SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Topics in Nonlinear Ordinary Differential Equations

av

Teodor Alfson

2015 - No 22

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Topics in Nonlinear Ordinary Differential Equations

Teodor Alfson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Andrzej Szulkin

2015

Topics in Nonlinear Ordinary Differential Equations

Teodor Alfson

Abstract

In this paper we describe different methods used in showing existence of solutions to boundary value problems for nonlinear ordinary differential equations. We describe the shooting method and give an example how it can be applied to the pendulum equation. We look at examples of how the shooting method can be used together with a priori bounds and Poincaré maps. We look at contraction maps and we prove Banach's contraction mapping principle and give an example of its use. Finally, we give a brief overview of the Mountain Pass Theorem and return to the pendulum equation.

Contents

1 Introduction

When trying to describe many physical phenomena, one of the more applicable theories in mathematics is that of differential equations. Such an equation defines a relation between a function and its derivatives. Applications are prominent throughout areas such as physics and engineering, among others.

Linear ordinary differential equations are most of the time very handy to work with, and solutions can often be found in exact form. However, for nonlinear equations it is a bit harder, since representing them with elementary functions is not always possible. With sufficient data (for example initial or boundary values) we can gather enough information about the behavior of the solution.

In this paper, by examples, we illustrate different methods for showing existence of solutions to certain boundary value problems for nonlinear differential equations, in particular the pendulum equation. In section 2 we start off with a few definitions and theorems that will be used throughout the text.

Section 3 is dedicated to the pendulum, and we derive the differential equation associated with it.

In section 4 we describe the shooting method and include an application. We continue with examples of how a priori bounds can be used together with the shooting method to show existence of solutions. Following that, periodic problems and Poincaré maps are described and we look at a couple of examples. All examples in section 4 are taken from [2].

In section 5 we describe the method of fixed points and Banach's contraction mapping principle. We then use said principle in an example to show existence of a unique solution for a differential equation.

Finally, in section 6, we engage in a brief overview of the Mountain Pass Theorem and an application to the pendulum equation. For more details on the Mountain Pass Theorem, see e.g. [4].

2 Prerequisites

Here we will list a few things that are needed later on. We start off with the definition of a Cauchy sequence.

Definition A sequence $\{x_n\}$ in a metric space X, with metric d, is said to be a **Cauchy sequence** if for every $\epsilon > 0$ there exists some integer N such that $d(x_n, x_m) < \epsilon$, if $n, m \geq N$.

Definition A metric space X is said to be **complete** if every *Cauchy sequence* has a limit in X. I.e. for any *Cauchy sequence* $\{x_n\}$, there exists some $x \in X$ such that $\lim\limits_{n\to\infty} x_n = x$.

Definition Given a vector space X over the real (or complex) numbers, a *norm* is a function $\|\cdot\| : X \to \mathbb{R}$ with the properties:

1. $\|a\mathbf{u}\| = |a|\|\mathbf{u}\|$ for all $a \in \mathbb{R}$ (or $\mathbb{C}$) and $u \in X$,
2. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ for all $u, v \in X$,
3. $\|\mathbf{u}\| \geq 0$ for all $\mathbf{u} \in X$, and $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u}$ is the zero vector.

When several vector spaces and their respective norms are considered, it is useful to denote the norm for a vector space X as $\|\cdot\|_X$.

We now look at a couple of different spaces we will be using:

Definition A vector space X over the real (or complex) numbers equipped with a norm, and such that X is complete with respect to that norm, is called a **Banach Space**.

Definition A **Hilbert Space** is a vector space X over the real (or complex) numbers with an inner product $\langle f, g\rangle$, such that the norm $\|f\| = \sqrt{\langle f, f\rangle}$ turns X in to a complete metric space.

Definition We define the space $L^p(a, b)$ as the completion of continuous functions in $[a, b]$ with respect to the norm

$$\|f\|_{L^p} = \left(\int_a^b |f(x)|^p \, dx\right)^{1/p}, \ 1 \leq p < \infty.$$

Two functions f, g on L^p are considered as equivalent if they are equal almost everywhere (a.e.). In other words, f is equivalent to g except on a set A of measure zero, i.e. for all $\epsilon > 0$, there exists a covering of A by intervals whose joint length is less than ϵ.

We also define the space $L^\infty(a, b)$ as the completion of continuous functions in $[a, b]$, but instead with respect to the norm $\|f\|_\infty := \sup\limits_{x\in[a,b]} |f(x)|$.

The norm in L^∞ then becomes

$$\|f\|_{L^\infty} = \operatorname*{ess\,sup}_{x\in[a,b]} |f(x)|,$$

where ess sup is the essential supremum of $|f|$, i.e. the smallest constant c such that $|f(x)| \le c$ a.e.

We also need the definition of a strong form of continuity:

Definition *(Lipschitz Continuity)* Let X and Y be two metric spaces, with metrics d_X and d_Y respectively. A map $f : X \to Y$ is said to be **Lipschitz continuous** if there exists a real constant $C > 0$ such that

$$d_Y(f(x_1), f(x_2)) \le C d_X(x_1, x_2), \text{ for all } x_1, x_2 \in X,$$

where the smallest such constant C is called the **Lipschitz constant** for f.

If for every $x \in X$ there exists a neighborhood N of x such that f restricted to N is Lipschitz continuous, then f is said to be **locally Lipschitz continuous**.

Another useful fact we need later on is that a linear operator is bounded if and only if it is continuous.

Lemma 2.1 *Let X and Y be two normed spaces, and let $G : X \to Y$ be a linear operator. Then G is continuous if and only if it is bounded.*

When we say that G is bounded we mean that there exists some $K > 0$ such that for all $u \in X$,

$$\|Gu\|_Y \le K\|u\|_X.$$

Proof Suppose G is bounded. Then for all $u, v \in X$,

$$\|Gu - Gv\|_Y \le K\|u - v\|_X.$$

By our definition above we see that G is Lipschitz continuous and hence continuous.

Now for the converse. Since G is continuous, it is continuous at 0. This implies that there exists some $\delta > 0$ such that $\|u\|_X \le \delta \Rightarrow \|Gu\|_Y \le 1$. For every non-zero $u \in X$ we see that

$$\begin{aligned} \left\| \frac{\delta u}{\|u\|_X} \right\|_X &= \delta \\ \Rightarrow \left\| \frac{\delta G u}{\|u\|_X} \right\|_Y &\le 1 \\ \Rightarrow \|Gu\|_Y &\le \frac{1}{\delta}\|u\|_X. \end{aligned}$$

This shows that G is bounded, and completes the proof.

Definition Let X be a real (or complex) vector space. A linear operator $G : X \to \mathbb{R}$ (or $\mathbb{C}$) is called a **linear functional**.

The set of all bounded linear functionals on a Banach space X is itself a Banach space and is called the dual space of X, which is denoted by X^*.

We now look at a few theorems we will be needing later on.

Theorem 2.2 (Poincaré-Miranda theorem) *Let*

$$f = (f_x, f_y) : [-1,1] \times [-1,1] \to \mathbb{R}^2$$

be continuous such that

$$f_x(-1, y) < 0 < f_x(1, y),$$

and

$$f_y(x, -1) < 0 < f_y(x, 1),$$

for all $x, y \in [-1, 1]$. Then there exists $c \in (-1,1) \times (-1,1)$ such that $f(c) = (0,0)$.

Before we continue we define a fixed point of a function $f(x)$, as a point x_0 such that $f(x_0) = x_0$. We also denote the open ball of radius $r > 0$, centered at the point c, as $B_r(c) = \{x \in X : d_X(x,c) < r\}$, where X is a metric space.

Theorem 2.3 (Brouwer's fixed point theorem) *Any continuous mapping*

$$f : \overline{B_1(0)} \to \overline{B_1(0)}$$

has at least one fixed point.

(For proofs of theorems 2.2, 2.3: see theorems 1.8, 1.7 respectively in [2].)

Here are a few inequalities to help us. First, Gronwall's inequality [3], theorem III.1.1.

Theorem 2.4 *Let $u(t), v(t)$ be two non-negative, continuous functions on $[a, b]$ and let $D \geq 0$ be a constant. Assume that*

$$v(t) \leq D + \int_a^t v(s)u(s)\, ds \qquad for\ a \leq t \leq b. \tag{2.1}$$

Then

$$v(t) \leq De^{I(t)} \qquad for\ a \leq t \leq b, \tag{2.2}$$

where $I(t) := \displaystyle\int_a^t u(s)\, ds$. In particular, if $D = 0$, then $v(t) \equiv 0$.

Proof For the case when $D > 0$: Let $V(t) := D + \int_a^t v(s)u(s)\,ds$. Then, $v(t) \leq V(t)$ by (2.1) and $V(t) \geq D$ on $[a, b]$. We note that $V'(t) = u(t)v(t) \leq u(t)V(t)$.

If we integrate $\frac{V'(t)}{V(t)} \leq u(t)$ over $[a, t]$, since $V(a) = D$, we get

$$V(t) \leq De^{I(t)},$$

where $I(t)$ is defined as above. So (2.2) follows from $V(t) \geq v(t)$.

For the second case, when $D = 0$: If (2.1) holds with $D = 0$, then our first case implies (2.2) for every $D > 0$. If we let D tend to 0 the desired result follows.

The last thing we need before we go on is the *Poincaré inequality.*

Lemma 2.5 *If f is a C^1 function, such that $f(0) = f(T) = 0$, then*

$$\|f(x)\|_{L^2} \leq \frac{T}{\pi}\|f'(x)\|_{L^2}.$$

For proof, see [2], lemma B.1.

In what follows all functions will be real-valued and all vector spaces will be over the real numbers.

3 The Pendulum Equation

A rather famous, often studied equation is that of the pendulum. In this section we will derive the equation.

3.1 Deriving the Equation

We start off with the force of gravity $\mathbf{F}_G$, directed downward, on the bob of the pendulum: $|\mathbf{F}_G| = mg$, where m is the mass of the bob, and g is the gravitational acceleration of the earth.

Using a moving coordinate system, with the origin at the center of the bob, the y-axis becomes aligned with the pendulum arm and the x-axis becomes tangent to the trajectory of the pendulum. In this coordinate system,

$$\mathbf{F}_P = -mg(\sin(\theta(t))\mathbf{i} + \cos(\theta(t))\mathbf{j}),$$

where $\theta(t)$ is the angle made by the pendulum arm, at the time t, relative to its resting position and $\mathbf{F}_P$ denotes the force acting on the pendulum. The vectors $\mathbf{i}$, $\mathbf{j}$ are the unit vectors for the x, y-axis respectively.

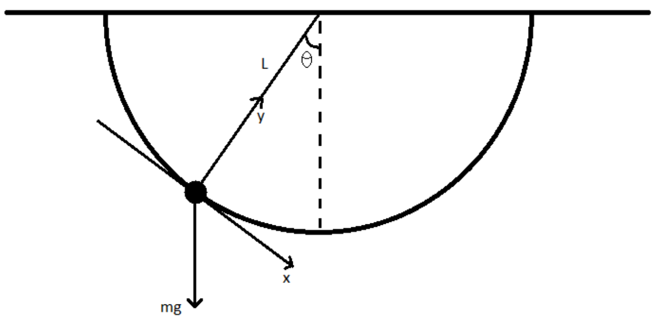
The tension T of the pendulum arm also acts on the bob, so the net force is equal to $\mathbf{F}_P + T\mathbf{j}$. Since the pendulum is forced to stay in a circular path, we conclude that the tension provided by the arm, $T\mathbf{j}$, is exactly cancelled by the $\mathbf{j}$-component of $\mathbf{F}_P$. Hence, the net force acting on the bob is equal to $-mg\sin(\theta(t))\mathbf{i}$.

By Newton's second law of motion we have $\mathbf{F} = m\mathbf{a}$. For our bob, the acceleration is equal to the length of the pendulum arm times the angular acceleration $\theta''(t)$. If we call the length of the arm L and denote the tangential component of $\mathbf{F}_P$ by $\mathbf{F}_T$, we get:

$$\mathbf{F}_T = mL\theta''(t)\mathbf{i} = -mg\sin\theta(t)\mathbf{i}.$$

Simplifying this expression we end up with our differential equation:

$$\theta''(t) = -\frac{g}{L}\sin\theta(t). \tag{3.1}$$



4 Shooting Method

Boundary value problems can be solved in an array of different methods, but here we will describe an elementary one that is called the *shooting method.* The gist of this method is to use a free parameter λ to solve an initial value problem, modified from our original problem, and then find the appropriate value(s) of λ such that the solution of our initial value problem takes the desired value on the boundary for the original problem.

Take for instance a second-order equation

$$u''(t) = f(t, u(t)), \tag{4.1}$$

with the conditions

$$u(0) = u(1) = 0. \tag{4.2}$$

Here is where the shooting comes in: We start off by solving (4.1) with the conditions

$$u(0) = 0, \quad u'(0) = \lambda, \tag{4.3}$$

for fixed $\lambda \in \mathbb{R}$. A solution u_λ, of our problem (4.1), (4.3), is well-defined and unique if we assume f to be continuous and locally Lipschitz in u.

Our problem has now been reduced to finding a zero of the function $\phi(\lambda) := u_\lambda(1)$. We now need to adjust the angle λ until an appropriate *shooting* angle is obtained, for which $u_\lambda(1) = 0$.

However, the solutions of the initial value problem may not be defined up to $t = 1$, so our $\phi(\lambda)$ may not be defined for all values of λ. However, if f is bounded we see, by writing (4.1) and (4.3) as a system of first order equations, that $\phi(\lambda)$ exists. In other words, if we let $u'(t) = v(t)$ and $v'(t) = f(t, u(t))$, where $u(0) = 0$ and $v(0) = \lambda$, then the result follows from theorem II.3.1 in [3].

4.1 Examples

To show how the shooting method can be used we will be looking at two examples.

Example 4.1.1 Let $g : \mathbb{R} \to \mathbb{R}$, $g \in C^1$. Also let g be bounded such that $g(0) = 0$ and $(2k-1)^2\pi^2 < g'(0) < (2k)^2\pi^2$ for some integer k. We will show that the Dirichlet problem

$$\begin{cases} u''(t) + g(u(t)) = 0, \\ u(1) = u(0) = 0, \end{cases} \tag{4.4}$$

has at least two different non-trivial solutions, i.e. solutions u such that $u \neq 0$.

First we consider the function $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(\lambda) = u_\lambda(1)$, where $\lambda := u'(0)$. Next we let $\lambda > \|g\|_\infty$, and then take the difference

$$u'_\lambda(t) - \lambda = u'_\lambda(t) - u'_\lambda(0) = \int_0^t u''_\lambda(s)\, ds = -\int_0^t g(u_\lambda(s))\, ds. \tag{4.5}$$

From what we know about the value of λ we see that

$$-\int_0^t g(u_\lambda(s))\, ds \geq -\int_0^t |g(u_\lambda(s))|\, ds \geq -\|g\|_\infty \int_0^t ds > -\lambda t.$$

Putting this together with (4.5) get $u'_\lambda(t) > \lambda(1-t)$. If we now integrate this inequality we see that

$$\int_0^t u'_\lambda(s)\, ds = u_\lambda(t) - u_\lambda(0) = u_\lambda(t) > \lambda\left(t - \frac{1}{2}t^2\right).$$

Finally, $\phi(\lambda) = u_\lambda(1) > \lambda(1 - \frac{1}{2}) = \frac{1}{2}\lambda > 0$.

On the other hand, if we look at $-\lambda$, we see that

$$u'_{-\lambda}(t) - (-\lambda) = u'_{-\lambda}(t) - u'_{-\lambda}(0) = \int_0^t u''_{-\lambda}(s)\, ds = -\int_0^t g(u_{-\lambda}(s))\, ds. \tag{4.6}$$

Following the same inequality chain as above, but with opposite signs, we get

$$-\int_0^t g(u_{-\lambda}(s))\, ds < \lambda t.$$

Again, combining this with (4.6) we see that $u'_{-\lambda}(t) < \lambda(t-1)$. Yet again we integrate:

$$\int_0^t u'_{-\lambda}(s)\, ds = u_{-\lambda}(t) - u_{-\lambda}(0) = u_{-\lambda}(t) < \lambda\left(\frac{1}{2}t^2 - t\right).$$

At last we see that $\phi(-\lambda) = u_{-\lambda}(1) < \lambda\left(\frac{1}{2} - 1\right) = -\frac{1}{2}\lambda < 0$.

So we have arrived at a situation where $\phi(-\lambda) < 0 < \phi(\lambda)$, for λ sufficiently large. Since $\phi(0) = 0$, by definition of ϕ, we have that

$$\phi'(0) = \lim_{\lambda \to 0} \frac{\phi(\lambda)}{\lambda} = \lim_{\lambda \to 0} \frac{u_\lambda(1)}{\lambda}.$$

Define $w_\lambda(t) := \dfrac{u_\lambda(t)}{\lambda}$, and $h_\lambda(t) := \dfrac{g(u_\lambda(t))}{u_\lambda(t)}$. Using this in the following equation

$$u''_\lambda(t) + g(u_\lambda(t)) = 0 \Leftrightarrow \frac{u''_\lambda(t)}{\lambda} + \frac{g(u_\lambda(t))}{u_\lambda(t)} \frac{u_\lambda(t)}{\lambda} = 0,$$

we get the equation

$$w''_\lambda(t) + h_\lambda(t) w_\lambda(t) = 0. \tag{4.7}$$

Observe that $w_\lambda(0) = 0$, $w'_\lambda(0) = \frac{u'_\lambda(0)}{\lambda} = \frac{\lambda}{\lambda} = 1$. We will later show that, when $\lambda \to 0$, the limiting problem is

$$\begin{cases} w''_0(t) + g'(0)w_0(t) = 0, \\ w_0(0) = 0, w'_0(0) = 1. \end{cases} \tag{4.8}$$

Solving this by standard methods, noting that the characteristic equation equals $r^2 + g'(0) = 0$, we get the solution $w_0(t) = \dfrac{\sin\sqrt{g'(0)}t}{\sqrt{g'(0)}}$.

By the restrictions we had on $g'(0)$ we see that $w_0(1) < 0$. Below we will show that $\|w_\lambda(t) - w_0(t)\|_\infty \to 0$ as $\lambda \to 0$. It will then follow that $\phi'(0) = \lim_{\lambda\to 0} w_\lambda(1) = w_0(1) < 0$.

Combining the above together with the fact that, for sufficiently large λ,

$$\phi(-\lambda) < \phi(0) = 0 < \phi(\lambda),$$

we can draw a rough picture of $\phi(\lambda)$ and see that it has at least two non-trivial zeroes, which is our desired result. We are left with proving that the limiting problem actually is (4.8).

We want to show that $\|w_\lambda(t) - w_0(t)\|_\infty \to 0$ as $\lambda \to 0$. In other words, we want to go from

$$w''_\lambda(t) + h_\lambda(t)w_\lambda(t) = 0,$$

to

$$w''_0(t) + g'(0)w_0(t) = 0.$$

We subtract one equation from the other and define

$$v_\lambda(t) := w_\lambda(t) - w_0(t),$$

and add the term $-h_\lambda(t)w_0(t)$ on both sides and get:

$$\begin{cases} v''_\lambda(t) + h_\lambda v_\lambda(t) = (g'(0) - h_\lambda(t))w_0(t), \\ v_\lambda(0) = v'_\lambda(0) = 0. \end{cases}$$

Next, we rewrite it as a system of first-order equations:

$$\begin{cases} v'_\lambda(t) = z_\lambda(t), \\ z'_\lambda(t) = (g'(0) - h_\lambda(t))w_0 - h_\lambda(t)v_\lambda(t), \\ v_\lambda(0) = z_\lambda(0) = 0. \end{cases} \tag{4.9}$$

Now, we integrate z'_λ and v'_λ:

1. $v_\lambda(t) = \int_0^t z_\lambda(s)\, ds$.

2. $z_\lambda(t) = -\int_0^t h_\lambda(s)v_\lambda(s)\, ds + \int_0^t (g'(0) - h_\lambda(s))w_0(s)\, ds$.

From 1. we see that

$$|v_\lambda(t)| \le \int_0^t |z_\lambda(s)|\, ds,$$

and from 2. we get

$$|z_\lambda(t)| \le \int_0^t C|v_\lambda(s)|\, ds + \epsilon_\lambda,$$

where $\epsilon_\lambda := \displaystyle\int_0^1 |g'(0) - h_\lambda(s)||w_0(s)|\, ds$ and $|h_\lambda(s)| \le C$, for small λ and a suitable constant C. We see that $\epsilon_\lambda \to 0$ when $\lambda \to 0$, since $u_\lambda(t) \to 0$ as $\lambda \to 0$ by continuous dependence of initial conditions, see section V.2 in [3].

As a final step we write

$$\begin{aligned} |v_\lambda(t)| + |z_\lambda(t)| &\le \epsilon_\lambda + \int_0^t |z_\lambda(s)|\, ds + \int_0^t C\, |v_\lambda(s)|\, ds \\ &\le \epsilon_\lambda + (1+C)\int_0^t (|v_\lambda(s)| + |z_\lambda(s)|\, ds. \end{aligned}$$

By Gronwall's inequality (theorem 2.4):

$$|v_\lambda(t)| + |z_\lambda(t)| \le \epsilon_\lambda e^{(1+C)t},$$

and hence $v_\lambda = w_\lambda - w_0 \to 0$ uniformly in t as $\lambda \to 0$, since $\epsilon_\lambda \to 0$ when $\lambda \to 0$.

Example 4.1.2 We will show that the forced pendulum equation with friction, under arbitrary Dirichlet conditions, has at least one solution. We write this as

$$\begin{cases} \theta''(t) + \frac{b}{m}\theta'(t) + \frac{g}{L}\sin\theta(t) = p(t), \\ \theta(0) = \theta_0,\ \ \theta(1) = \theta_1, \end{cases} \tag{4.10}$$

where the forcing term $p : [0,1] \to \mathbb{R}$ is continuous, b is a so called damping coefficient and g, L, m are the same as in section 3.1. For convenience we will define

$$\begin{cases} \alpha := \frac{b}{m}, \\ \beta := \frac{g}{L}, \\ f(t) := p(t) - \beta\sin\theta_\lambda(t). \end{cases}$$

As usual with the shooting method, we first look at the problem with $\lambda := \theta'(0)$:

$$\begin{cases} \theta''(t) + \alpha\,\theta'(t) = f(t), \\ \theta(0) = \theta_0,\ \ \theta'(0) = \lambda. \end{cases} \tag{4.11}$$

If we assume we already have a solution to (4.11), call it $\theta_\lambda(t)$, we multiply both sides of our equation by $e^{\alpha t}$ and get

$$(\theta_\lambda''(t) + \alpha\,\theta_\lambda'(t))e^{\alpha t} = \frac{d}{dt}(\theta_\lambda'(t)e^{\alpha t}) = f(t)e^{\alpha t},$$

which we then integrate, using the fact that $\theta'_\lambda(0) = \lambda$:

$$\int_0^s \frac{d}{dt}(\theta'_\lambda(t)e^{\alpha t})\,dt = \left[\theta'_\lambda(t)e^{\alpha t}\right]_0^s = \theta'_\lambda(s)e^{\alpha s} - \lambda = \int_0^s f(t)e^{\alpha t}\,dt.$$

Rearranging this, and dividing through by $e^{\alpha s}$, we get

$$\theta'_\lambda(s) = \lambda e^{-\alpha s} + \int_0^s f(t)e^{\alpha(t-s)}\,dt,$$

which we then integrate yet again, using the fact that $\theta_\lambda(0) = \theta_0$, and get

$$\int_0^r \theta'_\lambda(s)\,ds = \theta_\lambda(r) - \theta_0 = \lambda \int_0^r e^{-\alpha s}\,ds + \int_0^r \int_0^s f(t)e^{\alpha(t-s)}\,dt\,ds.$$

We now have an expression for our solution θ_λ:

$$\theta_\lambda(r) = \theta_0 + \frac{\lambda}{\alpha}(1 - e^{-\alpha r}) + \int_0^r \int_0^s [p(t) - \beta \sin\,\theta_\lambda(t)]e^{\alpha(t-s)}\,dt\,ds.$$

Defining $\phi(\lambda) := \theta_\lambda(1)$, we see that for R large enough, in $\phi(\pm R)$, the term $\frac{\pm R}{\alpha}(1 - e^{-\alpha})$ dominates the expression since the double integral is bounded, and hence

$$\phi(-R) < \theta_1 < \phi(R),$$

which by the intermediate value theorem tells us that there exists a point $c \in (-R, R)$ such that $\phi(c) = \theta_c(1) = \theta_1$, and we are done.

4.2 A Priori Bounds

Here we will look at more general cases and try to apply the shooting method. One idea we will be using is that of **a priori** bounds.

If we know that the solutions of a problem are bounded by some constant R, we may replace the function f by some bounded one. We must however be careful in order to ensure that the solutions of our modified problem are bounded by the same R, otherwise they would not solve our original problem. Here we will look at two examples.

Example 4.2.1 Let $f : [0, 1] \times \mathbb{R} \to \mathbb{R}$ be continuous and locally Lipschitz in its second variable. Assume there exists a positive constant $R > 0$ such that

$$f(t, -R) < 0 < f(t, R) \text{ for all } t \in [0, 1]. \tag{4.12}$$

Then (4.1)-(4.2) has at least one solution.

Now we define the following bounded function f_B:

$$f_B(t, u(t)) := \begin{cases} f(t, u(t)) & \text{if } |u(t)| \leq R, \\ f(t, R) & \text{if } u(t) > R, \\ f(t, -R) & \text{if } u(t) < -R. \end{cases}$$

Since f_B is bounded, continuous and locally Lipschitz in u, we see (as in example 4.1.2) that the shooting method provides a solution $\tilde{u}(t)$ of $u''(t) = f_B(t, u(t))$ satisfying (4.2). In summary, we are left with showing that $|\tilde{u}(t)| \leq R$ for all t.

In order to show this, assume $\tilde{u}$ achieves its maximum at some t_0, with $\tilde{u}(t_0) > R$, then $t_0 \in (0, 1)$ and

$$\tilde{u}''(t_0) = f_B(t_0, \tilde{u}(t_0)) = f(t_0, R) > 0,$$

which is a contradiction since we assumed it was a maximum.

If we on the other hand assume $\tilde{u}$ achieves its minimum at t_0, with $\tilde{u}(t_0) < -R$, then $t_0 \in (0, 1)$ and

$$\tilde{u}''(t_0) = f_B(t_0, \tilde{u}(t_0)) = f(t_0, -R) < 0.$$

Again, a contradiction, and hence we see that $|\tilde{u}(t)| \leq R$ for all t.

Note: The original problem may still admit more solutions, but we have shown that the absolute value of an arbitrary solution to the problem with f_B, is bounded by R.

The next example uses the monotonicity of a function to show uniqueness of a solution to (4.1)-(4.2). Showing the existence of a solution, however, is a bit trickier.

Example 4.2.2 Let $f : [0, 1] \times \mathbb{R} \to \mathbb{R}$ be continuous and locally Lipschitz in u. Now, let us also assume that

$$f(t, u) \leq f(t, v) \text{ for all } t \in [0, 1],$$

where $u, v \in \mathbb{R}$ and $u \leq v$. Then (4.1)-(4.2) has a unique solution.

To prove uniqueness, assume u, v solve our problem and write $w := u - v$. Then,

$$w''(t)w(t) = [f(t, u(t)) - f(t, v(t))](u(t) - v(t)) \geq 0,$$

and since $w(0) = w(1) = 0$, partial integration in the above inequality yields

$$-\int_0^1 w''(t)w(t)\,dt = -\left[w'(t)w(t)\right]_0^1 + \int_0^1 w'(t)^2\,dt = \int_0^1 w'(t)^2\,dt \leq 0.$$

From this we see that $w' \equiv 0$, which in turn shows that w is constant, and hence must be zero. We conclude that $u = v$ and hence the solution must be unique.

Now, do we know that a solution exists? Assume v solves our problem. We can then write

$$v''(t) = f(t, v(t)) - f(t, 0) + f(t, 0),$$

and since $v(0) = 0$, by monotonicity, we get

$$v''(t)v(t) = [f(t, v(t)) - f(t, 0)]v(t) + f(t, 0)v(t) \geq f(t, 0)v(t).$$

Integration on both sides yields

$$\int_0^1 v'(t)^2\, dt \leq -\int_0^1 f(t, 0)v(t)\, dt \leq \|v\|_\infty \int_0^1 |f(t, 0)|\, dt. \tag{4.13}$$

If we write $v(t) = \int_0^t v'(s)\, ds$ we see that, for all t,

$$-\int_0^1 (v'(s))^-\, ds \leq -\int_0^t (v'(s))^-\, ds \leq v(t),$$

and

$$v(t) \leq \int_0^t (v'(s))^+\, ds \leq \int_0^1 (v'(s))^+\, ds,$$

where $(v'(s))^+ := \max\{v'(s), 0\}$ and $(v'(s))^- := \max\{-v'(s), 0\}$. We recall that $0 = v(0) = v(1)$, and hence

$$0 = \int_0^1 v'(s)\, ds = \int_0^1 [(v'(s))^+ - (v'(s))^-]\, ds,$$

and so

$$\int_0^1 (v'(s))^-\, ds = \int_0^1 (v'(s))^+\, ds = \frac{1}{2}\int_0^1 [(v'(s))^+ + (v'(s))^-]\, ds = \frac{1}{2}\int_0^1 |v'(s)|\, ds.$$

We have found that

$$|v(t)| \leq \frac{1}{2}\int_0^1 |v'(s)|\, ds,$$

by combining our inequalities.

By the Cauchy-Schwarz inequality, we get

$$\frac{1}{2}\int_0^1 1 \cdot |v'(s)|\, ds \leq \frac{1}{2}\left(\int_0^1 1^2\, ds\right)^{1/2}\left(\int_0^1 |v'(s)|^2\, ds\right)^{1/2} = \frac{1}{2}\|v'(s)\|_{L^2},$$

and so

$$\|v(t)\|_\infty^2 \leq \frac{1}{4}\|v'(t)\|_{L^2}^2,$$

which together with the previous inequality, (4.13), yields

$$\|v(t)\|_\infty \leq \frac{1}{4}\int_0^1 |f(t, 0)|\, dt := R.$$

We can now define a cutoff function, as in example 4.2.1, and see that we actually have a (unique) solution $v(t)$ to (4.1)-(4.2), for our cutoff function f_B. Since f_B is nondecreasing in the second variable and $f_B(t, 0) = f(t, 0)$, we see that $v(t)$ also solves our original problem, since the same bound R applies.

4.3 Poincaré Maps

The purpose of this section will be to look at applications of the shooting method on periodic problems. To start things off, we will look at the first order system with $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}^2$ continuous and locally Lipschitz in $u \in \mathbb{R}^2$:

$$\begin{cases} u'(t) = f(t, u(t)), \\ u(1) = u(0). \end{cases} \tag{4.14}$$

Again, we let u_λ be the solution to the equation, but with initial value $\lambda \in \mathbb{R}^2$; $u_\lambda(0) = \lambda$. From here we try to show existence of a fixed point of the map $P(\lambda) := u_\lambda(1)$, i.e. existence of an appropriate value of λ such that $u_\lambda(1) = \lambda$.

Definition Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}^2$ and assume f satisfies the equation in (4.14). Let $u_\lambda(0) = \lambda$. The map $P(\lambda) := u_\lambda(1)$ is called the *Poincaré map*.

What happens should the solution not be defined up to $t = 1$? Well, in some cases we have enough information to show that there exists a fixed point in some subset of the plane, which the following example shows:

Example 4.3.1 We will show that if the function $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}^2$ is continuous and satisfies $f(t,v) \cdot v < 0$ for $|v| = R$, where $R > 0$ is a constant, then there exists at least one solution of (4.14). We also show that any such solution v, with $|v(0)| \le R$, satisfies $\|v\|_\infty \le R$.

Let $\lambda \in \mathbb{R}^2$, with $|\lambda| \le R$ and let v_λ be the solution to the modified problem

$$\begin{cases} v'(t) = f(t, v(t)), \\ v(0) = \lambda. \end{cases}$$

Assume v_λ is defined on $[0, K]$, for some K. We define $g(t) := |v(t)|^2$ and get the derivative $g'(t) = 2v(t) \cdot v'(t) = 2v(t) \cdot f(t, v(t))$. Letting $|v(t)| = R$ we see that $g'(t) < 0$, by our bound.

If we now assume $v(0) \in B_R(0)$, then $v(t)$ can never reach the boundary of $B_R(0)$ for any t. But on the other hand should $|v(0)| = R$, then $g(t)$ is at first decreasing and hence $v(t)$ must remain inside $B_R(0)$, for all $t \in (0, K]$.

What we gather from this is that v_λ is defined on $[0,1]$ for $|\lambda| \le R$ by theorem II.3.1 in [3], and hence the Poincaré map $P(\lambda)$ is well defined and continuous (according to this theorem $v_\lambda(t)$ is defined on $[0,1]$ unless $|v_\lambda(t)| \to \infty$ as $t \to a^+$ for some $a \le 1$). Also, $P(\overline{B_R(0)}) \subset \overline{B_R(0)}$, so Brouwer's fixed point theorem 2.3 shows that there exists a fixed point of P and hence a solution of (4.14).

Example 4.3.2 Let $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ be continuous and assume f satisfies (4.12), for some positive constant R. We want to solve the scalar problem

$$\begin{cases} u''(t) = f(t, u(t)), \\ u(0) = u(1), \\ u'(0) = u'(1). \end{cases}$$

The corresponding Poincaré map has two parameters: one for the initial value of u, and one for the initial value of u'.

We can then find a bounded continuous function f_B, even the same as in example 4.1.2, such that $f_B \equiv f$ over $[0,1] \times [-R, R]$ and

$$f_B(t, -u) < 0 < f_B(t, u),$$

for all $u \geq R$ and $t \in [0,1]$. Should u solve our problem we shall see that $\|u\|_\infty \leq R$.

Suppose u achieves its global maximum at t_0 with $u(t_0) > R$. If we take $t_0 \in (0,1)$, we see from the equation that $u''(t_0) = f_B(t_0, u(t_0)) > 0$, which is a contradiction. Hence, we see that

$$u(0) = u(1) = \max_{0 \leq t \leq 1} u(t).$$

So $u'(0) \leq 0 \leq u'(1)$ and, again from the boundary conditions, we see that $u'(0) = 0 = u'(1)$.

Since $u(0) > R$ we see from the equation that $u''(0) > 0$, and hence $u'(t)$ is initially increasing, i.e. u itself is initially increasing, since $u'(0) = 0$. But then u cannot achieve its maximum at $t = 0$. So we see that $u(t) \leq R$ for all t.

On the other hand, should u instead achieve its minimum at t_0 with $u(t_0) < -R$, and $t_0 \in (0,1)$, we see that $u''(t_0) = f_B(t_0, u(t_0)) < 0$, which is a contradiction. Hence

$$u(0) = u(1) = \min_{0 \leq t \leq 1} u(t),$$

and so $u'(0) \geq 0 \geq u'(1)$. By the same reasoning as above, $u''(0) < 0$, i.e. u is initially decreasing, which is impossible, and hence $u(t) \geq -R$ for all t.

What we conclude is that we only need to look at the modified problem, and verify that it has a solution. For fixed $\lambda, \gamma \in \mathbb{R}$, let $v_{\lambda,\gamma}$ be the unique solution to the problem

$$\begin{cases} v''(t) = f_B(t, v(t)), \\ v(0) = \lambda, \ v'(0) = \gamma. \end{cases}$$

We will be using a small trick of looking at zeroes of the function

$$g(\lambda, \gamma) := (v'_{\lambda,\gamma}(1) - \gamma, v_{\lambda,\gamma}(1) - \lambda), \tag{4.15}$$

instead of the ordinary Poincaré map $P(\lambda, \gamma) := v_{\lambda,\gamma}(1)$.

We have

$$\int_0^t v''_{\lambda,\gamma}(s)\, ds = \left[v'_{\lambda,\gamma}(s)\right]_0^t = v'_{\lambda,\gamma}(t) - \gamma.$$

So

$$v'_{\lambda,\gamma}(t) - \gamma = \int_0^t f_B(s, v_{\lambda,\gamma}(s))\, ds.$$

Hence we get

$$v_{\lambda,\gamma}(t) = \lambda + t\gamma + \int_0^t (t-s) f_B(s, v_{\lambda,\gamma}(s))\, ds,$$

since the derivative of $\displaystyle\int_0^t (t-s) f_B(s, v_{\lambda,\gamma}(s))\, ds$, with respect to t, is equal to $\displaystyle\int_0^t f_B(s, v_{\lambda,\gamma}(s))\, ds = v'_{\lambda,\gamma} - \gamma$ and $v_{\lambda,\gamma}(0) = \lambda$.

Now, let $K > \|f_B\|_\infty$. Since

$$\int_0^1 (1-s) f_B(s, v_{\lambda,\gamma}(s))\, ds \leq \|f_B\|_\infty \int_0^1 (1-s)\, ds = \frac{1}{2}\|f_B\|_\infty \leq \|f_B\|_\infty,$$

we get

$$v_{\lambda,K}(1) - \lambda = K + \int_0^1 (1-s) f_B(s, v_{\lambda,K}(s))\, ds \geq K - \|f_B\|_\infty > 0,$$

as well as

$$v_{\lambda,-K}(1) - \lambda = -K + \int_0^t (1-s) f_B(s, v_{\lambda,-K}(s))\, ds \leq -K + \|f_B\|_\infty < 0,$$

for all λ.

We now fix $M := R + 2K$, and let $\gamma \in [-K, K]$. Then,

$$v_{M,\gamma}(t) = M + \gamma t + \int_0^t (t-s) f_B(s, v_{M,\gamma}(s))\, ds > M - 2K = R.$$

Hence,

$$v'_{M,\gamma}(1) - \gamma = \int_0^1 f_B(t, v_{M,\gamma}(t))\, dt > 0,$$

since the conditions on f_B, i.e. (4.12), tell us that $f_B(t, v_{M,\gamma}(t)) > 0$. In a similar manner $v'_{-M,\gamma}(1) - \gamma < 0$. So, by theorem 2.2 there exists a zero of g and hence a solution to our problem. To be more precise, we have replaced the square $[-1,1] \times [-1,1]$ in theorem 2.2, with the rectangle $[-M, M] \times [-K, K]$.

5 Solutions by the Contraction Mapping Principle

We start off this section by returning to the map in the definition of Lipschitz continuity and especially what happens when the Lipschitz constant C is less than 1.

Definition *(Contraction map)* Let, X and Y be metric spaces. If a map $f : X \to Y$ is Lipschitz continuous with Lipschitz constant $C < 1$, we call f a **contraction**.

Theorem 5.1 (Banach's contraction mapping principle) *Let X be a complete metric space and $T : X \to X$ a contraction. Then T has a unique fixed point $\tilde{x}$. Further, if we arbitrarily choose a point $x_0 \in X$ and define a recursive sequence by $x_{n+1} := T(x_n)$, then $\tilde{x} = \lim_{n\to\infty} x_n$.*

Proof Argument taken from [2]. We note that $\forall x, y \in X$:

$$d(x, y) \leq d(x, T(x)) + d(T(x), T(y)) + d(y, T(y)),$$

and since T is a contraction, this becomes

$$(1 - C)d(x, y) \leq d(x, T(x)) + d(y, T(y)). \tag{5.1}$$

Note that this implies that if both x and y are fixed points of T, then $d(x, y) = 0$, i.e. $x = y$. So T has at most one fixed point. Define $T^n := T \circ T \circ \cdots \circ T$ (n times), which has a Lipschitz constant less than, or equal to, C^n. Now, let $x = x_n$ and $y = x_{n+k}$. Then, by (5.1), we get:

$$\begin{aligned} d(x_n, x_{n+k}) &\leq \frac{1}{1-C}(d(x_n, T(x_n)) + d(x_{n+k}, T(x_{n+k}))) \\ &= \frac{1}{1-C}(d(T^n(x_0), T^n(x_1)) + d(T^{n+k}(x_0), T^{n+k}(x_1))) \\ &\leq \frac{C^n}{1-C}d(x_0, x_1) + \frac{C^{n+k}}{1-C}d(x_0, x_1) \\ &= \frac{C^n + C^{n+k}}{1-C}d(x_0, x_1). \end{aligned}$$

We see now that for each $\epsilon > 0$ there exists an N such that if $n \geq N$, then $d(x_n, x_{n+k}) < \epsilon$. Hence our sequence is in fact a Cauchy sequence (see prerequisites), and converges to some $\tilde{x}$, since X is complete.

5.1 Using Banach's Theorem

We will here look at an example of how Banach's fixed point theorem can be used to show existence of solutions of some differential equations. The problem we will look at is this:

$$\begin{cases} -u''(t) = g(t, u(t)), \\ u(1) = u(0) = 0. \end{cases}$$

But before we dive in, we shall first look at what space we will be working in. A seemingly fitting space to define would be

$$C_0^1(0,1) = \{u \in C^1 : u(0) = u(1) = 0\},$$

equipped with the inner product $\langle u, v\rangle := \int_0^1 u'(t)v'(t)\,dt$. But this space is not actually complete, and since that is a requirement in theorem 5.1, we shall instead look to its completion; the Hilbert space $H_0^1(0,1)$.

It can be shown that this space, $H_0^1(0,1)$, consists of continuous functions u, that are almost everywhere differentiable, which satisfy $u(0) = u(1) = 0$ and whose derivative is in $L^2(0,1)$. Here, we shall only use the fact that such functions u satisfy the Poincaré inequality (lemma 2.5). For a detailed account of this type of spaces (called Sobolev spaces), see [5].

We can now return to the problem:

$$\begin{cases} -u''(t) = g(t, u(t)), \\ u(1) = u(0) = 0, \end{cases} \tag{5.2}$$

where g is continuous and Lipschitz continuous in u. As a reminder, this simply means that $|g(t, u) - g(t, v)| \leq C|u - v|$, for some constant C, much the same as we defined earlier. Now, here is what we want to prove:

Theorem 5.2 *Let g satisfy $|g(t, u) - g(t, v)| \leq C|u - v|$. If $C < \pi^2$ there exists a unique(!) solution to (5.2).*

Before we begin our proof, let us look at what properties a solution to our problem has. If y solves (5.2), then

$$-\int_0^1 y''(t)v(t)\,dt = \int_0^1 g(t, y(t))v(t)\,dt,$$

for all $v(t)$ smooth and such that $v(0) = v(1) = 0$. I will henceforth leave out the variable t, since it is clear to be the variable of integration. Partial integration yields

$$-\int_0^1 y''v\,dt = \left[-y'v\right]_0^1 + \int_0^1 y'v'\,dt,$$

but since $v(0) = v(1) = 0$ we simply get $\int_0^1 y'v'\,dt = \int_0^1 g(t, y)v\,dt$.

From now on we let $y, v \in H_0^1$, which allows us to write

$$\langle y, v\rangle = \int_0^1 g(t, y)v\,dt.$$

The map $G^*(y)\colon v \mapsto \displaystyle\int_0^1 g(t,y)v\,dt$ is a linear functional, since

$$G^*(y)(u+v) = \int_0^1 g(t,y)(u+v)\,dt = \int_0^1 g(t,y)u\,dt + \int_0^1 g(t,y)v\,dt,$$

and, for any real number α,

$$G^*(y)(\alpha u) = \int_0^1 g(t,y)\alpha u\,dt = \alpha \int_0^1 g(t,y)u\,dt.$$

Unless stated otherwise, the norm we will be using from now on is that in $H_0^1(0,1)$, where $\|y\|_{H_0^1(0,1)} = \|y\| = \sqrt{\langle y, y\rangle}$.

We can now start looking at the proof of our theorem. Since g is Lipschitz continuous in the second variable, we have $|g(t,y) - g(t,0)| \le C|y|$. So $|g(t,y)| \le C|y| + |g(t,0)| \le C|y| + D$ for some constants C, D, and

$$\begin{aligned}
|G^*(y)u| &= \left|\int_0^1 g(t,y)u\,dt\right| \\
&\le \int_0^1 (C|y| + D)|u|\,dt \le \text{ [by Cauchy-Schwarz]} \\
&\le (C\|y\|_{L^2} + D)\,\|u\|_{L^2} \le \text{ [by Poincaré]} \\
&\le \frac{1}{\pi}\left(\frac{C}{\pi}\|y'\|_{L^2} + D\right)\|u'\|_{L^2} \\
&= \frac{1}{\pi}\left(\frac{C}{\pi}\|y\| + D\right)\|u\| = E\,\|u\|,
\end{aligned}$$

for constants C, D, E. Hence, $G^*(y)$ is bounded.

This shows that $G^*(y)$ is in the dual space $H_0^1(0,1)^*$. By *Riesz' representation theorem* from functional analysis, there exists a $G(y) \in H_0^1(0,1)$ such that $G^*(y)v = \langle G(y), v\rangle$.

Bringing all our acquired knowledge together, we see that (5.2) is equivalent to

$$\langle y, v\rangle = \langle G(y), v\rangle = \int_0^1 g(t,y)v\,dt, \text{ for all } v \in H_0^1(0,1),$$

and hence to $y = G(y)$. We might now ask ourselves if G is a contraction?

To show that it is, we start off with y_1, y_2 and

$$\begin{aligned}
\|G(y_1) - G(y_2)\| &= \sup_{\|v\|\leq 1} \langle G(y_1) - G(y_2), v\rangle \\
&= \sup_{\|v\|\leq 1} \int_0^1 \big[g(t,y_1) - g(t,y_2)\big] v \, dt \\
&\leq \sup_{\|v\|\leq 1} \int_0^1 |g(t,y_1) - g(t,y_2)| \, |v| \, dt \\
&\leq C \sup_{\|v\|\leq 1} \int_0^1 |y_1 - y_2| \, |v| \, dt.
\end{aligned}$$

Using the Cauchy-Schwarz inequality we see that

$$C \int_0^1 |y_1 - y_2| \, |v| \, dt \leq C \|y_1 - y_2\|_{L^2} \|v\|_{L^2}.$$

By the Poincaré inequality, and since $\|v\| \leq 1$, we get

$$\begin{aligned}
C\|y_1 - y_2\|_{L^2}\|v\|_{L^2} &\leq \frac{C}{\pi^2}\|y_1' - y_2'\|_{L^2}\|v'\|_{L^2} \\
&= \frac{C}{\pi^2}\|y_1 - y_2\|\|v\| \\
&\leq \frac{C}{\pi^2}\|y_1 - y_2\|.
\end{aligned}$$

In summary we have that

$$\|G(y_1) - G(y_2)\| \leq \frac{C}{\pi^2}\|y_1 - y_2\|,$$

and by our assumption that $C < \pi^2$, we see that we in fact have a contraction, and by theorem 5.1 we hence have a unique solution to (5.2), which is what we wanted to prove.

6 The Mountain Pass Theorem and the Pendulum Equation

6.1 Least-action Principle

In this section we return to the pendulum, $u''(t) + a\sin(u(t)) = p(t)$, where the forcing term p is a 2π-periodic function and $\int_0^{2\pi} p(t)\,dt = 0$. This time we consider the periodic conditions

$$\begin{cases} u(0) = u(2\pi), \\ u'(0) = u'(2\pi). \end{cases}$$

We will be working in the space $H^1_{per}(0, 2\pi)$, defined as the closure of the set $\{u \in C^1(0, 2\pi) : u(0) = u(2\pi)\}$ with respect to the norm corresponding to the inner product $\langle u, v\rangle = \int_0^{2\pi} (u'(t)v'(t) + u(t)v(t))\,dt$.

Define the functional

$$J(u) := \frac{1}{2}\int_0^{2\pi} (u'(t))^2\,dt - a\int_0^{2\pi} (1 - \cos u)\,dt + \int_0^{2\pi} u(t)p(t)\,dt.$$

I will henceforth not write out our variable t, for simplicity. We now look at

$$J(u+sv) = \frac{1}{2}\int_0^{2\pi} (u')^2 + 2su'v' + s^2v^2\,dt - a\int_0^{2\pi} (1 - \cos(u+sv))\,dt + \int_0^{2\pi} p(u+sv)\,dt,$$

and look at the directional derivative of J at u, in the direction v:

$$\left.\frac{d}{ds}\right|_{s=0} J(u+sv) = \frac{1}{2}\int_0^{2\pi} u'v'\,dt - a\int_0^{2\pi} v\sin(u)\,dt + \int_0^{2\pi} vp\,dt. \qquad (6.1)$$

This formal computation can be justified, see e.g. [1]. If we assume u to be C^2, we see from our conditions that

$$\int_0^{2\pi} u'v'\,dt = \big[u'v\big]_0^{2\pi} - \int_0^{2\pi} u''v\,dt = -\int_0^{2\pi} u''v\,dt.$$

So, the right hand side of (6.1) is equal to $\int_0^{2\pi} (-u'' - a\sin u + p)v\,dt$. We see that if the right hand side of (6.1) is 0 for all $v \in H^1_{per}(0, 2\pi)$ then u solves our problem.

If we instead consider when u is not necessarily C^2, we say that $u \in H^1_{per}(0, 2\pi)$ is a weak solution if the right hand side of (6.1) is 0 for all $v \in H^1_{per}$. Fortunately, one can show that a weak solution is actually equal to an ordinary solution, i.e. C^2.

Definition *(Weak Convergence in Hilbert Space)* For $u_n, u \in H$, where H is a Hilbert Space, we say that u_n converges weakly to u, denoted $u_n \rightharpoonup u$, if $\langle u_n, v\rangle_H \to \langle u, v\rangle_H$ for all $v \in H$.

We now look at a few facts about weak convergence for our problem at hand. The first fact is valid for any Hilbert space H, but the second and third fact are both only valid for our problem.

1. If the sequence (u_n) is bounded in H it implies that there exists a subsequence (u_{n_j}) in H that converges weakly to some u. Moreover, $\liminf\limits_{j\to\infty} \|u_{n_j}\| \geq \|u\|$.

2. If $u_n \rightharpoonup u$ then $\displaystyle\int_0^{2\pi} (1-\cos u_n)\, dt \to \int_0^{2\pi} (1-\cos u)\, dt$.

3. If $u_n \rightharpoonup u$ then $N(u_n) \to N(u)$, where $\langle N(u), v\rangle := \displaystyle\int_0^{2\pi} (\sin u) v\, dt$.

The last two facts follow from more general results which may be found in [1]. The important part here is that the embedding of H^1_{per} in L^2 is compact. In other words, $u_n \rightharpoonup u$ in H^1_{per} implies $u_n \to u$ in L^2, see for example [5] or [1].

Returning to our problem, we write $u = \bar{u} + \tilde{u}$ where $\bar{u} = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} u\, dt$, and $\tilde{u}$ has mean value 0 over the period 2π. We note that $J(u+2k\pi) = J(u)$, so if u is a solution, then so is $u + 2k\pi$. We say that two solutions u_1, u_2 are geometrically different if $u_1 - u_2 \neq 2k\pi$ for any $k \in \mathbb{Z}$.

Now, for every $u \in H^1_{per}$ there is a unique element $\nabla J(u) \in H^1_{per}$, such that $\langle \nabla J(u), v\rangle$ is equal to the right hand side of (6.1). The scalar product used is that of H^1_{per}.

Lemma 6.1 $\|\tilde{u}\|_{L^2} \leq \|\tilde{u}'\|_{L^2}$.

The proof of this follows by using the Fourier expansion of $\tilde{u}$ and Parseval's equality. See lemma B.2 and the remarks following it in [2].

Theorem 6.2 *The functional J assumes a smallest value. If $J(u)$ is said smallest value, then $\nabla J(u) = 0$.*

Proof For the second part, assume $J(u)$ is the smallest value. Then

$$0 \leq \lim_{s\to 0^+} \frac{J(u+sv) - J(u)}{s} = \left.\frac{d}{ds}\right|_{s=0^+} J(u+sv) = \text{ RHS of (6.1).}$$

But, we know that the RHS of (6.1) is equal to $\langle \nabla J(u), v\rangle$, for which we have $\langle \nabla J(u), v\rangle \geq 0$ for all v. So, we must have $\nabla J(u) = 0$.

The easiest way to see the equality is to simply take $v = -\nabla J(u)$:

$$0 \le \langle \nabla J(u), -\nabla J(u) \rangle = -\|\nabla J(u)\|^2 \le 0.$$

Since p has mean value 0 over the period 2π, we have $\displaystyle\int_0^{2\pi} up\,dt = \int_0^{2\pi} \tilde{u}p\,dt$, and obviously $\tilde{u}' = u'$. We may therefore rewrite J as

$$J(u) = \frac{1}{2}\int_0^{2\pi} (\tilde{u}')^2\,dt - a\int_0^{2\pi} (1-\cos u)\,dt + \int_0^{2\pi} \tilde{u}p\,dt.$$

Now,

$$\begin{aligned} J(u) &= \frac{1}{2}\int_0^{2\pi} (\tilde{u}')^2\,dt - a\int_0^{2\pi} (1-\cos u)\,dt + \int_0^{2\pi} \tilde{u}p\,dt \\ &\ge \frac{1}{2}\|\tilde{u}'\|_{L^2}^2 - 4a\pi - \|p\|_{L^2}\|\tilde{u}\|_{L^2} \\ &\ge \frac{1}{2}\|\tilde{u}'\|_{L^2}^2 - 4a\pi - \|p\|_{L^2}\|\tilde{u}'\|_{L^2}. \end{aligned}$$

We see that $J(u)$ is bounded from below. We also see that $J(u_n) \to \infty$ as $\|\tilde{u}_n'\|_{L^2} \to \infty$. If we let $J(u_n) \to \inf J$, then $\|\tilde{u}_n'\|_{L^2}^2$ is bounded. Since the cosine is 2π-periodic we can subtract appropriate multiples of 2π from $\bar{u}_n$ without changing the value of J; we get $\bar{u}_n \in [0, 2\pi]$. Hence $\|u_n\|_{H^1_{per}}$ is bounded.

By going to a subsequence we may assume that $u_n \rightharpoonup u$. By using 1. and 2. above, we get

$$\begin{aligned} \inf J &= \lim_{n\to\infty} J(u_n) \\ &= \lim_{n\to\infty} \frac{1}{2}\int_0^{2\pi} (u_n')^2\,dt - a\int_0^{2\pi} (1-\cos u_n)\,dt + \int_0^{2\pi} pu_n\,dt \\ &\ge \int_0^{2\pi} (u')^2 - a\int_0^{2\pi} (1-\cos u)\,dt + \int_0^{2\pi} pu\,dt \\ &= J(u) \ge \inf J. \end{aligned}$$

We see that we must have equality, and hence we have found a solution.

6.2 Climbing the Pass

This is a non-thorough overview of the mountain pass theorem (MPT) itself. We will not prove any of the theorems.

We will start off with a condition we will be using throughout this section. Let $J \in C^1(H, \mathbb{R})$, where H is a Hilbert space, and let (u_n) be any sequence such that $|J(u_n)|$ is bounded and $\nabla J(u_n) \to 0$ as $n \to \infty$. If, for each such (u_n), the sequence has a convergent subsequence (u_{n_k}), then J is said to satisfy the Palais-Smale Condition (PS).

If the sequence $\nabla J(u_{n_k}) \to 0$, and if $u_{n_k} \to u$, we have that $\nabla J(u_{n_k}) \to \nabla J(u)$. This means that $\nabla J(u) = 0$, i.e. u is a critical point for J. Now we move on to the classic version of the (MPT).

Theorem 6.3 (Mountain Pass Theorem [4, Theorem 7.1]) *Let $J \in C^1(H, \mathbb{R})$ be such that*

1. *$J(0) = 0$,*
2. *$J(u) \geq \alpha > 0$ for all u with $\|u\|_H = \rho$,*
3. *there exists some e, with $\|e\|_H > \rho$, such that $J(e) \leq 0$.*

Define

$$\Gamma := \{h \in C([0,1] \times H) : h(0) = 0, h(1) = e\}$$

and

$$c := \inf_{h \in \Gamma} \max_{t \in [0,1]} J(h(t)),$$

then $c \geq \alpha$. If J satisfies (PS), then c is a critical value for J. In other words, there exists some u such that $J(u) = c$ and $\nabla J(u) = 0$.

As the name hints, there is a geometric interpretation of the MPT. As a real world example, imagine that we are standing at the lowest point in a valley (i.e. the origin) and set out to try to climb the mountains around us in order to get out of the valley. The theorem then ensures that there is a lowest mountain pass which seperates the valley we are in from the region outside.

6.3 Generalization of the MPT, and the Pendulum

Suppose u_0 is a strict local minimum of J, and that $J(u_1) \leq J(u_0)$ for some $u_1 \neq u_0$. If J satisfies (PS), then $c > J(u_0)$ is a critical value. Here,

$$\Gamma := \{h \in C([0,1] \times H) : h(0) = u_0, h(1) = u_1\}$$

and

$$c = \inf_{h \in \Gamma} \max_{t \in [0,1]} J(h(t)).$$

This follows from more general results which may be found in [4].

Returning to the pendulum equation, we recall the Hilbert space $H^1_{per}(0, 2\pi)$ and the functional

$$J(u) = \frac{1}{2} \int_0^{2\pi} (u'(t))^2 \, dt - a \int_0^{2\pi} (1 - \cos u(t)) \, dt + \int_0^{2\pi} p(t) u(t) \, dt,$$

where p is 2π-periodic and has mean value 0 over one such period.

Since $J(u(t)) = J(u(t) + 2k\pi)$, we note that if u_0 is a minimizer of J, then $u_1 = u_0 + 2k\pi$, for any $k \in \mathbb{Z}$, is also a minimizer of J.

Theorem 6.4 *There exists a solution v_0 of J such that $v_0 \neq u_0 + 2k\pi$, for any $k \in \mathbb{Z}$.*

This theorem has been proven in [6, Corollary 4]. To give a short proof sketch, suppose u_0 is a minimizer and that J has a strict local minimum in u_0 (and u_1), since there are infinitely many distinct solutions otherwise. We show that (PS) is satisfied. Let (u_n) be such that $J(u_n) \to c$ and $\nabla J(u_n) \to 0$. If $\|u_n'\|_{L^2} \to \infty$, then $J(u_n) \to \infty$. It follows that $\|u_n'\|_{L^2}$ is bounded.

Recall that we can write $u = \tilde{u} + \bar{u}$ and also that we may assume that $\bar{u} \in [0, 2\pi]$. It follows that (u_n) is bounded in H^1_{per}, so there exists some subsequence (u_{n_k}) such that $u_{n_k} \rightharpoonup v_0$ for some v_0.

In this case we know that in L^2, $u_{n_k} \to v_0$. Together with fact 3. of section 6.1 and the fact that $\nabla J(u_{n_k}) \to 0$, this implies that $u_{n_k} \to v_0$ in H^1_{per}, i.e. J satisfies (PS). In other words $J(v_0) = c$ and $\nabla J(v_0) = 0$. Moreover, $v_0 \neq u_0 + 2k\pi$ since $J(v_0) > J(u_0)$.

Remark When $p = 0$ we get

$$J(u) = \int_0^{2\pi} (u'(t))^2 \, dt - a \int_0^{2\pi} (1 - \cos u) \, dt,$$

which is minimal if $u(t) = \pi$. This corresponds to having the pendulum standing straight up! But what does this tell us? Well, we see that some minimizing solutions may not even be physically relevant and hence looking for other solutions is meaningful.

References

[1] A. Ambrosetti and G. Prodi: A Primer of Nonlinear Analysis, Cambridge University Press 1995.

[2] P. Amster: Topological Methods in the Study of Boundary Value Problems, Springer 2014.

[3] P. Hartman: Ordinary Differential Equations, Wiley 1964.

[4] Y. Jabri: The Mountain Pass Theorem, Cambridge University Press 2003.

[5] L. Lidman Bergqvist: An Introduction to Sobolev Spaces, Bachelor thesis, Stockholm University 2014.

[6] J. Mawhin and M. Willem: Multiple Solutions of the Periodic Boundary Value Problem for Some Forced Pendulum-Type Equations, Journal of Differential Equations 52 (1984), 264-287.