} u_k^* u_j \left[\int_a^b \phi_k^*(x)\Big(L\phi_j(x)\Big)\right] dx \\ &= \sum_{k,j\in\mathbb{Z}^+} u_k^* u_j \int_a^b \phi_k^*(x)\lambda_j\phi_j(x)\, dx \\ &= \sum_{k,j\in\mathbb{Z}^+} u_k^* u_j \Big[\lambda_j\langle \phi_k, \phi_j \rangle\Big] \\ &= \sum_{k\in\mathbb{Z}^+} \lambda_k |u_k|^2. \end{aligned}$$

From the ordering of the eigenvalues,

$$E[u] = \sum_{k\in\mathbb{Z}^+} \lambda_k|u_k|^2 \geq \lambda_1 \sum_{k\in\mathbb{Z}^+} |u_k|^2 = \lambda_1 \sum_{k\in\mathbb{Z}^+} |\langle u, \phi_k \rangle|^2 = \lambda_1\|u\|^2 = \lambda_1,$$

where we use the **Parseval's identity** and the assumption that all eigenvalues are non-negative. We see that the smallest eigenvalue λ_1 is obtained by minimising $E[u]$ with respect to all admissible functions u satisfying $\|u\|^2 = 1$. Moreover, a minimum occurs at $u = \phi_1$, since

$$E[\phi_1] = \sum_{k\in\mathbb{Z}^+} \lambda_k |\langle \phi_1, \phi_k \rangle|^2 = \lambda_1.$$

Observe that by normalising all functions $u \in \mathcal{D}(L)$ and thus relaxing the constraint $\|u\|^2 = 1$, the problem of finding the smallest eigenvalue λ_1 is equivalent to the following minimisation problem, called the **Rayleigh quotient**

$$\lambda_1 = \inf_{u \in \mathcal{D}(L)} \frac{\langle u, Lu \rangle}{\langle u, u \rangle}.$$

Similarly, if we restrict the class of admissible functions by requiring u also satisfying $u_k = \langle u, \phi_k \rangle = 0$ for all $k = 1, 2, \ldots, n-1$, *i.e.* u is orthogonal to lower eigenfunctions, then

$$E[u] = \sum_{j=n}^{\infty} \lambda_j |u_j|^2 \geq \lambda_n.$$

By defining $V_n = \text{span}(\phi_1, \phi_2, \ldots, \phi_{n-1})$, the problem of finding the nth eigenvalue λ_n is equivalent to the following minimisation problem

$$\lambda_n = \inf_{u \in \mathcal{D}(L), u \in V_n^{\perp}} \frac{\langle u, Lu \rangle}{\langle u, u \rangle}.$$

Of course, all of these are true with the assumptions that all eigenvalues are non-negative and their corresponding eigenfunctions form a complete orthonormal set. We now show that this is indeed the case for the Sturm-Liouville eigenvalue problem.

Theorem 4.3.1. *For the Sturm-Liouville operator* $L = -\dfrac{d}{dx}\left(p(x)\dfrac{d}{dx}\right) + q(x)$ *with* p, q *non-negative real-valued functions, we have the following results:*

(a) The eigenvalues $\lambda_n \longrightarrow \infty$ *as* $n \longrightarrow \infty$.

(b) The corresponding eigenfunctions (ϕ_n) *form a complete orthonormal set.*

Proof. We now use the variational principle that we just developed to prove these two statements. For simplicity, consider the **Dirichlet boundary condition**, *i.e.* $u(a) = u(b) = 0$.

(a) Assuming u is real. Expanding $E[u]$ and integrating by parts gives

$$\begin{aligned} E[u] = \langle u, Lu \rangle &= \int_a^b u^* Lu \, dx \\ &= \int_a^b u\Big[-(pu')' + qu\Big]\, dx \\ &= \int_a^b \Big[pu'^2 + qu^2\Big]\, dx - \Big[u^* pu'\Big]\Big|_a^b \\ &= \int_a^b \Big[pu'^2 + qu^2\Big]\, dx. \end{aligned}$$

where the boundary term vanishes. Consider the following expression

$$\frac{E[u]}{\|u\|^2} = \frac{\displaystyle\int_a^b \Big[pu'^2 + qu^2\Big]\, dx}{\displaystyle\int_a^b u^2 \, dx},$$

and denote the following quantity

$$p_M = \max_{x\in[a,b]} p(x), \quad q_M = \max_{x\in[a,b]} q(x)$$

$$p_m = \min_{x\in[a,b]} p(x), \quad q_m = \min_{x\in[a,b]} q(x)$$

Replacing $p(x), q(x)$ by the constant p_M, q_M, we obtain the following **modified Rayleigh Quotient**

$$\frac{E_M[u]}{\|u\|^2} \geq \frac{E[u]}{\|u\|^2} \implies \lambda_n^{(M)} \geq \lambda_n.$$

Similarly, replacing $p(x), q(x)$ by the constant p_m, q_m, we obtain the other **modified Rayleigh Quotient**

$$\frac{E_m[u]}{\|u\|^2} \leq \frac{E[u]}{\|u\|^2} \implies \lambda_n^{(m)} \leq \lambda_n.$$

Combining these two inequalities gives $\lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)}$.

Now, consider two intervals I_m, I_M such that $I_m \subset [a,b] \subset I_M$.

- Consider the variational problem for $\dfrac{E_M[u]}{\|u\|^2}$ on I_m. The admissible functions u must now vanish on ∂I_m, making the space of admissible functions smaller. This additional constraint implies that $\lambda_n^{(M)} \leq \hat{\lambda}_n^{(M)}$.
- Similarly, consider the variational problem for $\dfrac{E_m[u]}{\|u\|^2}$ on I_M. The admissible functions u do not have to vanish at $x = a, b$, making the space of admissible function larger. This implies that $\hat{\lambda}_n^{(m)} \leq \lambda_n^{(m)}$.

In summary, we have the hierarchy

$$\hat{\lambda}_n^{(m)} \leq \lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)} \leq \hat{\lambda}_n^{(M)}. \tag{4.3.5}$$

The eigenvalues $\hat{\lambda}_n^{(m)}$ and $\hat{\lambda}_n^{(M)}$ can be solved exactly since we have a constant coefficient ODE. We find that both $\hat{\lambda}_n^{(m)}$, $\hat{\lambda}_n^{(M)} \longrightarrow \infty$ as $n \longrightarrow \infty$. Hence, $\lambda_n \longrightarrow \infty$ as $n \longrightarrow \infty$.

(b) Consider the orthonormal set of eigenfunctions (ϕ_k), they satisfy the eigenvalue equation and boundary conditions

$$L\phi_k = -\frac{d}{dx}\left(p(x)\frac{d\phi_k}{dx}(x)\right) + q(x)\phi_k(x), \quad \phi_k(x) = 0 \text{ at } x = a, b.$$

Define the remainder term $R_n(x) = u(x) - \sum_{k=1}^{n} u_k\phi_k$, with $u_k = \langle u, \phi_k\rangle$. We need to show that $\lim_{n\to\infty} \|R_n(x)\| = 0$. Since $u(x)$ and $\phi_k(x)$ are admissable functions for the variational problem, so is the remainder $R_n(x)$. For all $j = 1, \ldots, n$,

$$\langle R_n, \phi_j\rangle = \left\langle u - \sum_{k=1}^{n} u_k\phi_k, \phi_j \right\rangle = \langle u, \phi_j\rangle - \sum_{k=1}^{n} u_k^*\langle \phi_k, \phi_j\rangle = 0. \tag{4.3.6}$$

By construction, we then have

$$\frac{E[R_n]}{\langle R_n, R_n\rangle} \geq \lambda_{n+1} \implies \|R_n\|^2 \leq \frac{E[R_n]}{\lambda_{n+1}}.$$

We have just shown that $\lambda_{n+1} \longrightarrow \infty$ as $n \longrightarrow \infty$, so $\|R_n\| \longrightarrow 0$ as $n \longrightarrow \infty$ provided $E[R_n]$ is uniformly bounded in n.

Let $\Omega_n = \displaystyle\sum_{k=1}^{n} u_k\phi_k$. Then

$$\begin{aligned}
E[u] &= E[R_n + \Omega_n] \\
&= \langle R_n + \Omega_n, L(R_n + \Omega_n)\rangle \\
&= \langle R_n, LR_n\rangle + \langle R_n, L\Omega_n\rangle + \langle \Omega_n, LR_n\rangle + \langle \Omega_n, L\Omega_n\rangle \\
&= E[R_n] + E[\Omega_n] + \langle LR_n, \Omega_n\rangle + \langle \Omega_n, LR_n\rangle \\
&= E[R_n] + E[\Omega_N] + 2\mathrm{Re}(\langle LR_n, \Omega_n\rangle) \\
&= E[R_n] + E[\Omega_n] + 2\langle LR_n, \Omega_n\rangle
\end{aligned}$$

where we use the self-adjointness of L on the fourth equality and the assumption that u, p, q are all real-valued functions. By repeating integration by parts,

$$\begin{aligned}
\langle LR_n, \Omega_n\rangle &= \left\langle LR_n, \sum_{k=1}^{n} u_k\phi_k \right\rangle = \sum_{k=1}^{n} u_k\langle LR_n, \phi_k\rangle \\
\langle LR_n, \phi_k\rangle &= \int_a^b \Big(-(pR_n')' + qR_n\Big)\phi_k\,dx \\
&= \Big[-pR_n'\phi_k\Big]\Big|_a^b + \int_a^b \Big(pR_n'\phi_k' + qR_n\phi_k\Big)\,dx \\
&= \Big[-pR_n'\phi_k + R_np\phi_k'\Big]\Big|_a^b + \int_a^b \Big(-R_n(p\phi_k')' + R_nq\phi_k\Big)\,dx
\end{aligned}$$

Because the Dirichlet boundary condition is homogeneous, $\mathcal{D}(L)$ is a vector space. Thus, $\phi_k(x), u(x) \in \mathcal{D}(L) \implies R_n(x) \in \mathcal{D}(L)$ and we immediately see that the boundary term vanishes. We are left with

$$\begin{aligned}
\int_a^b \Big(-R_n(p\phi_k')' + qR_n\phi_k\Big)\,dx &= \int_a^b R_n\Big(-(p\phi_k')' + q\phi_k\Big)\,dx \\
&= \lambda_k \int_a^b R_n\phi_k\,dx \\
&= \lambda_k\langle R_n, \phi_k\rangle = 0,
\end{aligned}$$

which follows from (4.3.6) since $1 \leq k \leq n$. Thus, $\langle LR_n, \Omega_n\rangle = 0$ and

$$\begin{aligned}
E[R_n] &= E[u] - E[\Omega_n] \\
&= E[u] - E\left[\sum_{k=1}^{n} u_k\phi_k\right]
\end{aligned}$$

$$
\begin{aligned}
&= E[u] - \int_a^b \left(\sum_{k=1}^n u_k^* \phi_k^* \right) L \left(\sum_{j=1}^n u_j \phi_j \right) dx \\
&= E[u] - \sum_{k,j=1}^n u_k^* u_j \langle \phi_k, L\phi_j \rangle \\
&= E[u] - \sum_{k,j=1}^n u_k^* u_j \lambda_j \langle \phi_k, \phi_j \rangle \\
&= E[u] - \sum_{k=1}^n \lambda_k |u_k|^2 \\
&\leq E[u],
\end{aligned}
$$

where we crucially use the fact that $\lambda_k \geq 0$ for all $k \geq 1$. Since u is an admissible function, $E[u]$ is a finite value. Hence, $E(R_n)$ is uniformly bounded in n, and we conclude that $\|R_n\| \longrightarrow 0$ as $n \longrightarrow \infty$.

■

Remark 4.3.2. Let us justfiy the eigenvalue inequality (4.3.5) we proved in the first part of the theorem.

1. It seems like a lower bound of the form $\lambda_n^{(m)} \leq \lambda_n$ is sufficient. Indeed, this inequality tells us that $\lambda_n^{(m)} \longrightarrow \infty$ as $n \longrightarrow \infty \implies \lambda_n \longrightarrow \infty$ as $n \longrightarrow \infty$. However, it doesn't guarantee that $\lambda_n < \infty$ for finite n.

2. The upper bound now comes into play. Now, we have the following inequality:

$$\lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)}$$

If additionally, we have that $\lambda_n^{(M)} \longrightarrow \infty$ as $n \longrightarrow \infty$, this "sandwich" guarantees that $\lambda_n < \infty$ for arbitrary but finite n.

3. Why is the inequality $\lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)}$ not enough then? We are implicitly assuming that we can show $\lambda_n^{(m)}, \lambda_n^{(M)} \longrightarrow \infty$ as $n \longrightarrow \infty$. In general (even in 2D), they can be really difficult to solve explicitly!

4.4 Distribution

In this section, we explore the analogy between matrices and linear operators on infinite-dimensional spaces. In the case of a finite-dimensional space, we choose a basis set and represent the linear operator in terms of the action of a matrix $A \in \mathbb{R}^{m \times n}$, *i.e.*

$$y = Ax \quad \text{or} \quad y_i = \sum_{j=1}^n A_{ij} x_j, 1 \leq i \leq m.$$

Formally speaking, the function space analogue of this, $g = Af$, is

$$g(x) = \int_a^b A(x, y) f(y)\, dy,$$

where we replace the summation over j by integration over a dummy variable y. If $A(x, y)$ is an ordinary function, then $A(x, y)$ is called an **integral kernel**. The **Dirac delta function** is the analogue of identity operator in infinite-dimensional vector spaces.

$$f(x) = \int_a^b \delta(x-y)f(y)\,dy.$$

Note that $\delta(x-y)$ is not an ordinary function, it is actually a generalised function/distribution as we will see later.

4.4.1 Distributions and Test Functions

We often think of $\delta(x)$ as being a *"limit"* of a sequence of functions with increasing height and decreasing width, such that its area under the curve is fixed.

Example 4.4.1. Consider the spike function

$$\delta_\varepsilon(x-a) = \begin{cases} \dfrac{1}{\varepsilon} & \text{if } x \in \left[a - \dfrac{\varepsilon}{2}, a + \dfrac{\varepsilon}{2}\right], \\ 0 & \text{otherwise.} \end{cases}$$

The L^2-norm of δ_ε is

$$\|\delta_\varepsilon\|^2 = \int_{-\infty}^{\infty} |\delta_\varepsilon(x)|^2\,dx = \frac{1}{\varepsilon^2}\int_{a-\varepsilon/2}^{a+\varepsilon/2} dx = \frac{1}{\varepsilon} \longrightarrow \infty \text{ as } \varepsilon \longrightarrow 0.$$

Thus, as $\varepsilon \longrightarrow 0$, the spike function δ_ε doesn't converge to any function in $L^2(\mathbb{R})$.

Example 4.4.2. Another example arises from Fourier theory:

$$\delta_\lambda(x) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda} e^{ikx}\,dk = \frac{1}{2\pi}\frac{e^{ikx}}{ix}\bigg|_{-\lambda}^{\lambda} = \frac{1}{\pi}\left(\frac{\sin(\lambda x)}{x}\right).$$

The L^2 norm of δ_λ is

$$\begin{aligned} \|\delta_\lambda\|^2 = \int_{-\infty}^{\infty} \frac{\sin^2(\lambda x)}{\pi^2 x^2}\,dx = \frac{\lambda^2}{\pi^2}\int_{-\infty}^{\infty}\frac{\sin^2(\lambda x)}{\lambda^2 x^2}\,dx &= \frac{\lambda}{\pi^2}\int_{-\infty}^{\infty}\frac{\sin^2(x)}{x^2}\,dx \\ &= \frac{\lambda}{\pi^2}\pi = \frac{\lambda}{\pi} \longrightarrow \infty \text{ as } \lambda \longrightarrow \infty. \end{aligned}$$

It wasn't until late 1940s, that Laurent Schwartz developed the **theory of distribution** and succeeded in explaining all these singular objects. To make sense of the delta function, Schwartz exploited the concept of a dual space from linear algebra. Recall that, the **dual space** V^* of a vector space V is the vector space of all linear functions from V to $\mathbb{R}$ (or $\mathbb{C}$). We can interpret $\delta(x)$ as an element of the dual space of a vector space $\mathcal{D}$, called the **space of test functions**. For our purposes, test functions are smooth (infinitely differentiable) functions that converges rapidly to zero at infinity. Precise definition of test functions is problem-dependent. A common example is C_c^∞, the space of smooth functions with compact support. In Fourier theory and harmonic analysis, one usually works with the **Schwartz space**, $\mathcal{S}(\mathbb{R})$.

Remark 4.4.3. The *"nice"* behaviour of test functions compensate for the *"nasty"* behaviour of $\delta(x)$ and its relatives. However, not every linear map $\mathcal{D} \longrightarrow \mathbb{R}$ is admissible, we actually require these maps to be continuous. More precisely, if $\varphi_n \longrightarrow \varphi$ in $\mathcal{D}$, then $u \in \mathcal{D}^*$ must obey $u(\varphi_n) \longrightarrow u(\varphi)$. A topology must be specified on $\mathcal{D}$ in order to talk about convergence/continuity.

Definition 4.4.4. In terms of the dual space formulation, we define the **Dirac delta function** as the linear map satisfying

$$(\delta, \varphi) = \varphi(0) \quad \text{for all } \varphi \in \mathcal{D}.$$

The notation is not to be confused with an inner product. It is actually a **duality pairing**.

Example 4.4.5. Assuming integration by parts is valid, we have that

$$\int_{-\infty}^{\infty} \delta'(x)\varphi(x)\,dx = -\int_{-\infty}^{\infty} \delta(x)\frac{d\varphi}{dx}\,dx = -\varphi'(0).$$

This suggests that we can define the distribution $\delta^{(n)}$ as the linear map satisfying

$$(\delta^{(n)}, \varphi) = (-1)^n \varphi^{(n)}(0).$$

L^2 functions are not "nice" enough for their dual space to accommodate the Dirac delta function. The first observation is that L^2 is a reflexive Hilbert space. The **Riesz-Frechet representation theorem** asserts that any continuous, linear map $F\colon L^2 \longrightarrow \mathbb{R}$ can be written as $F(\varphi) = \langle f, \varphi\rangle$ for some unique $f \in L^2$. However, δ is not a continuous map from L^2 to $\mathbb{R}$. Another reason is that L^2 functions are only defined up to sets of measure zero, consequently $\varphi(0)$ for $\varphi \in L^2$ is undefined. The upshot is, to work with much more exotic distributions, the space of test functions has to be as nice as they can be.

4.4.2 Weak Derivatives

Definition 4.4.6. We define the **weak** or **distributional derivative**, $v(x)$ of a distribution u by requiring the following to be true for all test functions $\varphi \in \mathcal{D}$:

$$\int v(x)\varphi(x)\,dx = -\int u(x)\varphi'(x)\,dx.$$

Remark 4.4.7. The weak derivative is a well-defined distribution. This definition is obtained by formally performing integration by parts. The same idea also applies to ordinary functions that are not differentiable in the classical sense.

Example 4.4.8. In the weak sense, $\dfrac{d}{dx}|x| = \operatorname{sgn}(x)$.

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{d}{dx}|x|\varphi(x)\,dx &= -\int_{-\infty}^{\infty} |x|\varphi'(x)\,dx \\
&= -\int_{0}^{\infty} x\varphi'(x)\,dx + \int_{-\infty}^{0} x\varphi'(x)\,dx \\
&= -\Big[x\varphi(x)\Big]\Big|_0^{\infty} + \int_0^{\infty} \varphi(x)\,dx + \Big[x\varphi(x)\Big]\Big|_{-\infty}^{0} - \int_{-\infty}^{0} \varphi(x)\,dx \\
&= \int_0^{\infty} \varphi(x)\,dx - \int_{-\infty}^{0} \varphi(x)\,dx \\
&= \int_{-\infty}^{\infty} \operatorname{sgn}(x)\varphi(x)\,dx.
\end{aligned}$$

Example 4.4.9. In the weak sense, $\dfrac{d}{dx}\operatorname{sgn}(x) = 2\delta(x)$.

$$\begin{aligned}
\int_{-\infty}^{\infty} \left(\frac{d}{dx}\operatorname{sgn}(x)\right)\varphi(x)\,dx &= -\int_{-\infty}^{\infty} \operatorname{sgn}(x)\varphi'(x)\,dx \\
&= -\int_{0}^{\infty} \varphi'(x)\,dx + \int_{-\infty}^{0} \varphi'(x)\,dx \\
&= 2\varphi(0) \\
&= \int_{-\infty}^{\infty} 2\delta(x)\varphi(x)\,dx.
\end{aligned}$$

Example 4.4.10. Solve the differential equation $u' = 0$ in the sense of distribution. That is, we need to find u such that

$$0 = (u', \varphi) = -(u, \varphi') \qquad \text{for all } \varphi \in C_c^{\infty}(\mathbb{R}).$$

First, we need to determine the action of u on all test functions $\varphi \in C_c^{\infty}(\mathbb{R})$ such that $(u, \Psi) = 0$ whenever $\Psi(x) = \varphi'(x)$.

- W claim that $\Psi(x) = \varphi'(x)$ for some $\varphi \in C_c^{\infty}(\mathbb{R})$ if and only if $\displaystyle\int_{-\infty}^{\infty} \Psi(x)\,dx = 0$.

  Proof. If $\Psi(x) = \varphi'(x)$ for some $\varphi \in C_c^{\infty}(\mathbb{R})$, then

  $$\int_{-\infty}^{\infty} \Psi(x)\,dx = \varphi(x)\Big|_{-\infty}^{\infty} = 0.$$

  Conversely, suppose $\displaystyle\int_{-\infty}^{\infty} \Psi(x)\,dx = 0$. Define

  $$\varphi(x) = \int_{-\infty}^{x} \Psi(s)\,ds.$$

Then $\varphi'(x) = \Psi(x)$ and $\varphi \in C_c^\infty(\mathbb{R})$ since $\Psi \in C_c^\infty(\mathbb{R})$ and $\displaystyle\int_{-\infty}^{\infty} \Psi(x)dx = 0$.

■

- Choose an arbitrary $\varphi_0 \in C_c^\infty(\mathbb{R})$ such that $\displaystyle\int_{-\infty}^{\infty} \varphi_0(x)\,dx = 1$. We can rewrite any $\varphi \in C_c^\infty(\mathbb{R})$ in the following form:

$$\begin{aligned}\varphi(x) &= \varphi_0(x)\int_{-\infty}^{\infty} \varphi(s)\,ds + \left[\varphi(x) - \varphi_0(x)\int_{-\infty}^{\infty} \varphi(s)\,ds\right] \\ &= \varphi_0(x)\int_{-\infty}^{\infty} \varphi(s)\,ds + \Psi(x).\end{aligned}$$

Since

$$\int_{-\infty}^{\infty} \Psi(x)\,dx = \int_{-\infty}^{\infty} \varphi(x)\,dx - \left(\int_{-\infty}^{\infty} \varphi(s)\,ds\right)\int_{-\infty}^{\infty} \varphi_0(x)\,dx = 0,$$

the claim above tells us that there exists $\hat{\varphi} \in C_c^\infty(\mathbb{R})$ such that $\Psi(x) = \hat{\varphi}'(x)$. Consequently, $(u, \Psi) = 0$ and

$$\begin{aligned}(u, \varphi) &= \left(u, \varphi_0(x)\int_{-\infty}^{\infty} \varphi(s)\,ds + \Psi(x)\right) \\ &= \left[\int_{-\infty}^{\infty} \varphi(s)\,ds\right](u, \varphi_0) + (u, \Psi) \\ &= (u, \varphi_0)\left[\int_{-\infty}^{\infty} \varphi(s)\,ds\right].\end{aligned}$$

Since $(u, \varphi_0) \in \mathbb{R}$, we have that for any $\varphi \in C_c^\infty(\mathbb{R})$, $(u, \varphi) = (c, \varphi)$. The distributional solution is therefore $u = c$ for some constant $c \in \mathbb{R}$.

4.5 Green's Function

Green's function is useful for solving the inhomogeneous linear equation $Ly = f$, where L is a linear differential operator. Roughly speaking, it is an **integral kernel** representing the inverse operator L^{-1}.

4.5.1 Fredholm Alternative

Theorem 4.5.1 (Fredholm Alternative)**.** *Let V be a finite-dimensional vector space equipped with an inner product and $A\colon V \longrightarrow V$ a linear operator. Exactly one of the following must hold:*

(a) $Ax = b$ has a unique solution. (A^{-1} exists)

(b) $Ax = 0$ has a non-trivial solution. In this case, then $Ax = b$ has no solution unless b is orthogonal to all solutions of $A^\dagger x = \mathbf{0}$.

Remark 4.5.2. This result continues to hold for linear differential operators on function spaces, defined on a finite interval, provided that $L^\dagger$ is defined using adjoint boundary conditions. In general, $\mathcal{N}(L) \neq \mathcal{N}(L^\dagger)$ unless the number of boundary conditions is equal to the order of the equation.

Example 4.5.3. Consider $Ly = \dfrac{dy}{dx}$, with boundary conditions $y(0) = y(1) = 0$, $x \in [0,1]$. $Ly = 0$ only has the trivial solution $y \equiv 0$ due to the boundary conditions. Hence, if solution to $Ly = f$ exists, it will be unique. For any $y_1 \in \mathcal{D}(L)$,

$$\langle y_2, Ly_1 \rangle = \int_0^1 y_2 \left(\frac{dy_1}{dx}\right) dx = -\int_0^1 \left(\frac{dy_2}{dx}\right) y_1 \, dx = \langle L^\dagger y_2, y_1 \rangle.$$

Thus, $L^\dagger = -\dfrac{d}{dx}$ with no boundary conditions needed to define $\mathcal{D}(L^\dagger)$. Clearly, $L^\dagger y = 0$ has a non-trivial solution $y = 1$. It follows from Fredholm alternative that $Ly = f$ has no solution unless f is orthogonal to 1, *i.e.*

$$\langle f, 1 \rangle = \int_0^1 f(x)\, dx = 0.$$

If this condition is satisfied, then $y(x) = \displaystyle\int_0^x f(s)\, ds$. Clearly, $y(0) = y(1) = 0$.

4.5.2 Green's Functions for Homogeneous Boundary Conditions

Consider a linear differential operator L, with $\mathcal{N}(L) = \mathcal{N}(L^\dagger) = \{\mathbf{0}\}$. To solve the equation $Ly = f$ for a given f, the most natural/naive method is to find the inverse operator L^{-1}. We represent L^{-1} as an integral kernel $(L^{-1})_{x,\xi} = G(x,\xi)$, satisfying

$$L_x G(x,\xi) = \delta(x-\xi). \tag{4.5.1}$$

Here, the subscript on L_x means that L acts on the first argument, x, of G. The solution to $Ly = f$ can now be written as

$$y(x) = \int G(x,\xi) f(\xi)\, d\xi, \tag{4.5.2}$$

since

$$Ly(x) = \int L_x G(x,\xi) f(\xi)\, d\xi = \int \delta(x-\xi) f(\xi)\, d\xi = f(x).$$

The problem is now reduces to constructing $G(x,\xi)$. There are 3 necessary conditions on $G(x,\xi)$:

1. For a fixed ξ, the function $\chi(x) = G(x,\xi)$ must have some discontinuous behaviour at $x = \xi$ in order to generate the Dirac delta function $\delta(x-\xi)$. This is motivated by the following fact
$$\frac{d}{dx} H(x-\xi) = \delta(x-\xi) \quad \text{in the weak sense,}$$
where $H(x-\xi)$ is the **Heaviside function**.

2. $\chi(x)$ must satisfy $L\chi = 0$ for all $x \neq \xi$ so that (4.5.1) is satisfied.

3. $\chi(x)$ must obey the same homogeneous boundary conditions as $y(x)$. This ensures that the solution $y(x)$ given by (4.5.2) satisfies the required boundary conditions. It also ensures that $\mathcal{R}(G) \subset \mathcal{D}(L)$, otherwise the composition $LG = I$ does not make sense.

Example 4.5.4. Consider the Sturm-Liouville differential operator

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y(x) = f(x), \quad x \in [a,b] \subset \mathbb{R},$$

with homogeneous self-adjoint boundary conditions

$$\frac{y'(a)}{y(a)} = \frac{\beta}{\alpha}, \quad \frac{y'(b)}{y(b)} = \frac{\hat{\beta}}{\hat{\alpha}}.$$

We want to construct a function $G(x,\xi)$ such that $LG = (pG')' + qG = \delta(x-\xi)$.

First, we require G to be continuous at $x = \xi$; otherwise the second derivative applied to a jump function will generate δ' term. Let $G(x,\xi)$ be defined as follows

$$G(x,\xi) = \begin{cases} Ay_L(x)y_R(\xi) & \text{if } x < \xi, \\ Ay_L(\xi)y_R(x) & \text{if } x > \xi. \end{cases}$$

This is continuous at $x = \xi$ by construction. We choose y_L, y_R such that

- $y_L(x)$ satisfies $Ly_L = \mathbf{0}$ and the left boundary condition.
- $y_R(x)$ satisfies $Ly_R = \mathbf{0}$ and the right boundary condition.

With these choices, we see that $G(x,\xi)$ satisfies the homogeneous boundary conditions and $LG = \delta(x-\xi)$.

In order to determine conditions on G at $x = \xi$, we integrate $LG = \delta$ from $x = \xi - \varepsilon$ to $x = \xi + \varepsilon$:

$$\int_{\xi-\varepsilon}^{\xi+\varepsilon} (pG')' + qG\,dx = \int_{\xi-\varepsilon}^{\xi+\varepsilon} \delta(x-\xi)\,dx = 1.$$

As $\varepsilon \longrightarrow 0$, $\displaystyle\int_{\xi-\varepsilon}^{\xi+\varepsilon} qG\,dx \longrightarrow 0$ since $G(x,\xi)$ and $q(x)$ are continuous functions. Thus, we are left with

$$\begin{aligned} \lim_{\varepsilon\to 0}\Big[p(\xi+\varepsilon)G'(\xi+\varepsilon,\xi) - p(\xi-\varepsilon)G'(\xi-\varepsilon,\xi)\Big] &= 1 \\ \lim_{\varepsilon\to 0} p(\xi)\Big[\Big(G'(\xi+\varepsilon,\xi) - G'(\xi-\varepsilon,\xi)\Big)\Big] &= 1 \\ \lim_{\varepsilon\to 0} Ap(\xi)\Big[y_L(\xi)y_R'(\xi+\varepsilon) - y_L'(\xi-\varepsilon)y_R(\xi)\Big] &= 1 \end{aligned}$$

$$Ap(\xi)\left[y_L(\xi)y_R'(\xi) - y_L'(\xi)y_R(\xi)\right] = 1$$
$$Ap(\xi)W(\xi) = 1$$

where $W(\xi)$ is the Wronskian of y_L, y_R. Thus, provided $p(\xi)W(\xi) \neq 0$,

$$G(x,\xi) = \begin{cases} \dfrac{1}{p(\xi)W(\xi)} y_L(x)y_R(\xi) & \text{if } x < \xi, \\ \dfrac{1}{p(\xi)W(\xi)} y_L(\xi)y_R(x) & \text{if } x > \xi. \end{cases}$$

For the Sturm-Liouville equation, $A = p(\xi)W(\xi)$ is constant. To see this, we first rewrite the homogeneous Sturm-Liouville equation $py'' + p'y' + qy = \mathbf{0}$ into systems of first order ODE:

$$\begin{cases} y' &= z \\ z' &= y'' = -\dfrac{p'z}{p} - \dfrac{qy}{p} \end{cases} \implies \begin{bmatrix} y' \\ z' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q/p & -p'/p \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix}.$$

Invoking the **Liouville's formula** for non-autonomous, homogeneous linear ODE:

$$W(\xi) = W(0)\exp\left[-\int_0^\xi \frac{p'(x)}{p(x)}\,dx\right] = W(0)\exp\left[-\ln\left(\frac{p(\xi)}{p(0)}\right)\right] \quad = \frac{W(0)p(0)}{p(\xi)} \text{ for all } \xi.$$

Hence, the solution to $Ly = f$ is given by

$$y(x) = \frac{1}{pW}\left\{y_L(x)\int_x^b y_R(\xi)f(\xi)\,d\xi + y_R(x)\int_a^x y_L(\xi)f(\xi)\,d\xi\right\}.$$

Remark 4.5.5. The constancy of pW means that G is symmetric, *i.e.* $G(x,\xi) = G(\xi,x)$. We require that $W \neq 0$, which is an indication of linear independency. If $W = 0$, this means that $y_L \sim y_R$ and the single function y_R would satisfy $Ly_R = 0$ and the boundary conditions. That is, $\mathcal{N}(L)$ is non-trivial and solutions (if they exists) to $Ly = f$ are not unique.

Example 4.5.6. Consider the boundary value problem

$$-\frac{d^2y}{dx^2} = f(x), \quad y(0) = y(1) = 0.$$

Solutions to $Ly = 0$ are of the form $y(x) = Ax + B$. Solving for the boundary conditions yields $y_L(x) = x$ and $y_R(x) = 1 - x$. Also,

$$W = y_L y_R' - y_L' y_R = x(-1) - 1(1-x) = -1 \neq 0.$$

Therefore,

$$G(x,\xi) = \begin{cases} x(1-\xi) & \text{if } 0 < x < \xi, \\ \xi(1-x) & \text{if } \xi < x < 1. \end{cases}$$

and

$$\begin{aligned} y(x) = \int_0^1 G(x,\xi)f(\xi)\,d\xi &= \int_x^1 x(1-\xi)f(\xi)\,d\xi + \int_0^x \xi(1-x)f(\xi)\,d\xi \\ &= (1-x)\int_0^x \xi f(\xi)\,d\xi + x\int_x^1 (1-\xi)f(\xi)\,d\xi. \end{aligned}$$

4.5.3 Modified Green's Function

When the equation $Ly = 0$ has a non-trivial solution, there can be no unique solution to $Ly = f$, but solution will still exists if $f \in \mathcal{N}(L^\dagger)^\perp$. One can express the family of solutions using the modified Green's function.

Example 4.5.7. Consider the following boundary value problem

$$Ly = -\frac{d^2y}{dx^2} = f(x), \quad y'(0) = y'(1) = 0.$$

$Ly = 0$ has a non-trivial solution $y = 1$. The operator is self-adjoint in $L^2[0,1]$ and Fredholm alternative says that there will be no solution to $Ly = f$ unless

$$\int_0^1 f(x)\,dx = 0.$$

Observe that we cannot define Green's function as a solution to

$$-\frac{\partial^2}{\partial x^2}G(x,\xi) = L_xG(x,\xi) = \delta(x-\xi),$$

since $\displaystyle\int_0^1 L_xG(x,\xi)\,dx = 0$, but $\displaystyle\int_0^1 \delta(x-\xi)\,dx = 1$. This can be rectified by modifying the definition of Green's function. The above suggests defining $G(x,\xi)$ according to the modified equation

$$-\frac{\partial^2}{\partial x^2}G(x,\xi) = L_xG(x,\xi) = \delta(x-\xi) - 1.$$

- Integrating once with respect to x the equation $-\dfrac{\partial^2 G}{\partial x^2} = -1$ yields

$$G_x(x,\xi) = \begin{cases} x + A_L = x & \text{if } 0 < x < \xi, \\ x + A_R = x - 1 & \text{if } \xi < x < 1, \end{cases}$$

  where A_L and A_R are found by using the boundary conditions $G_x(0,\xi) = G_x(1,\xi) = 0$. Integrating $G_x(x,\xi)$ with respect to x yields

$$G(x,\xi) = \begin{cases} \dfrac{1}{2}x^2 + B_L & \text{if } 0 < x < \xi, \\ \dfrac{1}{2}x^2 - x + B_R & \text{if } \xi < x < 1. \end{cases}$$

  where B_L, B_R are functions of ξ.

- Continuity at $x = \xi$ gives $B_L = B_R - \xi$. Thus,

$$G(x,\xi) = \begin{cases} \dfrac{1}{2}x^2 + B_R - \xi & \text{if } 0 < x < \xi, \\ \dfrac{1}{2}x^2 - x + B_R & \text{if } \xi < x < 1. \end{cases}$$

Furthermore, for this particular structure of $G(x, \xi)$, we find that the jump discontinuity condition

$$\lim_{\varepsilon \to 0} - \left[G_x(\xi + \varepsilon, \xi) - G_x(\xi - \varepsilon, \xi) \right] = 1,$$

is automatically satisfied, which means that there is no restriction on $B_R(\xi)$.

- We choose $B_R(\xi) = \dfrac{1}{2}\xi^2 + \dfrac{1}{3}$ so that $G(x, \xi)$ is symmetric. Hence,

$$G(x, \xi) = \begin{cases} \dfrac{1}{3} - \xi + \dfrac{1}{2}(x^2 + \xi^2) & \text{if } 0 < x < \xi, \\ \dfrac{1}{3} - x + \dfrac{1}{2}(x^2 + \xi^2) & \text{if } \xi < x < 1, \end{cases}$$

and a solution is given by

$$y(x) = \int_0^1 G(x, \xi) f(\xi)\, d\xi + A,$$

where A is an arbitrary constant. Note that an additive constant C in G does not contribute to a solution due to the requirement $\displaystyle\int_0^1 f(x) dx = 0$, *i.e.*

$$\int_0^1 \Big(G(x, \xi) + C \Big) f(\xi)\, d\xi = \int_0^1 \Big(G(x, \xi) f(\xi) + C f(\xi) \Big)\, d\xi = \int_0^1 G(x, \xi) f(\xi)\, d\xi.$$

4.5.4 Eigenfunction Expansions

This subsection is to address why we can always modified the Green's function definition and how to modify it. Recall that self-adjoint operators possess a complete set of eigenfunctions, which we can use to expand the Green's function. Let $L\phi_n = \lambda_n \phi_n$ and assume $\lambda_n \neq 0$ for all $n \geq 1$; this means that L is invertible since absence of zero eigenvalue implies $\mathcal{N}(L) = \{\mathbf{0}\}$. The Green's function then has eigenfunctions expansion given by

$$G(x, \xi) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n^*(\xi)}{\lambda_n},$$

since

$$\begin{aligned} L_x G(x, \xi) = L_x \left(\sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n^*(\xi)}{\lambda_n} \right) &= \sum_{n=1}^{\infty} L_x(\phi_n(x)) \frac{\phi_n^*(\xi)}{\lambda_n} \\ &= \sum_{n=1}^{\infty} \phi_n(x)\phi_n^*(\xi) \\ &= \delta(x - \xi), \end{aligned}$$

where the last equality follows from completeness of eigenfunctions.

Example 4.5.8. Suppose that $L = -\dfrac{d}{dx^2}$, with domain

$$\mathcal{D}(L) = \Big\{y, Ly \in L^2[0,1] \colon y(0) = y(1) = 0\Big\}.$$

The Green's function is given by

$$G(x,\xi) = \begin{cases} x(1-\xi) & \text{if } 0 < x < \xi, \\ \xi(1-x) & \text{if } \xi < x < 1. \end{cases}$$

Alternatively, the eigenvalue equation $L\phi_n = \lambda_n\phi_n$ has normalised eigenfunctions $\phi_n(x) = \sqrt{2}\sin(n\pi x)$ with corresponding eigenvalues $\lambda_n = n^2\pi^2$. Thus,

$$G(x,\xi) = \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2}\sin(n\pi x)\sin(n\pi\xi).$$

When one or more eigenvalues are zero, *i.e.* $\mathcal{N}(L)$ is non-trivial, a modified Green's function is obtained by simply omitting the corresponding terms in the series. More precisely, we define the modified Green's function as

$$G_{mod}(x,\xi) = \sum_{\lambda_n \neq 0} \frac{\phi_n(x)\phi_n^*(\xi)}{\lambda_n},$$

which gives

$$LG_{mod}(x,\xi) = \sum_{\lambda_n \neq 0} \phi_n(x)\phi_n^*(\xi) = \delta(x-\xi) - \sum_{\lambda_n = 0} \phi_n(x)\phi_n^*(x).$$

4.5.5 Inhomogeneous Boundary Conditions

Suppose we wish to solve the following boundary value problem

$$-\frac{d^2y}{dx^2} = f(x), \quad y(0) = a, y(1) = b.$$

We can still use the Green's function for homogeneous boundary conditions. We need to go back to the first principle, where we will pick up boundary terms. We saw previously that

$$G(x,\xi) = \begin{cases} x(1-\xi) & \text{if } 0 < x < \xi, \\ \xi(1-x) & \text{if } \xi < x < 1. \end{cases}$$

Using Green's theorem,

$$\begin{aligned} \langle y, LG\rangle - \langle L^\dagger y, G\rangle &= \int_0^1 \Big[y(x)\Big(-G_{xx}(x,\xi)\Big) - \Big(-y''(x)\Big)G(x,\xi)\Big]\,dx \\ &= \Big[-G_x y\Big]\Big|_{x=0}^{x=1} + \int_0^1 \Big[y'(x)G_x(x,\xi)\Big]\,dx \\ &\quad - \Big[-y'G\Big]\Big|_{x=0}^{x=1} - \int_0^1 \Big[y'(x)G_x(x,\xi)\Big]\,dx \end{aligned}$$

$$
\begin{aligned}
&= \Big[G(x,\xi)y'(x) - G_x(x,\xi)y(x)\Big]\Big|_{x=0}^{x=1} \\
&= G(1,\xi)y'(1) - G_x(1,\xi)y(1) - G(0,\xi)y'(0) + G_x(0,\xi)y(0) \\
&= (0)y'(1) - (-\xi)y(1) - (0)y'(0) + (1-\xi)y(0) \\
&= (1-\xi)y(0) + \xi y(1).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\int_0^1 \Big[y(x)\Big(-G_{xx}(x,\xi)\Big) - \Big(-y''(x)\Big)G(x,\xi)\Big]\,dx \\
&= \int_0^1 \Big[y(x)\delta(x-\xi) - f(x)G(x,\xi)\Big]\,dx \\
&= y(\xi) - \int_0^1 G(x,\xi)f(x)\,dx.
\end{aligned}
$$

Hence,

$$
y(\xi) = \int_0^1 G(x,\xi)f(x)\,dx + (1-\xi)y(0) + \xi y(1).
$$

4.6 Problems

1. (a) Show that if a norm is derived from an inner product, then it obeys the parallelogram law
$$\|f+g\|^2+\|f-g\|^2 = 2(\|f\|^2+\|g\|^2).$$

Solution:

(b) Show that a compact linear operator is always bounded.

Solution:

(c) Show that the eigenvalues of a compact symmetric operator satisfy $|\lambda| \leq \|A\|$.

Solution:

2. Find the adjoint operator $L^\dagger$ and its domain (assuming standard inner product on $L^2[0,1]$).

(a) $Lu = u'' + a(x)u' + b(x)u, u(0) = u'(1), u(1) = u'(0)$, where a is continuously differentiable and b is continuous.

Solution:

(b) $Lu = -(p(x)u')' + q(x)u, u(0) = u(1), u'(0) = u'(1)$, where p is continuously differentiable and q is continuous.

Solution:

3. Consider the differential operator acting on $L^2(\mathbb{R})$
$$L = -\frac{d^2}{dx^2}, \quad 0 \leq x < \infty,$$
with self-adjoint boundary conditions $\phi(0)/\phi'(0) = \tan\theta$ for some fixed angle θ.

(a) Show that when $\tan\theta < 0$, there is a single negative eigenvalue with a normalisable eigenfunction $\phi_0(x)$ localised near the origin, but none when $\tan\theta > 0$.

Solution:

(b) Show that there is a continuum of eigenvalues $\lambda = k^2$ with eigenfunctions $\phi_k(x) = \sin(kx + \eta(k))$, where the phase shift η is found from
$$e^{i\eta(k)} = \frac{1+ik\tan\theta}{\sqrt{1+k^2\tan^2\theta}}.$$

Solution:

(c) Evaluate the integral

$$I(x,x') = \frac{2}{\pi}\int_0^\infty \sin(kx+\eta(k))\sin(kx'+\eta(k))\,dk,$$

and interpret the result with regards the relationship to the Dirac delta function and completeness, that is,

$$\delta(x-x') - I(x,x') = \phi_0(x)\phi_0(x').$$

You will need the following standard integral

$$\int_{-\infty}^{\infty} \frac{e^{ikx}}{1+k^2t^2}\frac{dk}{2\pi} = \frac{1}{2|t|}e^{-|x/t|}.$$

Hint: You should monitor how the bound state contribution (for $\tan\theta < 0$) switches on and off as θ is varied. Keeping track of the modulus signs in the given standard integral is crucial for this.

Solution:

4. Consider the so-called Poschel-Teller (PT) equation

$$-\frac{d^2\phi}{dx^2} - 2\text{sech}^2(x)\phi = E\phi, \quad -\infty < x < \infty.$$

(a) Show that under the change of variable $u = \tanh(x)$, the equation becomes

$$L\phi := -\frac{d^2\phi}{dx^2} + [u^2 - u']\phi = \lambda\phi, \quad \lambda = E+1.$$

Solution:

(b) Show that the differential operator L can be rewritten as $L = M^\dagger M$, where

$$M = \left(\frac{d}{dx} + u(x)\right), \quad M^\dagger = \left(-\frac{d}{dx} + u(x)\right).$$

Solution:

(c) Suppose that $\phi_-(x)$ is an eigenfunction of the equation $M^\dagger M\phi = \lambda\phi$. By left-multiplying both sides by M, show that there exists an eigenfunction ϕ_+ satisfying $MM^\dagger\phi_+ = \lambda\phi_+$ for the same eigenvalue λ. Derive the inverse relationship

$$(E+1)\phi_- = M^\dagger\phi_+,$$

and hence show that the mapping $\phi_+ \longleftrightarrow \phi_-$ breaks down in the special case $\lambda = 0$, that is, $E = -1$.

Solution:

(d) By writing down the explicit form for the equation $MM^\dagger\phi_+ = \lambda\phi_+$, show that it has eigensolutions of the form $\phi_{+,k}(x) = e^{ikx}$ with $E = k^2$. Using the relationship between ϕ_- and ϕ_+ (for $\lambda \neq 0$), deduce that the differential operator L has a continuous spectrum $k^2 + 1$ with eigenfunctions

$$\phi_k(x) = \frac{1}{\sqrt{1+k^2}} e^{ikx}(-ik + \tanh(x)).$$

The normalisation is chosen such that at large $|x|$, where $\tanh(x) \longrightarrow \pm 1$, we have

$$\phi_k^*(x)\phi_k(x') \longrightarrow e^{-ik(x-x')}.$$

Solution:

(e) Also show that L has a discrete eigenvalue $\lambda = 0$ with normalised eigenfunction

$$\phi_0(x) = \frac{1}{\sqrt{2}}\operatorname{sech}(x).$$

Solution:

(f) Compute the difference

$$I = \delta(x-x') - \int_{-\infty}^{\infty} \phi_k^*(x)\phi_k(x')\,\frac{dk}{2\pi},$$

by using

$$\delta(x-x') = \int_{-\infty}^{\infty} e^{-ik(x-x')}\,\frac{dk}{2\pi},$$

and the standard integral

$$\int_{-\infty}^{\infty} \frac{e^{-ik(x-x')}}{1+k^2}\,\frac{dk}{2\pi} = \frac{1}{2}e^{-|x-x'|},$$

together with its x' derivative. Hence derive the expected completeness condition

$$\delta(x-x') = \phi_0(x)\phi_0(x') + \int_{-\infty}^{\infty} \phi_k^*(x)\phi_k(x')\,\frac{dk}{2\pi}.$$

Solution:

5. (a) Solve the following differential equations in the sense of distribution:

   i. $x^2u' = 0$,

Solution:

ii. $u'' = \delta''(x)$.

Solution:

(b) Express $\delta(x^2 - a^2)$ in terms of the usual δ-function.

Solution:

6. Construct the Green's function for the following problems:

(a) $u'' + \alpha^2 u = f(x), u(0) = u(1), u'(0) = u'(1)$. For what value of α does the Green's function fail to exist?

Solution:

(b) $u'' + \dfrac{3}{2x}u' - \dfrac{3}{2x^2}u = f(x), u(0) = 0, u'(1) = 0$.

Solution:

7. Consider the ODE

$$-y''(x) = f(x), \quad 0 < x < 1, \quad y(0) = y(1), y'(0) = y'(1).$$

(a) Show that the associated linear operator has a zero eigenvalue and determine the eigenfunction. What condition must f satisfy for a solution to exist?

Solution:

(b) Show that the modified Green's function for this problem is

$$\hat{G}(x, \xi) = \frac{1}{2}(x - \xi)^2 - \frac{1}{2}|x - \xi|.$$

You should make sure that this Green's function satisfies the ODE and the boundary conditions.

Solution: