Discussion Notes for Methods of Applied Mathematics (MATH 146)

UCLA

Fall 2017

Created by

Howard Heaton

Purpose: This document is a compilation of notes generated for discussion in MATH 146 with reference credit due to J. David Logan's text *Applied Mathematics*. If the reader finds any errors/typos, please feel free to email me at `heaton@math.ucla.edu` and I will address these and post an updated set of notes to my webpage.

Contents

Section 1: Introduction

These notes are provided to compliment the TA discussion sessions on Thursdays for MATH 146. Typically, more detail is provided here than on the board during discussion. And, the solutions to problems provided here illustrate the level of rigor desired from students this quarter.

Section 2: Review Material

Definition: Define $f : [a, b] \to \mathbb{R}$. For any $x \in [a, b]$, define the quotient

$$\phi(t) := \frac{f(t) - f(x)}{t - x} \quad (a < t < b,\ t \neq x), \tag{1}$$

and define

$$f'(x) := \lim_{t \to x} \phi(t), \tag{2}$$

provided the limit exists. We associate the function f' with f at the points where the limit (2) exists. The function f' is called the *derivative* of f. If f' is defined at a point x, we say f is *differentiable* at x. And if f' is defined at every point in a set $I \subset [a, b]$, then we say f is differentiable on I. △

Example 1: Use the above definition to compute $f'(1)$ for the function $f(x) = x^2$.

Solution:

Through direct computation, we find

$$\begin{aligned} f'(1) &= \lim_{t \to 1} \frac{f(t) - f(1)}{t - 1} \\ &= \lim_{h \to 0} \frac{f(1 + h) - f(1)}{(1 + h) - 1} \\ &= \lim_{h \to 0} \frac{(1 + h)^2 - 1}{h} \\ &= \lim_{h \to 0} \frac{1 + 2h + h^2 - 1}{h} \\ &= \lim_{h \to 0} \frac{2h + h^2}{h} \\ &= \lim_{h \to 0} 2 + h \\ &= 2 + 0 \\ &= 2. \end{aligned} \tag{3}$$

□

Taylor's Theorem: Let $I \subset \mathbb{R}$ be a neighborhood of x_0 and n be a nonnegative integer. Suppose the function $f : I \to \mathbb{R}$ has $n+1$ derivatives. Then, for each point $x \neq x_0$ in I, there is a point ξ strictly between x and x_0 such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \tag{4}$$

△

Remark 1: The second term in (4) is known as the *Lagrange Remainder.* ⋄

Consider using Taylor's theorem when $n = 1$. That is, suppose f is twice differentiable at x and define

$$\varepsilon(h) := \frac{f^{(2)}(\xi(h))}{2}h^2 \tag{5}$$

where $\xi(h)$ is the point strictly between x and $x + h$ such that

$$f(x + h) = f(x) + f'(x)h + \varepsilon(h), \tag{6}$$

which we know exists by Taylor's theorem. This form of expansion will be useful for us to remember when we look at differentiation of more abstract quantities known as functionals. Furthermore, this shows

$$f'(x) = \lim_{h\to 0} f'(x) = \lim_{h\to 0}\left(\frac{f(x+h) - f(x) - \varepsilon(h)}{h}\right) = \lim_{h\to 0}\left(\frac{f(x+h) - f(x)}{h} - \frac{\varepsilon(h)}{h}\right) = f'(x) - \lim_{h\to 0}\frac{\varepsilon(h)}{h}. \tag{7}$$

Thus $\lim_{h\to 0} \varepsilon(h)/h = 0$. Using little-oh notation (defined below), we write this as $\varepsilon(h) = o(h)$.

Definition: Assume $g(x)$ is nonzero. Then we say $f(x) = o(g(x))$ as $x \longrightarrow x^*$ provided

$$\lim_{x\to x^*}\left|\frac{f(x)}{g(x)}\right| = 0. \tag{8}$$

This notation is referred to as *little-oh notation.* △

Example 2: Define $f(x) := x^2$. Express $f(x+h)$ explicitly in the form of (6).

Solution:

First observe $f'(x) = 2x$ and $f''(x) = 2$. Then we see

$$f(x+h) = (x+h)^2 = x^2 + 2xh + h^2 = f(x) + f'(x)h + \varepsilon(h) \tag{9}$$

where $\varepsilon(h) := h^2$. □

We now turn our attention to a necessary condition for a point $\overline{x}$ to be a local minimizer of f.

Theorem: If $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function and $\overline{x}$ is a local minimizer of f, then $f'(\overline{x}) = 0$. △

Proof:

Let $\overline{x}$ be a minimizer of f, i.e., there is a $\delta^* > 0$ such that $f(\overline{x}) \leq f(x)$ for all $x \in (\overline{x} - \delta^*, \overline{x} + \delta^*)$. We proceed by way of contradiction, i.e., suppose $f'(\overline{x}) \neq 0$. By hypothesis f' is continuous, and so there is a $\delta > 0$ such that

$$|z - \overline{x}| < \delta \quad \Rightarrow \quad \left|f'(z) - f'(\overline{x})\right| < \frac{|f'(\overline{x})|}{2}. \tag{10}$$

But, using the reverse triangle inequality, we see

$$|f'(\overline{x})| - |f'(z)| \leq \left|f'(z) - f'(\overline{x})\right| < \frac{|f'(\overline{x})|}{2} \quad \Rightarrow \quad \frac{|f'(\overline{x})|}{2} < |f'(z)|. \tag{11}$$

Suppose $f'(\overline{x}) > 0$ and pick $z \in (\overline{x} - \delta/2, \overline{x})$. Taylor's theorem asserts there is $\xi \in (z, \overline{x})$ such that

$$f(z) = f(\overline{x}) + f'(\xi)(z - \overline{x}) = f(\overline{x}) - f'(\xi)|z - \overline{x}| < f(\overline{x}) - \frac{|f'(\overline{x})|}{2}|z - \overline{x}| < f(\overline{x}). \tag{12}$$

This shows $f(z) < f(\overline{x})$ for all $z \in (\overline{x} - \delta/2, \overline{x})$. Thus $\overline{x}$ cannot be a local minimizer of f, contradicting our initial assumption. Whence $f'(\overline{x}) \leq 0$. By analogous argument to above, if instead $f'(\overline{x}) < 0$, we pick $z \in (\overline{x}, \overline{x} + \delta/2)$ to deduce

$$f(z) = f(\overline{x}) + f'(\xi)(z - \overline{x}) = f(\overline{x}) + f'(\xi)|z - \overline{x}| < f(\overline{x}) - \frac{|f'(\overline{x})|}{2}|z - \overline{x}| < f(\overline{x}), \tag{13}$$

again giving a contradiction. This shows $f'(\overline{x}) \geq 0$. Therefore, combining our results, we conclude $f'(\overline{x}) = 0$, as desired. ■

Remark 2: The above theorem shows that a necessary condition for $\overline{x}$ to be a local minimizer of f is that $f'(\overline{x}) = 0$. Below we provide several examples illustrating the use and limitations of this theorem. ⋄

Example 3: Define $f(x) = (x-3)^2 + 5x + 3$. Solve the optimization problem

$$\min_{x \in \mathbb{R}} f(x), \tag{14}$$

using only the above theorem and definition of a minimizer.

Solution:

First note f is continuously differentiable since it is a polynomial. And,

$$f'(x) = 2(x-3) + 5 + 0 = 2x - 1. \tag{15}$$

The single critical point of f is at $x = 1/2$. The above theorem shows this is the only candidate solution to the optimization problem.

All that remains is to verify $x = 1/2$ is, in fact, a minimizer. We can rewrite f as $f(x) = x^2 - x + 12$. Pick any $z \in \mathbb{R}$ and set $\delta := z - 1/2$ so that $z = 1/2 + \delta$. Then

$$\begin{aligned} f(z) = f\left(\frac{1}{2} + \delta\right) &= \left(\frac{1}{2} + \delta\right)^2 - \left(\frac{1}{2} + \delta\right) + 12 \\ &= \left(\frac{1}{4} + \delta + \delta^2\right) - \left(\frac{1}{2} + \delta\right) + 12 \\ &= \left(\frac{1}{4} - \frac{1}{2} + 12\right) + \delta^2 \\ &= f\left(\frac{1}{2}\right) + \delta^2 \\ &\geq f\left(\frac{1}{2}\right). \end{aligned} \tag{16}$$

This shows $f(1/2) \leq f(z)$ for all $z \in \mathbb{R}$, i.e., $1/2$ is the global minimizer of f, and we are done. □

Example 4: Define $f(x) = x^3$. Can the above theorem be applied to find a local minimum?

Solution:

Observe $f'(x) = 3x^2$ and so $f'(x) = 0$ if and only if $x = 0$. But, $f(0) = 0 > -\varepsilon^3 = f(-\varepsilon)$ for every $\varepsilon > 0$ and so 0 is not a local minimum of f. Thus the above theorem cannot be applied to find a local minimum. Moreover, because this was the only candidate for a minimizer, we are able to further conclude f has no global minimizer over $\mathbb{R}$. □

REMARK 3: The above theorem shows that the condition $f'(x) = 0$ is necessary, but *not* sufficient. We illustrate this again with the following example. ◇

Example 5: Define $f(x) = -x^2$. Can the above theorem be applied to find a local minimum?

Solution:

Observe $f'(x) = -2x$ and so $f'(x) = 0$ if and only if $x = 0$. But, $f(z) = -z^2 < 0 = f(0)$ for all $z \neq 0$. This shows 0 is not a local minimum of f. Thus the above theorem cannot be applied to find a local minimum. In fact, the above shows $x = 0$ is a global maximizer of f. □

Example 6: Define $f(x) := 3|x - 5|$. What is the global minimizer of f and can the above theorem be applied? Explain.

Solution:

The global minimizer is $x = 5$. Indeed,

$$f(5) = 0 \leq 3|x-5| = f(x) \quad \forall\, x \in \mathbb{R}. \tag{17}$$

However, f is not continuously differentiable since f' is not continuous at $x = 5$. Indeed,

$$\lim_{x \to 5-} f'(x) = -3 \neq 3 = \lim_{x \to 5+} f'(x). \tag{18}$$

Thus a condition for the theorem does not hold and so it cannot be applied. □

Remark 4: Here we review integration by parts. Let $f, g \in C^1[a,b]$. Then using the product rule we write

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)g(x)\right] = f'(x)g(x) + f(x)g'(x). \tag{19}$$

So,

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)g(x)\right] \ \mathrm{d}x = \int_a^b f'(x)g(x) + f(x)g'(x) \ \mathrm{d}x. \tag{20}$$

But, the left hand side can be rewritten as

$$\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left[f(x)g(x)\right] \ \mathrm{d}x = \int_{f(a)g(a)}^{f(b)g(b)} \mathrm{d}\,(fg) = \left[f(x)g(x)\right]_{x=a}^{x=b} = f(b)g(b) - f(a)g(a). \tag{21}$$

Thus the integration by parts formula becomes

$$\int_a^b f(x)g'(x) \ \mathrm{d}x = -\int_a^b f'(x)g'(x) \ \mathrm{d}x + \left[f(x)g(x)\right]_{x=a}^{x=b}. \tag{22}$$

This will be especially useful tool for us and is important to have at our disposal. ⋄

SECTION 3: FUNCTIONALS $J : V \to \mathbb{R}$

Definition: A *functional* $J : V \to F$ is a mapping from a vector space V to a field of scalars F, e.g., the real numbers $\mathbb{R}$. $\triangle$

Definition: We define the space $C^n[a, b]$ to be the set of all functions $f : [a, b] \to \mathbb{R}$ such that $f^{(n)}$ is continuous. $\triangle$

Definition: We say $y_0 \in \mathcal{A}$ is a **local minimizer** for J over $\mathcal{A}$ if there exists $\varepsilon > 0$ such that $J(y) \geq J(y_0)$ for all $y \in \mathcal{A}$ satisfying $\|y - y_0\| \leq \varepsilon$. $\triangle$

REMARK 5: We take the following approach in the examples below.

1. Let $y \in \mathcal{A}$ and find a reasonable lower bound ℓ for $J(y)$ using the definitions of J and $\mathcal{A}$.
2. If we are able to find $y_0 \in \mathcal{A}$ such that $J(y_0) = \ell$, then y_0 is a minimizer of J over $\mathcal{A}$ and $J(y_0)$ is the minimum. Otherwise, we look for a sequence $\{y_n\}$ contained in $\mathcal{A}$ for which $J(y_n) \longrightarrow \ell$. If this is can be done, then the lower bound ℓ is, in fact, an infimum for J over $\mathcal{A}$. $\diamond$

REMARK 6: We will have more sophisticated methods at our disposal later. The examples below are given to familiarize students with the notions of the admissibility class $\mathcal{A}$, a functional $J : V \to \mathbb{R}$, and the existence and values of minimums/infimums and minimizers. $\diamond$

REMARK 7: Note the infimum of a set is the greatest lower bound of that set. Consider the set $S := (0, 1]$. Here $\inf S = 0$. Note, however, that 0 is a lower bound for S and α is also a lower bound for S for each $\alpha < 0$. And, 0 is *not* a minimizer for S since $0 \notin S$. $\diamond$

3.1 – Simple Methods for Identifying Lower Bounds/Infimums:

Example 7: Define the admissibility class $\mathcal{A} := \{f \in C[a,b] \ : \ f(x) \geq 5\}$ and let $J : C[a,b] \to \mathbb{R}$ be the functional defined by

$$J(y) := \int_a^b y(x)^2 - 8y(x) + 20 \ \mathrm{d}x. \tag{23}$$

Find the minimum of $J(y)$ for $y \in \mathcal{A}$. What is the minimizer?

Solution:

Let $y \in \mathcal{A}$. Then

$$\begin{aligned} J(y) &= \int_a^b y(x)^2 - 8y(x) + 20 \ \mathrm{d}x \\ &= \int_a^b \big(y(x)^2 - 8y(x) + 16\big) + 4 \ \mathrm{d}x \\ &= \int_a^b (y(x) - 4)^2 + 4 \ \mathrm{d}x \\ &\geq \int_a^b (5-4)^2 + 4 \ \mathrm{d}x \\ &= \int_a^b 5 \ \mathrm{d}x \\ &= 5(b-a). \end{aligned} \tag{24}$$

This shows $5(b-a)$ is a lower bound for $J(y)$. To verify this is the minimum for $J(y)$, it suffices to find $f \in \mathcal{A}$ such that $J(f) = 5(b-a)$. This is accomplished if and only if the inequality in (24) is a strict equality. The only candidate is $f(x) = 5$. Since f is continuous on $[a,b]$ and $f \geq 5$, we see $f \in \mathcal{A}$. Thus we conclude $\boxed{f(x) = 5}$ is the minimizer of $J(y)$ over $\mathcal{A}$ and $\boxed{5(b-a)}$ is the minimum of $J(y)$ over $\mathcal{A}$. □

Remark 8: Note in the above example we say $f(x) = 5$ is "the" minimizer. This is because the is the only function in $\mathcal{A}$ that gives $J(f) = 5(b-a)$. In the next example, multiple minimizers exist. ⋄

Remark 9: Note $\mathcal{A}$ does not form a vector space in the following example. This follows from the fact it is not closed under scalar multiplication. For example, if $f \in \mathcal{A}$, then $-f \notin \mathcal{A}$. ⋄

Example 8: Define the admissibility class $\mathcal{A} := \{f \in C[0,1] : \ f(x) \geq x^2 - 10x + 28\}$. Then let $J : C[0,1] \to \mathbb{R}$ be the functional defined by

$$J(f) := \inf_{x\in[0,1]} f(x). \tag{25}$$

Find $\inf_{f\in A} J(f)$. Does $J(f)$ attains its infimum?

Solution:
Let $f \in \mathcal{A}$. Then, for each $x \in [0,1]$,

$$f(x) \geq x^2 - 10x + 28 = (x^2 - 10x + 25) + 3 = (x-5)^2 + 3. \tag{26}$$

Set $g(x) := (x-5)^2 + 3$. Also note $g'(x) = 2(x-5) < 0$ for $x < 5$, and so g is strictly decreasing on $[0,1]$. This implies $\inf_{x\in[0,1]} g(x) = g(1)$. Using this fast, we see

$$J(f) = \inf_{x\in[0,1]} f(x) \geq \inf_{x\in[0,1]} g(x) = g(1) = (1-5)^2 + 3 = 19. \tag{27}$$

This shows $J(f) \geq 19$, i.e., 19 is a lower bound. Moreover, because g is a polynomial, it is continuous. Whence $g \in \mathcal{A}$ and

$$J(f) \geq J(g) = 19 \quad \forall\ f \in \mathcal{A}. \tag{28}$$

Thus g is a minimizer of J over $\mathcal{A}$ and so $\boxed{\inf_{f\in\mathcal{A}} J(f) = 19.}$ Yes, $J(f)$ attains its infimum. □

Remark 10: Note in the above example we say g is "a" minimizer. In general, there may be multiple minimizers. For instance, in the above example consider defining $q(x) := g(x) + (x-1)^2$. Then $q \in C[0,1]$ and $q(x) = g(x) + (x-1)^2 \geq g(x)$, which implies $q \in \mathcal{A}$. Moreover,

$$q'(x) = g'(x) + 2(x-1) = 2(x-5) + 2(x-1) \leq 2(x-5) + 0 < 0 \quad \forall\ x \in [0,1]. \tag{29}$$

This shows q is strictly decreasing on $[0,1]$. Thus

$$J(q) = \inf_{x\in[0,1]} q(x) = q(1) = g(1) + (1-1)^2 = g(1) = 19. \tag{30}$$

This shows g and q are minimizers of J over $\mathcal{A}$. ⋄

Example 9: Define the function $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) := \begin{cases} 0 & \text{if } |x| < 1, \\ 1 & \text{if } |x| \geq 1. \end{cases} \tag{31}$$

Define the admissibility class $\mathcal{A} := \{f \in C^1(\mathbb{R}) \ : \ f(x) \geq h(x)\}$. Then let $J : C^1(\mathbb{R}) \to \mathbb{R}$ be the functional

$$J(y) := \int_{-1}^{1} y(x) \ \mathrm{d}x. \tag{32}$$

Compute $\inf\limits_{y \in \mathcal{A}} J(y)$. Does J attain its infimum?

Solution:

We proceed as follows. First we find a lower bound for J over $\mathcal{A}$. Then we show this is the greatest lower bound for J over $\mathcal{A}$. Lastly, we remark why J dose *not* attain its infimum, i.e., there is no minimizer in $\mathcal{A}$. Note, for $y \in \mathcal{A}$,

$$J(y) = \int_{-1}^{1} y(x) \ \mathrm{d}x \geq \int_{-1}^{1} h(x) \ \mathrm{d}x = \int_{-1}^{1} 0 \ \mathrm{d}x = 0. \tag{33}$$

This shows 0 is a lower bound for $J(y)$. We claim there is a sequence of functions $\{f_n\}_{n=1}^{\infty}$ contained in $\mathcal{A}$ such that $J(f_n) \longrightarrow 0$. This implies there is no lower bound greater than zero and, therefore, 0 must be the greatest lower bound for J. In other words, $0 = \inf\limits_{y \in \mathcal{A}} J(y)$.

All that remains is to verify the claimed sequence $\{f_n\}_{n=1}^{\infty}$ exists. Define $f_n(x) := x^{2n}$ for $n \geq 1$. Then $f_n(x) = x^{2n} \geq 0 = h(x)$ for $|x| < 1$ and $f_n(x) = x^{2n} \geq 1^{2n} = 1 = h(x)$ for $|x| \geq 1$. Hence $f_n \geq h$ and, with the fact f is a polynomial (and thus smooth), we see $f_n \in \mathcal{A}$ for each n. Then computing $J(f_n)$ gives

$$J(f_n) = \int_{-1}^{1} f_n(x) \ \mathrm{d}x = \int_{-1}^{1} x^{2n} \ \mathrm{d}x = 2\int_{0}^{1} x^{2n} \ \mathrm{d}x = 2\left(\frac{1^{2n+1}}{2n+1}\right) = \frac{2}{2n+1} \leq \frac{1}{n} \tag{34}$$

Taking the limit as $n \longrightarrow \infty$, we see

$$0 \leq \lim_{n \to \infty} J(f_n) \leq \lim_{n \to \infty} \frac{1}{n} = 0. \tag{35}$$

Thus $\lim\limits_{n \to \infty} J(f_n) = 0$, as desired.

Lastly, we note J does not attain its infimum. This is because the infimum is obtained if and only if $y(x) = 0$ for $|x| < 1$. But, because we need $y(\pm 1) \geq 1$, such a minimizer would necessarily have a jump discontinuity, contradicting the fact $y(x)$ must be continuous to be in $\mathcal{A}$. $\square$

Remark 11: After reading the above example, we may ask ourselves "But why did you pick $f_n(x) = x^{2n}$? How did you know to do that?". I encourage the reader to draw a picture. A good picture can go a long way.

We want a continuous function f with $f(-1) \geq 1$ and $f(1) \geq 1$, but approaches 0 for $|x| < 1$. To keep things simple, we may restrict our consideration to even functions. Perhaps an initial guess might be to use x^2 to get an even function with $(-1)^2 = 1 = 1^2$. Then because $|x| < 1$, we know $|x|^n \longrightarrow 0$ as $n \longrightarrow \infty$ (see Lemma below). So, we could try $(x^2)^n = x^{2n}$. Indeed, we see graphically below this does do the trick.

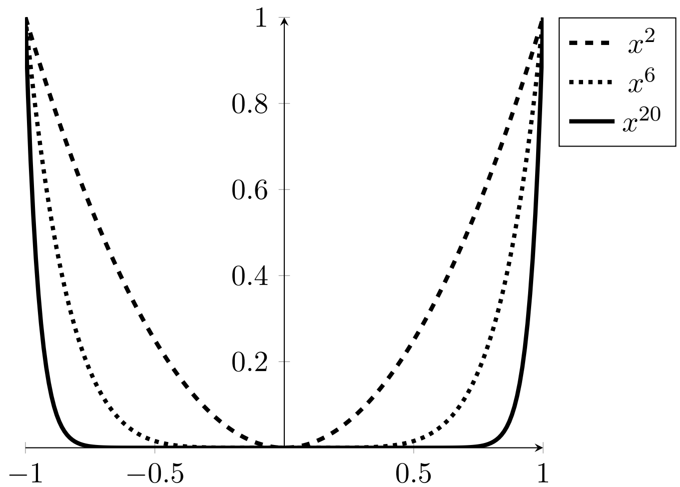

Figure 1: Plots of x^{2n} on $[-1, 1]$ for $n = 1, 3, 10$. ◇

Lemma: Let $c \in (0, 1)$. The $\lim_{n\to\infty} c^n = 0$. △

Proof:

Let $n \in \mathbb{N}$. Then $c^{n+1} = cc^n < 1c^n = c^n$. This shows the sequence $\{c^n\}_{n=1}^{\infty}$ is decreasing. And, the fact $c^n \geq 0^n = 0$ shows it is bounded from below. The Monotone Convergence Theorem then asserts $\{c^n\}_{n=1}^{\infty}$ converges to some limit $\alpha \in \mathbb{R}$. Observe

$$\alpha = \lim_{n\to\infty} c^n = \lim_{n\to\infty} c^{n+1} = c \lim_{n\to\infty} c^n = c\alpha. \tag{36}$$

Because $c \in (0, 1)$, the above can hold if and only if $\alpha = 0$. Thus $\lim_{n\to\infty} c^n = 0$. ■

3.2 – The Gâteaux Derivative and Its Applications:

Definition: Suppose V is a vector space. The **Gâteaux derivative**, denoted $DJ(y)h$, of $J : V \to \mathbb{R}$ at $y \in V$ in the direction of h is defined as a mapping $DJ(y)h$ from V into $\mathbb{R}$ such that

$$DJ(y)h := \lim_{\varepsilon \to 0} \frac{J(y+\varepsilon h) - J(y)}{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \left[J(y+\varepsilon h) \right]_{\varepsilon=0}, \tag{37}$$

provided the limit exists. If the limit exists for all $h \in V$, then we say J is **Gâteaux differentiable** at y. △

Definition: Note another common notation is to write $\delta J(y,h) := DJ(y)h$. These notations are interchangeable and $\delta J(y,h)$ is used in our text. △

Remark 12: **Computing the Gâteaux derivative:**

1. Identify J and $\mathcal{A}$.
2. Fix $y \in \mathcal{A}$. Let $h \in V$ such that $y + \varepsilon h \in \mathcal{A}$ for all ε with $|\varepsilon|$ sufficiently small.[1]
3. Compute $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y+\varepsilon h)\right]$ and then evaluate the result at $\varepsilon = 0$.

◇

Theorem: Suppose y_0 is a local minimizer of a functional $J : V \to \mathbb{R}$ over an open set contained in $\mathcal{A}$. Then $DJ(y_0)h = 0$ for every admissible variation h. △

Remark 13: We can state the result of the above theorem more intuitively, using knowledge from calculus. Fix h to be any admissible variation. Then define the function $f : \mathbb{R} \to \mathbb{R}$ by $f(\varepsilon) := J(y_0 + \varepsilon h)$. Since y_0 is a minimizer for J, it follows that 0 is a local minimizer of f. Therefore $f'(0) = 0$. In other words,

$$0 = f'(0) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y_0+\varepsilon h)\right]_{\varepsilon=0} = \lim_{\varepsilon\to 0} \frac{J(y_0+\varepsilon h) - J(y_0)}{\varepsilon} = DJ(y_0)h. \tag{38}$$

◇

[1] Note the choice of $\mathcal{A}$ may impose restrictions on h.

Lemma: Suppose $f \in C[a,b]$. If for all $h \in C[a,b]$ we have

$$0 = \int_a^b f(x)h(x)\ \mathrm{d}x, \tag{39}$$

then $f(x) = 0$ for all $x \in [a,b]$, i.e., f is identically zero. △

Remark 14: The above lemma will be of significant importance to use when trying to determine what differential equation a minimizer y of a functional J must satisfy. ⋄

Definition: We say that y is an **extremal** of J provided $\delta J(y,h) = 0$ for arbitrary h. △

Remark 15: From the above definition and our theorem giving a necessary condition for minimizers, we know every minimizer is an extremal. However, an extremal need not necessarily be a minimizer (e.g., it could be a maximizer as is the case in Example 13). ⋄

Remark 16: Below we list steps for finding extremals of $J : V \to \mathbb{R}$ over $\mathcal{A}$, assuming $\mathcal{A}$ is nonempty.

Finding Extremals:

1. Let $y \in \mathcal{A}$ and find conditions on $h \in V$ such that $y + \varepsilon h \in \mathcal{A}$ when $|\varepsilon|$ is sufficiently small.
2. Compute $DJ(y)h = \dfrac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y+\varepsilon h)\right]_{\varepsilon=0}$ and simplify using information known from $\mathcal{A}$.
3. Set $DJ(y) = 0$ and use our lemma to obtain an equation for y.
4. Use this equation to solve for each possible extremal $y_0 \in \mathcal{A}$.

If, in addition, we would like to verify an extremal y_0 is a *local minimizer* of J over $\mathcal{A}$, then we must show there is $\varepsilon > 0$ such that $J(y_0) \leq J(y)$ for all $y \in \mathcal{A}$ satisfying $\|y - y_0\| < \varepsilon$ for some appropriate norm $\|\cdot\|$. If $J(y_0) \leq J(y)$ for all $y \in \mathcal{A}$, then it is a *global minimizer*. Note every global minimizer is also a local minimizer. ⋄

Example 10: Define the admissibility class $\mathcal{A} := C[a,b]$. Then let $J : \mathcal{A} \to \mathbb{R}$ be the functional

$$J(y) = \int_a^b (y(x) - 4)^2 \, dx. \tag{40}$$

Find a minimizer of J over $\mathcal{A}$ using the Gâteaux derivative. You may suppose a minimizer $y_0 \in \mathcal{A}$ exists.

Solution:
Fix $y, h \in \mathcal{A}$. Then $y + \varepsilon h \in \mathcal{A}$ for all $\varepsilon \in \mathbb{R}$ since $\mathcal{A}$ is a vector space. We expand $J(y+\varepsilon h)$ to find

$$\begin{aligned} J(y+\varepsilon h) &= \int_a^b [y(x) + \varepsilon h(x) - 4]^2 \, dx \\ &= \int_a^b [y(x) + \varepsilon h(x)]^2 - 8\,[y(x) + \varepsilon h(x)] + 16 \, dx \\ &= \int_a^b y(x)^2 + 2\varepsilon y(x)h(x) + \varepsilon^2 h(x)^2 - 8\,[y(x) + \varepsilon h(x)] + 16 \, dx \\ &= \int_a^b y(x)^2 - 8y(x) + 16 \, dx + 2\varepsilon \int_a^b [y(x) - 4]\, h(x) \, dx + \varepsilon^2 \int_a^b h(x)^2 \, dx \\ &= J(y) + +2\varepsilon \int_a^b [y(x) - 4]\, h(x) \, dx + \varepsilon^2 \int_a^b h(x)^2 \, dx. \end{aligned} \tag{41}$$

Thus

$$\begin{aligned} DJ(y)h &= \lim_{\varepsilon \to 0} \frac{J(y+\varepsilon h) - J(y)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \left(2 \int_a^b [y(x) - 4]\, h(x) \, dx + \varepsilon \int_a^b h(x)^2 \, dx \right) \\ &= 2 \int_a^b [y(x) - 4]\, h(x) \, dx. \end{aligned} \tag{42}$$

By our theorem, we know that if y_0 is a minimizer of J over $\mathcal{A}$, then

$$0 = DJ(y_0)h = 2 \int_a^b [y_0(x) - 4]\, h(x) \, dx \tag{43}$$

for all admissible variations h. This implies the only candidate minimizer is $y_0(x) = 4$. Since $J(y) \geq 0$ for all $y \in \mathcal{A}$ and $J(y_0) = 0$, we conclude $\boxed{y_0(x) = 4}$ is the global minimizer of J over $\mathcal{A}$.

□

Remark 17: The above computations may seem quite tedious. This is because they are. It is important to note, if $q(\varepsilon) := J(y_0 + \varepsilon h)$, then the statement $q'(0) = 0$ is equivalent to saying $DJ(y_0)h = 0$ since

$$q'(0) = \lim_{\varepsilon \to 0} \frac{q(\varepsilon) - q(0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{J(y_0 + \varepsilon h) - J(y_0)}{\varepsilon} = DJ(y_0)h. \tag{44}$$

So, a more elegant approach for computing $DJ(y)h$ is given in the following reworking of the above example, utilizing our knowledge of derivatives for real valued functions. $\diamond$

Example 11: Repeat the previous example, making use of derivatives of real-valued functions.

Solution:

Fix $y, h \in \mathcal{A}$. Then $y + \varepsilon h \in \mathcal{A}$ for all $\varepsilon \in \mathbb{R}$ since $\mathcal{A}$ is a vector space. Through direct computation we find

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} J(y + \varepsilon h) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int_a^b \left[y(x) + \varepsilon h(x) - 4\right]^2 \ \mathrm{d}x = \int_a^b 2\left[y(x) + \varepsilon h(x) - 4\right] h(x) \ \mathrm{d}x. \tag{45}$$

Thus

$$DJ(y)h = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y + \varepsilon h)\right]_{\varepsilon = 0} = \int_a^b 2\left[y(x) - 4\right] h(x) \ \mathrm{d}x. \tag{46}$$

If y_0 is a minimizer of J, then it is an extremal of J, and so $DJ(y_0)h = 0$. Since this result holds for arbitrary admissible variations h, our lemma states the only candidate minimizer is $y_0(x) = 4$. Since $J(y) \geq 0$ for all $y \in \mathcal{A}$ and $J(y_0) = 0$, we conclude $\boxed{y_0(x) = 4}$ is the global minimizer of J over $\mathcal{A}$. $\square$

Example 12: Define the functional $J : C^2[0,1] \to \mathbb{R}$ by

$$J(y) := \int_0^1 \frac{1}{2}m\dot{y}^2 - mgy \ \mathrm{d}t, \tag{47}$$

here using the dot notation for time t derivatives of y. Let $\mathcal{A} := \{y \in C^2[0,1] \ : y(0) = \alpha_1, \ y(1) = \alpha_2\}$. Find an extremal $y_0 \in \mathcal{A}$ of J.

Solution:

Pick $y \in \mathcal{A}$. Since the end points of $y \in \mathcal{A}$ are fixed, if $h \in V$ and $y_0+\varepsilon h \in \mathcal{A}$ for any nonzero ε, then

$$\alpha_1 = y(0) + \varepsilon h(0) = \alpha_1 + \varepsilon h(0) \quad \Rightarrow \quad h(0) = 0. \tag{48}$$

Similarly, $h(1) = 0$. Fixing h, we compute

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y+\varepsilon h)\right] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_0^1 \frac{1}{2}m(\dot{y}+\varepsilon\dot{h})^2 - mg(\dot{y}+\varepsilon\dot{h}) \ \mathrm{d}t = \int_0^1 m\left(\dot{y}+\varepsilon\dot{h}\right)h' - mgh \ \mathrm{d}t. \tag{49}$$

To make this more useful, we use integration by parts with the first to rewrite the above in a more useful form. That is,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y+\varepsilon h)\right] = \int_0^1 -m(\ddot{y}+\varepsilon\ddot{h})h - mgh \ \mathrm{d}t = -m\int_0^1 \left(\ddot{y}+\varepsilon\ddot{h}+g\right)h \ \mathrm{d}t. \tag{50}$$

Evaluating the above at $\varepsilon = 0$, we deduce

$$\delta J(y,h) = -m\int_0^1 (\ddot{y}+g)h \ \mathrm{d}t. \tag{51}$$

Thus for a minimizer y_0 of J we see $\delta J(y,h) = 0$. Since this holds for an arbitrary admissible variation h, by our lemma we deduce the only candidate minimizer satisfies $\ddot{y}_0 = -g$. Then

$$\dot{y}_0 = -gt + c_1 \quad \Rightarrow \quad y_0 = -\frac{1}{2}gt^2 + c_1 t + c_2 \tag{52}$$

for some $c_1, c_2 \in \mathbb{R}$. Using the fact $y_0(0) = \alpha_1$, we know $c_2 = \alpha_1$. Then

$$\alpha_2 = y_0(1) = -\frac{1}{2}g + c_1 + \alpha_1 \quad \Rightarrow \quad c_1 = \alpha_2 - \alpha_1 + \frac{g}{2}, \tag{53}$$

and we conclude

$$\boxed{y_0(t) = -\frac{1}{2}gt^2 + \left(\alpha_2 - \alpha_1 + \frac{g}{2}\right)t + \alpha_1.} \tag{54}$$

□

Remark 18: From classical mechanics in physics, we know the kinetic energy of a ball of mass m is given by $T = \frac{1}{2}m\dot{y}^2$ and its potential energy is $U = mgh$. In the above, we are minimizing $T - U$ over a time interval. It turns out this is associated with *Hamilton's Principle* and it gives the same result as would be obtained using Newton's second law of motion. $\diamond$

Example 13: Let $V := C[0,1]$ and $\mathcal{A} := V$. Define $J : V \to \mathbb{R}$ by

$$J(y) = \int_0^1 -y(x)^2 + 6y(x) + 10 \text{ d}x. \tag{55}$$

Find $y_0 \in \mathcal{A}$ such that $DJ(y_0)h = 0$. Is y_0 a minimizer of J over $\mathcal{A}$?

Solution:

Let $y, h \in \mathcal{A}$ and $\varepsilon \in \mathbb{R}$. Differentiating, we see

$$\frac{\text{d}}{\text{d}\varepsilon}J(y+\varepsilon h) = \frac{\text{d}}{\text{d}\varepsilon}\int_0^1 -\left[y+\varepsilon h\right]^2 + 6[y+\varepsilon h] + 10 \text{ d}x = \int_0^1 -2\left[y+\varepsilon h\right]h + 6h \text{ d}x. \tag{56}$$

This implies

$$DJ(y)h = \int_0^1 -2yh + 6h \text{ d}x = \int_0^1 (-2y+6)\,h \text{ d}x. \tag{57}$$

So, taking $y_0(x) = 3$, we obtain $DJ(y_0)h = 0$. Now observe

$$J(y) = \int_0^1 -\left(y^2 - 6y + 9\right) + 19 \text{ d}x = \int_0^1 -(y-3)^2 + 19 \text{ d}x \le \int_0^1 19 \text{ d}x = 19. \tag{58}$$

This shows 19 is an upper bound for $J(y)$. However,

$$J(y_0) = \int_0^1 -(3-3)^2 + 19 \text{ d}x = 19. \tag{59}$$

Moreover, if $y \neq 3 = y_0$, then $J(y) < J(y_0)$. Thus y_0 is **not** a minimizer of J. In fact, this shows y_0 is a maximizer of J. $\square$

Remark 19: The above example shows the condition $DJ(y_0)h = 0$ is **not** a sufficient condition to conclude y_0 is a minimizer of J. $\diamond$

3.3 – An Introduction to the Euler Lagrange Equations:

Suppose we have a function $L(x, y, y')$ and

$$J(y) = \int_a^b L(x, y, y')\ \mathrm{d}x. \tag{60}$$

Herein we derive a general formulation for the equation satisfied by any minimizer y_0 of J over the admissibility class

$$\mathcal{A} := \{y \in C^2[a, b]\ :\ y(a) = c_0,\ y(b) = c_1\}. \tag{61}$$

Let us first compute the Gâteaux derivative $DJ(y)h$. Let $y \in \mathcal{A}$ and suppose $y + \varepsilon h \in \mathcal{A}$ for some $\varepsilon \in \mathbb{R}$ and $h \in C^2[a, b]$. Then the restrictions of $\mathcal{A}$ reveal

$$c_0 = y(a) + \varepsilon h(a) = c_0 + \varepsilon h(a) \quad \Rightarrow \quad \varepsilon h(a) = 0 \quad \Rightarrow \quad h(a) = 0. \tag{62}$$

Similarly, $h(b) = 0$. We will use this fact later. Now, computing we see

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y + \varepsilon h)\right] &= \int_a^b \frac{\partial}{\partial \varepsilon} L(x, y + \varepsilon h, y' + \varepsilon h')\ \mathrm{d}x \\ &= \int_a^b L_y(x, y + \varepsilon h, y' + \varepsilon h')h(x) + L_{y'}(x, y + \varepsilon h, y' + \varepsilon h')h'(x)\ \mathrm{d}x. \end{aligned} \tag{63}$$

The second line follows from the third line by using the chain rule. In more familiar calculus notation, we may write

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} L(x, y + \varepsilon h, y' + \varepsilon h') &= \frac{\partial L}{\partial x}\frac{\partial}{\partial \varepsilon}(x) + \frac{\partial L}{\partial y}\frac{\partial}{\partial \varepsilon}(y + \varepsilon h) + \frac{\partial L}{\partial y'}\left(y' + \varepsilon h'\right) \\ &= \frac{\partial L}{\partial x} 0 + \frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y'} h' \\ &= \frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y'} h' \\ &= L_y h + L_{y'} h'. \end{aligned} \tag{64}$$

With this partial differentiation review complete, let us return to (63). This shows

$$DJ(y)h = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[J(y + \varepsilon h)\right]_{\varepsilon = 0} = \int_a^b L_y(x, y, y')h + L_{y'}(x, y, y')h'\ \mathrm{d}x. \tag{65}$$

Although this equation is true, it isn't really useful for finding a minimizer y. What we would really like is to have an integrand where we have one big expression all multiplied by h. This can be accomplished by using integration by parts to move the derivative off of the h' term and onto the $L_{y'}$ term. We show this as follows:

$$DJ(y)h = \int_a^b L_y h - \left(\frac{\mathrm{d}}{\mathrm{d}x} L_{y'}\right) h \ \mathrm{d}x + \left[L_{y'} h\right]_{x=a}^{b} = \int_a^b L_y h - \left(\frac{\mathrm{d}}{\mathrm{d}x} L_{y'}\right) h \ \mathrm{d}x. \tag{66}$$

Note the boundary term canceled. This is because $h(a) = h(b) = 0$, and so $L_{y'}h$ evaluated at $x = a$ and $x = b$ is also zero. We now see

$$DJ(y)h = \int_a^b \left(L_y - \frac{\mathrm{d}}{\mathrm{d}x} L_{y'}\right) h \ \mathrm{d}x. \tag{67}$$

But for a minimizer y_0 of J we have $DJ(y_0)h = 0$. Since h was arbitrary in $C^2[a, b]$ with $h(a) = h(b) = 0$, we conclude the necessary condition for a function y_0 to be a minimizer of J over $\mathcal{A}$ is

$$L_y(x, y_0, y_0') - \frac{\mathrm{d}}{\mathrm{d}x} L_{y'}(x, y_0, y_0') = 0 \ \text{ for } x \in [a, b]. \tag{68}$$

This above equation will be of great practical importance for us. It is well-known as the **Euler-Lagrange equation.** For it simplifies and summarizes all of the previous work we have done this course into one simple formula. We illustrate how to make use of this with the following examples.

Remark 20: Note that when we take the derivative of $L_{y'}$ with respect to x we need to take into account the fact y and y' are functions of x. ◇

Remark 21: This next example does in two lines of equations what we previously this quarter would have done with over a page of computations. ◇

Example 14: Define $J : C^2[0,1] \to \mathbb{R}$ by

$$J(y) := \frac{1}{2}\int_0^1 (y')^2 + y^2 + 2ye^x \ \mathrm{d}x. \tag{69}$$

Find the form of any minimizer of J.

Solution:

Here $L = y^2 + (y')^2 + 2ye^x$. Through direct computation we see

$$L_y = 2y + 0 + 2e^x = 2(y + e^x), \quad L_{y'} = 0 + 2y' + 0 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x} L_{y'} = 2y''. \tag{70}$$

Thus the Euler-Lagrange equations show any minimizer satisfies

$$0 = L_y - \frac{\mathrm{d}}{\mathrm{d}x} L_{y'} = 2(y + e^x) - 2y'' \quad \Rightarrow \quad y'' - y = -e^x. \tag{71}$$

From the theory of ordinary differential equations (ODEs), we know this is a linear ODE with a solution of the form $y = y_p + y_h$ where y_h satisfies $y_h'' - y_h = 0$. This implies $y_h = c_1 e^x + c_2 e^{-x}$ for some scalars c_1 and c_2. And, it may be checked that $y_p = -(-x/2)e^x$. Thus extremals of J are of the form

$$\boxed{y(x) = y_p(x) + y_h(x) = \left(c_1 - \frac{x}{2}\right) e^x + c_2 e^{-x}.} \tag{72}$$

□

Remark 22: If we were given $y(0)$ and $y(1)$ then we could explicitly solve for c_1 and c_2. ◇

Remark 23: Again, this example reveals how much simpler it is to make use of the Euler-Lagrange equation. For this problem would have been extremely tedious to complete if we were to expand $J(y + \varepsilon h)$ in powers of ε, and then take limits. Looking at the bigger picture, this provides an example that it may be extremely advantage to first look at the problem abstractly and find the general form of a solution, and then use this general form to tackle specific problems. ◇

Example 15: Find the form of extremals for

$$J(y) = \int_a^b x^2(y')^2 + y^2 \, \mathrm{d}x. \tag{73}$$

Solution:
Here $L = x^2(y')^2 + y^2$. Through direct computation we see

$$L_y = 2y, \quad L_{y'} = 2x^2y' \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}L_{y'} = 4xy' + 2x^2y''. \tag{74}$$

Thus the Euler-Lagrange equation for this problem is

$$0 = L_y - \frac{\mathrm{d}}{\mathrm{d}x}L_{y'} = 2y - (4xy' + 2x^2y'') \quad \Rightarrow \quad -2x^2y'' - 4xy' + 2y = 0. \tag{75}$$

This is a second order Cauchy-Euler equation, which has a solution of the form

$$\boxed{y(x) = c_1x^{(\sqrt{5}-1)/2} + c_2x^{-(\sqrt{5}+1)/2}} \tag{76}$$

for some scalars $c_1, c_2 \in \mathbb{R}$. So functions of the form (76) give extremals for $J(y)$. □

3.4 – Physics Equations of Motion Examples:

Example 16: Suppose a ball is oscillating on top of a massless spring. The kinetic energy of the ball is $T = \frac{1}{2}m\dot{y}^2$ where here dots are used to represent time derivatives and m is the mass of the ball. The potential energy here is $U = mgy + \frac{1}{2}ky^2$, which is the sum of the potential energy from the ball being at a height y and the potential energy from the spring where $k > 0$ is the spring constant. (Here we take $y = 0$ when the spring is in its relaxed state.) Hamilton's Principle asserts the motion y of the ball for $t \in [0, t^*]$ will minimize

$$\int_0^{t^*} T - U \ \mathrm{d}t. \tag{77}$$

Use Hamilton's Principle to find an ODE describing the motion for this ball.

Solution:
Here the Lagrangian is

$$L = T - U = \frac{1}{2}m\dot{y}^2 - \left(mgy + \frac{1}{2}ky^2\right). \tag{78}$$

Then differentiating gives

$$L_y = -(ky + mg), \quad L_{y'} = m\dot{y}, \quad \frac{\mathrm{d}}{\mathrm{d}x}L_{y'} = m\ddot{y}. \tag{79}$$

Applying Hamilton's Principle, we know the ball will satisfy the Euler-Lagrange equation, i.e.,

$$0 = L_y - \frac{\mathrm{d}}{\mathrm{d}t}L_{\dot{y}} = -(ky + mg) - m\ddot{y} \quad \Rightarrow \quad \boxed{m\ddot{y} = -ky - mg.} \tag{80}$$

□

Remark 24: Consider again the above problem but instead using Newton's laws of motion. The forces acting on the ball will be a force from the spring ($F_s = -ky$) and a force from gravity ($F_g = -mg$) with the negative sign due to the fact gravity pushes the ball downward. Then Newton's second law tells us

$$m\ddot{y} = \sum_i F_i = F_s + F_g = -ky - mg, \tag{81}$$

which is precisely the result we obtained in the above example! In simple situations like this, it may be much easier to just write Newton's laws. But, it is quite easy to think of examples where writing Newton's laws out would be extremely cumbersome. However, in such cases the above approach with the Euler-Lagrange equations still works well. ◇

Example 17: A pendulum consists of a mass m suspended by a massless spring with unextended length b and spring constant k. Find Lagrange's equations of motion.

Solution:

For this problem we consider the following illustration:

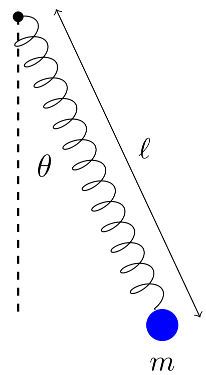

Figure 2: Simple Pendulum with Spring

We define the potential energy U such that $U = 0$ when $\theta = \pi/2$ and the spring is relaxed, i.e.,

$$U = mg(-\ell\cos\theta) + \frac{k}{2}(\ell - b)^2. \tag{82}$$

Using polar coordinates, the kinetic energy T is given by $T = \frac{1}{2}mv^2 = \frac{1}{2}m(\dot{\ell}^2 + (\ell\dot{\theta})^2)$. Thus the Lagrangian $\mathcal{L}$ for this system is given by

$$\mathcal{L} = T - U = \frac{1}{2}m(\dot{\ell}^2 + (\ell\dot{\theta})^2) + mg\ell\cos\theta - \frac{1}{2}k(\ell - b)^2$$

Then the first of Lagrange's equations of motion is

$$0 = \frac{\partial\mathcal{L}}{\partial\ell} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\ell}} = \left[m\ell\dot{\theta}^2 + mg\cos\theta - k(\ell - b)\right] - \frac{\mathrm{d}}{\mathrm{d}t}\left[m\dot{\ell}\right] = \left[m\ell\dot{\theta}^2 + mg\cos\theta - k(\ell - b)\right] - m\ddot{\ell}. \tag{83}$$

Dividing by m and rearranging yields

$$\boxed{\ddot{\ell} - \ell\dot{\theta}^2 - g\cos\theta + \frac{k}{m}(\ell - b) = 0.} \tag{84}$$

Similarly, observe

$$0 = \frac{\partial\mathcal{L}}{\partial\theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = -mg\ell\sin\theta - \frac{\mathrm{d}}{\mathrm{d}t}\left[m\ell\dot{\theta}\right] = -mg\ell\sin\theta - m\left[2\ell\dot{\ell}\dot{\theta} + \ell^2\ddot{\theta}\right]. \tag{85}$$

Dividing by m and rearranging gives the second equation

$$\boxed{\ell^2\ddot{\theta} + 2\ell\dot{\ell}\dot{\theta} + g\ell\sin\theta = 0.} \tag{86}$$

□

Example 18: A simple pendulum of length ℓ and bob with mass m is attaches to a massless support moving vertically upward with constant acceleration a. Determine a) the equation of motion and b) the period for small oscillations.

Solution:

Consider the following illustration for this problem.

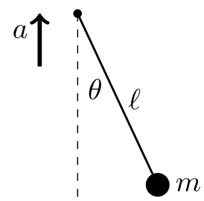

Figure 3: Accelerating Simple Pendulum

a) We define the potential energy U to be zero where we would obtain $\theta = t = 0$. Since the pendulum accelerates upward with velocity a, we write

$$U = mg\left(\ell(1-\cos\theta) + \iint a \ \mathrm{d}t\mathrm{d}t\right) = mg\left(\ell(1-\cos\theta) + \frac{1}{2}at^2\right). \tag{87}$$

Take the origin to be where $U = 0$ at $t = 0$. Then

$$(x, y) = \left(\ell\sin\theta, \ell\,(1-\cos\theta) + \frac{1}{2}at^2\right) \quad \Rightarrow \quad (\dot{x}, \dot{y}) = \left(\ell\dot{\theta}\cos\theta, \ell\dot{\theta}\sin\theta + at\right). \tag{88}$$

The kinetic energy T is given by

$$T = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\left[\left(\ell\dot{\theta}\cos\theta\right)^2 + \left(\ell\dot{\theta}\sin\theta + at\right)^2\right] = \frac{1}{2}m\left[\ell^2\dot{\theta}^2 + 2a\ell\dot{\theta}t\sin\theta + (at)^2\right]. \tag{89}$$

Thus, the Lagrangian for the system is

$$\mathcal{L} = T - U = m\left[\frac{1}{2}\left(\ell^2\dot{\theta}^2 + 2a\ell\dot{\theta}t\sin\theta + (at)^2\right) - g\left(\ell[1-\cos\theta] + \frac{1}{2}at^2\right)\right] \tag{90}$$

Since m is directly proportional to $\mathcal{L}$ we may, without loss of generality, choose the units of mass such that m is unity. From the Euler-Lagrange equation we obtain

$$\begin{aligned} 0 = \frac{\partial\mathcal{L}}{\partial\theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} &= \left[a\ell\dot{\theta}t\cos\theta - g\ell\sin\theta\right] - \frac{\mathrm{d}}{\mathrm{d}t}\left[\ell^2\dot{\theta} + a\ell t\sin\theta\right] \\ &= \left[a\ell\dot{\theta}t\cos\theta - g\ell\sin\theta\right] - \left[\ell^2\ddot{\theta} + a\ell\sin\theta + a\ell t\dot{\theta}\cos\theta\right] \\ &= -\ell\,[a+g]\sin\theta - \ell^2\ddot{\theta}. \end{aligned} \tag{91}$$

Returning to (88), we see the position of the mass is a function of θ and t (n.b. ℓ is constant in time). This means it suffices to find an equation of motion for θ. From (91), we see this is

$$\boxed{\ddot{\theta} = -\frac{a+g}{\ell}\sin\theta.}$$

b) For small angles we have that $\sin\theta \approx \theta$. Using a) this approximation yields the differential relation

$$\ddot{\theta} = -\frac{a+g}{\ell}\theta$$

This is precisely the equation for a harmonic oscillator with angular frequency $\omega^2 = \sqrt{(a+g)/b}$. The period t^* of oscillation is then

$$t^* = \frac{2\pi}{\omega} = \boxed{2\pi\sqrt{\frac{b}{a+b}}.}$$

$\square$

Example 19: A double pendulum consists of two simple pendula, with one pendulum suspended from the bob of the other. If the two pendula have equal lengths and have bobs of equal mass and if both pendula are confined to move in the same plane, find Lagrange's equations of motion for the system. Do not assume small angles.

Solution:

For this problem we consider the following illustration of the double pendulum. The point masses

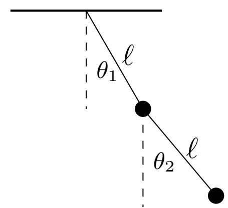

Figure 4: Double Pendulum Illustration

are located at $(x_1, y_1) = (\ell \sin\theta_1,\ -\ell\cos\theta_1)$ and $(x_2, y_2) = (\ell[\sin\theta_1 + \sin\theta_2],\ -\ell[\cos\theta_1 + \cos\theta_2])$. The kinetic energy is a scalar multiple of $v_1^2 + v_2^2$. So, observe

$$\begin{aligned}
v_1^2 + v_2^2 &= \dot{x}_1^2 + \dot{y}_1^2 + \dot{x}_2^2 + \dot{y}_2^2 \\
&= \left[\ell\cos\theta_1\dot{\theta}_1\right]^2 + \left[\ell\sin\theta_1\dot{\theta}_1\right]^2 + \left[\ell(\cos\theta_1\cdot\dot{\theta}_1 + \cos\theta_2\cdot\dot{\theta}_2)\right]^2 + \left[\ell(\sin\theta_1\dot{\theta}_1 + \sin\theta_2\dot{\theta}_2)\right]^2 \\
&= \ell^2\left[\dot{\theta}_1^2 + \dot{\theta}_1^2(\cos^2\theta_1 + \sin^2\theta_1) + \dot{\theta}_2^2(\cos^2\theta_2 + \sin^2\theta_2) + 2\cos\theta_1\cos\theta_2\cdot\dot{\theta}_1\dot{\theta}_2 + 2\sin\theta_1\sin\theta_2\cdot\dot{\theta}_1\dot{\theta}_2\right] \\
&= \ell^2\left[2\dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1\dot{\theta}_2\cos(\theta_2 - \theta_1)\right].
\end{aligned}$$

Letting the potential energy be zero at the top attachment of the double pendulum we have

$$U = mg(y_1 + y_2) = -mg\left(\ell\cos\theta_1 + [\ell\cos\theta_1 + \ell\cos\theta_2]\right) = -mg\ell\left(2\cos\theta_1 + \cos\theta_2\right).$$

Combining these results we obtain the Lagrangian

$$\begin{aligned}
\mathcal{L} = T - U &= \frac{1}{2}m(v_1^2 + v_2^2) + mg\ell\left(2\cos\theta_1 + \cos\theta_2\right) \\
&= \frac{1}{2}m\ell^2\left[2\dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1\dot{\theta}_2\cos(\theta_2 - \theta_1)\right] + mg\ell\left(2\cos\theta_1 + \cos\theta_2\right).
\end{aligned}$$

We make the following computations:

$$\begin{aligned}
\frac{\partial L}{\partial \theta_1} &= m\ell^2\dot\theta_1\dot\theta_2\sin(\theta_2-\theta_1) - 2mg\ell\sin\theta_1,\\
\frac{\partial L}{\partial \theta_2} &= -m\ell^2\dot\theta_1\dot\theta_2\sin(\theta_2-\theta_1) - mg\ell\sin\theta_2,\\
\frac{\partial L}{\partial \dot\theta_1} &= m\ell^2\left(2\dot\theta_1 + \dot\theta_2\cos(\theta_2-\theta_1)\right),\\
\frac{\partial L}{\partial \dot\theta_2} &= m\ell^2\left(\dot\theta_2 + \dot\theta_1\cos(\theta_2-\theta_1)\right),\\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot\theta_1} &= m\ell^2\left(2\ddot\theta_1 + \ddot\theta_2\cos(\theta_2-\theta_1) + \dot\theta_2[\dot\theta_1-\dot\theta_2]\sin(\theta_2-\theta_1)\right),\\
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot\theta_2} &= m\ell^2\left(\ddot\theta_2 + \ddot\theta_1\cos(\theta_2-\theta_1) + \dot\theta_1[\dot\theta_1-\dot\theta_2]\sin(\theta_2-\theta_1)\right).
\end{aligned}$$

Finally, we obtain Lagrange's equations of motion. Namely, we use the fact that

$$\begin{aligned}
0 &= \frac{\partial L}{\partial \theta_1} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot\theta_1}\\
&= \left[m\ell^2\dot\theta_1\dot\theta_2\sin(\theta_2-\theta_1) - 2m\ell^2(g/\ell)\sin\theta_1\right] - m\ell^2\left(2\ddot\theta_1 + \ddot\theta_2\cos(\theta_2-\theta_1) + \dot\theta_2[\dot\theta_1-\dot\theta_2]\sin(\theta_2-\theta_1)\right) \quad (92)\\
&= \dot\theta_1\dot\theta_2\sin(\theta_2-\theta_1) - \frac{2g}{\ell}\sin\theta_1 - 2\ddot\theta_1 - \ddot\theta_2\cos(\theta_2-\theta_1) + \dot\theta_2[\dot\theta_1-\dot\theta_2]\sin(\theta_2-\theta_1).
\end{aligned}$$

This implies

$$\boxed{\frac{2g}{\ell}\sin\theta_1 + 2\ddot\theta_1 + \ddot\theta_2\cos(\theta_2-\theta_1) + \dot\theta_2^2\sin(\theta_2-\theta_1) = 0.} \quad (93)$$

Similarly, we see

$$\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial \theta_2} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot\theta_2}\\
&= -m\ell^2\dot\theta_1\dot\theta_2\sin(\theta_2-\theta_1) - m\ell^2(g/\ell)\sin\theta_2 - m\ell^2\left(\ddot\theta_2 + \ddot\theta_1\cos(\theta_2-\theta_1) + \dot\theta_1[\dot\theta_1-\dot\theta_2]\sin(\theta_2-\theta_1)\right).
\end{aligned} \quad (94)$$

After simplifying, we obtain the second equation of motion

$$\boxed{\frac{g}{\ell}\sin\theta_2 + \ddot\theta_2 + \ddot\theta_1\cos(\theta_2-\theta_1) + \dot\theta_1^2\sin(\theta_2-\theta_1) = 0.} \quad (95)$$

□

Example 20: The potential for an anharmonic oscillator is $U = kx^2/2 + bx^4/4$ where k and b are constants. Find Hamilton's equations of motion.

Solution:

The Lagrangian $\mathcal{L}$ for this system is given by

$$\mathcal{L} = T - U = \frac{1}{2}m\dot{x}^2 - k\frac{x^2}{2} - b\frac{x^4}{4}.$$

This implies the momentum is given by

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$$

so that the Hamiltonian H is

$$H = p_x\dot{x} - \mathcal{L} = m\dot{x}^2 - \left[\frac{1}{2}m\dot{x}^2 - k\frac{x^2}{2} - b\frac{x^4}{4}\right] = \frac{1}{2}m\dot{x}^2 + k\frac{x^2}{2} + b\frac{x^4}{4} = \frac{p_x^2}{2m} + k\frac{x^2}{2} + b\frac{x^4}{4}.$$

Consequently, Hamilton's equations of motion are

$$\boxed{\dot{x} = \frac{\partial H}{\partial p_x} = \frac{p_x}{m}} \quad \text{and} \quad \boxed{\dot{p}_x = -\frac{\partial H}{\partial x} = -kx - bx^3.}$$

□

Example 21: Find a Lagrangian that yields the Euler-Lagrange equation $\ddot{y}+ay+b=0$ for fixed $a,b>0$.

Solution:
For a Lagrangian $\mathcal{L}$, the Euler-Lagrange equation is $0=\mathcal{L}_y-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{y}}$. Seeking a simple solution, assume $\mathcal{L}_y=ay+b$ so that

$$\mathcal{L}=\frac{ay^2}{2}+by+f(\dot{y},t) \tag{96}$$

for some function $f(\dot{y},t)$. Then we need

$$\ddot{y}=-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{y}}=-\frac{\mathrm{d}}{\mathrm{d}t}\left[f_{\dot{y}}\right]=-f_{\dot{y}\dot{y}}\ddot{y}+f_{\dot{y}t}. \tag{97}$$

This equation holds when we assume $f_{\dot{y}\dot{y}}=-1$ and $f_{\dot{y}t}=0$, which implies $f_{\dot{y}}=-\dot{y}$ and so

$$f(y,t)=-\frac{\dot{y}^2}{2}+g(t) \tag{98}$$

for some function $g(t)$. Taking $g(t)=0$ and compiling our results, we obtain the Lagrangian

$$\boxed{\mathcal{L}=-\frac{\dot{y}^2}{2}+\frac{ay^2}{2}+by.} \tag{99}$$

□

Remark 25: The above example can be checked by computing the appropriate derivatives of $\mathcal{L}$ and verifying the Euler-Lagrange equation holds. ⋄

Remark 26: Here we discuss whether the Hamiltonian for a system is conserved. Let $H = H(q_k, p_y, t)$ be the Hamiltonian for a system. (Note here we have position variables $q_1, \ldots, q_n$ and momentum variables $p_1, \ldots, p_n$.) Then

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{k=1}^{n} \left(\frac{\partial H}{\partial q_k}\dot{q}_k + \frac{\partial H}{\partial p_k}\dot{p}_k \right) + \frac{\partial H}{\partial t}. \tag{100}$$

But, recall

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \text{and} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k} \quad \text{for } k = 1, \ldots, n. \tag{101}$$

Thus

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{k} \left(\frac{\partial H}{\partial q_k}\frac{\partial H}{\partial p_k} + \frac{\partial H}{\partial p_k}\left(-\frac{\partial H}{\partial q_k} \right) \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}. \tag{102}$$

This shows that the total derivative of H with respect to time is equal to its partial derivative with respect to time. This is a neat result that enables us to immediately deduce H is conserved whenever H does not contain an explicit time dependence. And, when $H = E$, i.e., the Hamiltonian gives the total energy of the system, we see can quickly see whether energy is conserved by computing $\partial H/\partial t$. $\diamond$

Section 4: Minimization with Constraints

We now turn our attention to the minimization of a function with respect to constraints. From our unconstrained material, we know a function is minimized when its gradient is zero. However, how does this change when subject to certain constraints? Do we still seek a constraints-compatible point for which the gradient is zero? It turns out, this is not the case. However, we shall see our primary result for equality constrained optimization, in fact, yields the previous result of the gradient equaling zero in the special case where we set our constraint set to be all of $\mathbb{R}^n$. We first show this in the case of equality constraints.

4.1 – Equality Constraints:

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ be smooth functions. Then we can write the constrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad h(x) = 0. \tag{103}$$

The central result here of Lagrange is that a solution x^* to (103) satisfies

$$\nabla f(x^*) = \lambda \nabla h(x^*) \tag{104}$$

for some $\lambda \in \mathbb{R}$. In the case where h is the zero function (i.e., $h(x) = 0$ for each $x \in \mathbb{R}^n$), we see (104) implies the gradient of f will be zero for any solution x^* to (103). We also may use second order information of f to determine sufficient conditions for a point x to be a solution to (103). We now state the key result used in this section.

Lagrange's Theorem: Let x^* be a local minimizer of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfying $h(x^*) = 0$ where $h : \mathbb{R}^n \to \mathbb{R}$ with $m \leq n$. Then there is $\lambda^* \in \mathbb{R}$ such that

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0. \tag{105}$$

△

REMARK 27: For minimization problems like the following example, we consider roughly the following steps:

1. Define an appropriate function f to minimize and constraint function g.

2. Use Lagrange's multiplier theorem to assert any minimizer x^* is a solution to the equation $\nabla f(x^*) = \lambda \nabla g(x^*)$ for some $\lambda \in \mathbb{R}$.

3. Also note x^* satisfies the constraint $g(x^*) = 0$. Use this fact with the equation in the previous result to find a list of potential minimizers.

4. From this, hopefully small, list of candidate minimizers, we can then check to see which points are minimizers or maximizers or saddle points.

◇

Example 22: Consider the constrained optimization problem

$$\min \ 4x_1 + 3x_2 - 10 \quad \text{subject to} \quad x_1x_2 = 12. \tag{106}$$

Use Lagrange's theorem to find all possible local minimizers and maximizers. Which of these are local minimizers?

Solution:

First we set

$$f(x_1, x_2) := 4x_1 + 3x_2 - 10 \quad \text{and} \quad h(x) := x_1x_2 - 12 \tag{107}$$

so that, writing $x = (x_1, x_2)$, our Lagrangian $\mathcal{L}$ becomes $\mathcal{L}(x, \lambda) = f(x) + \lambda h(x)$. Then Lagrange's theorem for multipliers asserts any solution x^* to (106) satisfies

$$\nabla \mathcal{L}(x^*) = 0. \tag{108}$$

More explicitly, this means

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \nabla \mathcal{L}(x^*) = \nabla f(x^*) + \lambda \nabla h(x^*) = \begin{bmatrix} 4 \\ 3 \end{bmatrix} + \lambda \begin{bmatrix} x_2^* \\ x_1^* \end{bmatrix}, \tag{109}$$

which implies

$$x^* = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = -\frac{1}{\lambda} \begin{bmatrix} 3 \\ 4 \end{bmatrix}. \tag{110}$$

Using our constraint, we then discover

$$0 = h(x^*) = \left(-\frac{3}{\lambda}\right)\left(-\frac{4}{\lambda}\right) - 12 = 12\left(\frac{1}{\lambda^2} - 1\right) \quad \Rightarrow \quad \lambda \in \{-1, 1\}. \tag{111}$$

Thus the possible local extrema corresponding to $\lambda = -1$ and $\lambda = 1$, respectively, are

$$\boxed{\begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{and} \quad -\begin{bmatrix} 3 \\ 4 \end{bmatrix}}. \tag{112}$$

Then observe $f(3, 4) = 4(3) + 3(4) - 10 = 14$ and $f(-3, -4) = 4(-3) + 3(-4) - 10 = -34$, and so $\boxed{(-3, -4)}$ is the local minimizer.[2] □

[2] This is not really a rigorous argument for why this is a local minimizer. We could use second order information, but that is currently beyond the scope of what has been covered so far in class.

Example 23: Consider a disk rolling without slipping on an incline plane. Determine the equation of constraint in terms of the coordinates y and θ where θ is the angle of rotation, R is the radius of the disk, and y is the distance the disk has moved from its initial position on the slope.

Solution:

Here we recall $y = R\theta$ is the arc length through an angle θ. This implies the equation g of constraint is $g(y, \theta) = y - R\theta = 0$. □

Now suppose we wish to minimize a functional $J(y)$ subject to some imposed constraint. Suppose

$$J(y) = \int_a^b \mathcal{L}(x, y, \dot{y}) \, \mathrm{d}x \quad \text{and} \quad \int_a^b G(x, y, \dot{y}) \, \mathrm{d}x = 0. \tag{113}$$

Set $\mathcal{L}^* = \mathcal{L} + \lambda G$ for some $\lambda \in \mathbb{R}$. Then Lagrange's theorem of multipliers tells us a minimizer $\bar{y}$ satisfies

$$\mathcal{L}^*_y - \frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{L}^*_{y'}\right] = 0 \quad \text{and} \quad G(\bar{y}) = 0. \tag{114}$$

In a more general case, if we have several variables, e.g., $y_1(x), \ldots, y_n(x)$ and a single constraint, then the first order optimality conditions are

$$\mathcal{L}^*_{y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{L}^*_{y_i}\right] = 0 \quad \text{and} \quad G(y_1, \ldots, y_n) = 0 \quad \text{for } i = 1, \ldots, n. \tag{115}$$

Remark 28: Note the following problems follow similarly to before. Here we need to identify $\mathcal{L}$ and G and define $\mathcal{L}^*$. Then we write the corresponding Euler Lagrange equation from Lagrange's theorem for multipliers. From this, we obtain an ODE. ◇

Example 24: Write down the first order questions that determine the solution of the problem

$$\min_{y\in C^2[a,b]} J(y) \quad \text{subject to} \quad \int_a^b r(x)y^2 \, \mathrm{d}x = 1, \tag{116}$$

where

$$J(y) = \int_a^b p(x)y'^2 + q(x)y^2 \, \mathrm{d}x. \tag{117}$$

Solution:

First observe our Lagrangian $\mathcal{L}$ is given by the integrand of J and let G be the integrand of our constraint function. Then set $\mathcal{L}^* = \mathcal{L} + G$. Then Lagrange's theorem for multipliers implies any solution to the optimization problem satisfies the Euler Lagrange equation

$$0 = \mathcal{L}^*_y - \frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{L}^*_{y'}\right] = [2qy + \lambda 2ry] - \frac{\mathrm{d}}{\mathrm{d}x}\left[2py'\right] = 2\left[qy + \lambda ry - (py')'\right], \tag{118}$$

which implies

$$(py')' - qy = \lambda ry. \tag{119}$$

□

Example 25: Write down the first order questions that determine the solution of the problem

$$\min_{y\in C^2[a,b]} J(y) \quad \text{subject to} \quad \int_a^b y \; \mathrm{d}x = 0, \tag{120}$$

where

$$J(y) = \int_a^b y'^2 + y^2 \; \mathrm{d}x. \tag{121}$$

Solution:

First observe our Lagrangian $\mathcal{L}$ is given by the integrand of J and let G be the integrand of our constraint function. Then set $\mathcal{L}^* = \mathcal{L} + G = y'^2 + y^2 + \lambda y$. Then Lagrange's theorem for multipliers implies any solution to the optimization problem satisfies the Euler Lagrange equation

$$0 = \mathcal{L}^*_y - \frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{L}^*_{y'}\right] = [2y + \lambda] - \frac{\mathrm{d}}{\mathrm{d}x}\left[2y'\right] = 2y + \lambda - 2y''. \tag{122}$$

□

Remark 29: The above ODE can be solved for a particular solution, e.g., $y_P = -\lambda/2$ and a homogeneous solution $y_H = c_1 e^x + c_2 e^{-x}$ to get the solution $y = c_1 e^x + c_2 e^{-x} - \lambda/2$. Then we can plug this into our integral constraint and deduce further information about c_1 and c_2. ⋄

Section 5: Eigenvalue Problems and Fourier Series

We now turn our attention to expansions of functions in terms of basis functions and eigenvalue problems. With this the following definitions will be useful

Definition: A function $f : (a, b) \to \mathbb{R}$ is in $L^2(a, b)$ provided

$$\left(\int_a^b f^2 \, \mathrm{d}x \right)^{1/2} < \infty. \tag{123}$$

△

Definition: The space of functions $L^2(a, b)$ is a Hilbert space with inner product for $f, g \in L^2(a, b)$ given by

$$\langle f, g \rangle_{L^2(a,b)} := \int_a^b fg \, \mathrm{d}x \tag{124}$$

and norm

$$\|f\|_{L^2(a,b)} := \sqrt{\langle f, f \rangle_{L^2(a,b)}}. \tag{125}$$

△

Example 26: Consider the eigenvalue problem $y' = \lambda y$ on the interval $[0,1]$, with $\lambda \in \mathbb{R}$ unknown.

a) What are the eigenfunctions if we impose the boundary condition $y(0) = 2$?

b) Are these eigenfunctions orthogonal with respect to the L^2 inner product on $(0,1)$?

Solution:

a) If $\lambda = 0$, then y is constant and the boundary condition implies $y(x) = 2$ for each $x \in [0,1]$. If $\lambda \neq 0$, then using separation of variables we may write

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lambda y \quad \Rightarrow \quad \frac{\mathrm{d}y}{y} = \lambda \ \mathrm{d}x \quad \Rightarrow \quad \ln(y) = \int \frac{\mathrm{d}y}{y} = \lambda x + c_1. \tag{126}$$

This implies

$$y = \exp\left(\lambda x + c_1\right) = c_2 \exp\left(\lambda x\right), \tag{127}$$

where $c_2 = \exp(c_1)$. Using the boundary condition, we deduce

$$2 = y(0) = c_2 \exp(\lambda 0) = c_2. \tag{128}$$

Consequently, the family of eigenfunctions is given by $\{2\exp(\lambda x) \ : \ \lambda \in \mathbb{R}\}$.

b) We say two functions $f, g \in L^2(0,1)$ with $f \neq g$ are orthogonal provided

$$0 = \langle f, g\rangle_{L^2(0,1)} = \int_0^1 f(x)g(x) \ \mathrm{d}x. \tag{129}$$

Let $f = 2\exp(\lambda x)$ and $g(x) := 2\exp(\alpha x)$ for scalars $\lambda, \alpha \in \mathbb{R}$ with $\lambda \neq \alpha$. Then f and g are in the family of eigenfunctions and

$$\begin{aligned}
\langle f, g\rangle_{L^2(0,1)} &= \int_0^1 f(x)g(x) \ \mathrm{d}x \\
&= \int_0^1 \left[2\exp(\lambda x)\right]\left[2\exp\left(\alpha x\right)\right] \ \mathrm{d}x \\
&= 4\int_0^1 \exp\left((\alpha+\lambda)x\right) \ \mathrm{d}x \\
&= \begin{cases} 4 & \text{if } \alpha + \lambda = 0, \\ \dfrac{4}{\alpha+\lambda}\left[\exp(\alpha+\lambda) - 1\right] & \text{otherwise.} \end{cases}
\end{aligned} \tag{130}$$

This shows $\langle f, g\rangle_{L^2(0,1)} \neq 0$ for distinct eigenfunctions f and g, from which we conclude the eigenfunctions are *not orthogonal.* □

Example 27: Consider the family of functions $\{e_n\}$ where for each $n \in \mathbb{N}$

$$e_n := \frac{\sin(n\pi t)}{\|\sin(n\pi t)\|_{L^2(0,1)}}. \tag{131}$$

a) Verify the collection $\{e_n\}$ is orthonormal.

b) Find the projection of the function $f(t) = t(1-t)$ into the space spanned by $\{e_n\}$.

c) Is f contained in the span of $\{e_n\}$?

Solution:

a) We first verify the collection $\{e_n\}$ is orthonormal. By definition,

$$\|e_n\|_{L^2(0,1)} = \left\| \frac{\sin(n\pi t)}{\|\sin(n\pi t)\|_{L^2(0,1)}} \right\| = \frac{\|\sin(n\pi t)\|_{L^2(0,1)}}{\|\sin(n\pi t)\|_{L^2(0,1)}} = 1 \quad \text{for } n \in \mathbb{N}. \tag{132}$$

This verifies each e_n has norm unity. All that remains is to show they are orthogonal. Suppose $e_n \neq e_m$. Then

$$\begin{aligned}
\langle e_n, e_m \rangle_{L^2(0,1)} &= \int_0^1 e_n(t) e_m(t) \ \mathrm{d}t \\
&= \int_0^1 \sin(n\pi t)\sin(m\pi t) \ \mathrm{d}t \\
&= \frac{1}{2}\int_0^1 \cos([n-m]\pi t) - \cos([n+m]\pi t) \ \mathrm{d}t \\
&= \frac{1}{2}\left[\frac{\sin([n-m]\pi t)}{[n-m]\pi} - \frac{\sin([n+m]\pi t)}{[n+m]\pi} \right]_0^1 \\
&= 0,
\end{aligned} \tag{133}$$

where the final equality holds since $\sin(k\pi) = \sin(0) = 0$ for each integer k. Whence we conclude the collection $\{e_n\}$ is orthonormal.

b) For each $N \in \mathbb{N}$, we define the expansion S_N of f in terms of $\{e_n\}$ to be

$$S_N := \sum_{n=1}^{N} \langle f, e_n \rangle \, e_n. \tag{134}$$

Let $V := \operatorname{span}\{e_n\}$ and note V is a subspace of $L^2(0,1)$. Moreover, since the e_n's are orthonormal,

$$\text{proj}_V(f) = \lim_{N\to\infty} S_N. \tag{135}$$

To determine this expansion S_N, we must make a few computations. First note

$$\langle f, \sin(n\pi t)\rangle = \int_0^1 t(1-t)\sin(n\pi t)\ \mathrm{d}t = \int_0^1 t\sin(n\pi t) - t^2\sin(n\pi t)\ \mathrm{d}t. \tag{136}$$

Using integration by parts, we see

$$\begin{aligned}\int_0^1 t\sin(n\pi t)\ \mathrm{d}t &= \left[\frac{-t\cos(n\pi t)}{n\pi}\right]_0^1 - \int_0^1 -\frac{\cos(n\pi t)}{n\pi}\ \mathrm{d}t \\ &= \left[\frac{-t\cos(n\pi t)}{n\pi}\right]_0^1 + \left[\frac{\sin(n\pi t)}{(n\pi)^2}\right]_0^1 \\ &= \frac{(-1)^{n+1}}{n\pi}.\end{aligned} \tag{137}$$

Similarly,

$$\begin{aligned}\int_0^1 t^2\sin(n\pi t)\ \mathrm{d}t &= \left[\frac{-t^2\cos(n\pi t)}{n\pi}\right]_0^1 - \int_0^1 -\frac{2t\cos(n\pi t)}{(n\pi)^2}\ \mathrm{d}t \\ &= \left[\frac{-t^2\cos(n\pi t)}{n\pi}\right]_0^1 + 2\left[\frac{t\sin(n\pi t)}{(n\pi)^3}\right]_0^1 - 2\int_0^1 \frac{\sin(n\pi t)}{(n\pi)^2}\ \mathrm{d}t \\ &= \left[\frac{-t^2\cos(n\pi t)}{n\pi}\right]_0^1 + 2\left[\frac{t\sin(n\pi t)}{(n\pi)^3}\right]_0^1 + 2\left[\frac{\cos(n\pi t)}{(n\pi)^3}\right]_0^1 \\ &= \frac{(-1)^{n+1}}{n\pi} + \frac{2\left[(-1)^n - 1\right]}{(n\pi)^3}.\end{aligned} \tag{138}$$

Then (136), (137), and (138) together imply

$$\langle f, \sin(n\pi t)\rangle = \frac{(-1)^{n+1}}{n\pi} - \left[\frac{(-1)^{n+1}}{n\pi} + \frac{2\left[(-1)^n - 1\right]}{(n\pi)^3}\right] = \frac{2\left[1 - (-1)^n\right]}{(n\pi)^3}. \tag{139}$$

To compute $\langle f, e_n\rangle$, we need the additional fact that

$$\|\sin(n\pi t)\|^2_{L^2(0,1)} = \int_0^1 \sin^2(n\pi t)\ \mathrm{d}t = \frac{1}{2}\int_0^1 1 - \cos(2n\pi t)\ \mathrm{d}t = \frac{1}{2}\left[t - \frac{\sin(2n\pi t)}{2n\pi}\right]_0^1 = \frac{1}{2}. \tag{140}$$

Thus, using the definition of e_n and linearity of the inner product, we deduce

$$\langle f, e_n\rangle_{L^2(0,1)} = \sqrt{2}\,\langle f, \sin(n\pi t)\rangle_{L^2(0,1)} = \begin{cases} \dfrac{4\sqrt{2}}{(n\pi)^3} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.}\end{cases} \tag{141}$$

This implies for each nonnegative integer N the expansion is given by

$$\boxed{S_{2N+2} = S_{2N+1} = \sum_{n=0}^{N} \left(\frac{4\sqrt{2}}{[(2n+1)\pi]^3} \right) e_{2n+1}.} \tag{142}$$

Here we use the subscript $2n+1$ to only write the odd terms.

c) Because $\{e_n\}$ is orthonormal, the square of the norm of S_{2N+1} is given by the sum of the squares of the coefficients for each e_n. In mathematical terms, we write

$$\|\text{proj}_V(f)\|^2_{L^2(0,1)} = \lim_{N\to\infty} \|S_{2N+1}\|^2_{L^2(0,1)} = \lim_{N\to\infty} \sum_{n=0}^{N} \left(\frac{4\sqrt{2}}{[(2n+1)\pi]^3} \right)^2 = \frac{1}{30}, \tag{143}$$

where the sum was evaluated using Wolfram. And, because

$$\|f\|^2_{L^2(0,1)} = \int_0^1 [t(1-t)]^2 \ \text{d}t = \int_0^1 t^2 - 2t^3 + t^4 \ \text{d}t = \frac{1}{30}, \tag{144}$$

we see $\|\text{proj}_V(f)\|_{L^2(0,1)} = \|f\|_{L^2(0,1)}$. But, from linear algebra we know either

$$\|\text{proj}_V(f)\|_{L^2(0,1)} < \|f\|_{L^2(0,1)} \quad \text{or} \quad \text{proj}_V(f) = f. \tag{145}$$

Consequently, we conclude the expansion converges to f in $L^2(0,1)$ and $f \in V$.

□

Remark 30: Note f is symmetric about $t = 1/2$. So, we expect only the odd terms e_{2n+1} to be included in the expansion since these are symmetric about $t = 1/2$ while the even terms e_{2n} are anti-symmetric about $t = 1/2$. ◇

Remark 31: We can naturally extend f smoothly to the interval $[-1,1]$ so its extension is an odd function (i.e., $f(t) = f(-t)$) and so we should expect it to be expressible as an expansion of sine functions since these are also odd while cosine functions are even. ◇

Example 28: Consider the collection of polynomials

$$S := \{v_1, v_2, v_3, v_4\} := \left\{1, \quad x, \quad \frac{1}{2}\left(3x^2-1\right), \quad \frac{1}{2}\left(5x^3-3x\right)\right\}. \tag{146}$$

a) Verify the collection S contains orthogonal function with respect to the $L^2(-1,1)$ inner product.

b) In the space $L^2(-1,1)$, expand the function $f(x) := x^2$ in terms of the family of functions given. Is $f \in \text{span}(S)$?

Solution:

a) We must verify

$$\langle v_i, v_j \rangle_{L^2(-1,1)} = 0 \quad \text{for } i,j = 1,2,3,4 \text{ with } i \neq j. \tag{147}$$

For notational compactness, in this problem set $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{L^2(-1,1)}$. Then

$$\begin{aligned}
\langle v_1, v_2 \rangle &= \int_{-1}^{1} x \ \mathrm{d}x = \left[\frac{x^2}{2}\right]_{-1}^{1} = 0,\\
\langle v_1, v_3 \rangle &= \int_{-1}^{1} \frac{1}{2}\left(3x^2-1\right) \ \mathrm{d}x = \frac{1}{2}\left[x^3 - x\right]_{-1}^{1} = 0,\\
\langle v_1, v_4 \rangle &= \int_{-1}^{1} \frac{1}{2}\left(5x^2-3x\right) \ \mathrm{d}x = \frac{1}{2}\left[\frac{5x^4}{4} - \frac{3x^2}{2}\right]_{-1}^{1} = \frac{1}{2}\left[\left(\frac{5}{4}-\frac{3}{2}\right) - \left(\frac{5}{4}-\frac{3}{2}\right)\right] = 0,\\
\langle v_2, v_3 \rangle &= \int_{-1}^{1} \frac{1}{2}\left(3x^3-x\right) \ \mathrm{d}x = \frac{1}{2}\left[\frac{3x^4}{4} - \frac{x^2}{2}\right]_{-1}^{1} = \frac{1}{2}\left[\left(\frac{3}{4}-\frac{1}{2}\right) - \left(\frac{3}{4}-\frac{1}{2}\right)\right] = 0,\\
\langle v_2, v_4 \rangle &= \int_{-1}^{1} \frac{1}{2}\left(5x^4-3x^2\right) \ \mathrm{d}x = \frac{1}{2}\left[x^5 - x^3\right]_{-1}^{1} = 0,\\
\langle v_3, v_4 \rangle &= \int_{-1}^{1} \frac{1}{4}\left(15x^5-14x^3+3x\right) \ \mathrm{d}x = \frac{1}{4}\left[\frac{15x^6}{6} - \frac{14x^4}{4} + \frac{3x^2}{2}\right]_{-1}^{1} = 0.
\end{aligned} \tag{148}$$

Consequently, this shows (147) holds and we conclude the collection S forms an orthogonal set of functions.

b) Observe

$$\begin{aligned}
\langle v_1, f\rangle &= \int_{-1}^{1} x^2 \, \mathrm{d}x = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{2}{3}, \\
\langle v_2, f\rangle &= \int_{-1}^{1} x^3 \, \mathrm{d}x = 0 \\
\langle v_3, f\rangle &= \int_{-1}^{1} \frac{1}{2}\left(3x^4 - x^2\right) \, \mathrm{d}x = \frac{1}{2}\left[\frac{3x^5}{5} - \frac{x^3}{3}\right]_{-1}^{1} = \frac{4}{15}, \\
\langle v_4, f\rangle &= \int_{-1}^{1} \frac{1}{2}\left(5x^5 - 3x^3\right) \, \mathrm{d}x = \frac{1}{2}\left[x^6 - x^4\right]_{-1}^{1} = 0.
\end{aligned} \tag{149}$$

Then the projection of f into $\mathrm{span}(S)$ is given by

$$\mathrm{proj}_V(f) = \sum_{k=1}^{4} \frac{\langle v_k, f\rangle}{\langle v_k, v_k\rangle} v_k, \tag{150}$$

where the division by $\langle v_k, v_k\rangle$ is necessary here in order to normalize the v_k's to obtain an orthonormal family. Using (149) and the fact

$$\langle v_1, v_1\rangle = \int_{-1}^{1} 1 \, \mathrm{d}x = 2 \quad \text{and} \quad \langle v_3, v_3\rangle = \int_{-1}^{1} \frac{1}{4}\left(3x^2 - 1\right)^2 \, \mathrm{d}x = \frac{2}{5}, \tag{151}$$

we deduce

$$\begin{aligned}
\mathrm{proj}_V(f) &= \frac{\langle v_1, f\rangle}{\langle v_1, v_1\rangle} v_1 + \frac{\langle v_3, f\rangle}{\langle v_3, v_3\rangle} v_3 \\
&= \frac{2/3}{2} v_1 + \frac{4/15}{2/5} v_3 \\
&= \frac{1}{3} v_1 + \frac{2}{3} v_3 \\
&= \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{2}\left(3x^2 - 1\right) \\
&= x^2.
\end{aligned} \tag{152}$$

Whence $\mathrm{proj}_V(f) = f$, and so $f \in \mathrm{span}(S)$.

□

Example 29: Suppose we wish to use the collection S of functions in the previous example to expand a function instead on the interval $[1, 2]$.

a) Describe a change of variable formula that could be used to transform the collection S into corresponding functions on $[1, 2]$.

b) Plot v_2 and v_3 and the corresponding functions defined on $[1, 2]$.

Solution:

a) We seek a linear transformation that will take a function defined on $[-1, 1]$ to a function defined on $[1, 2]$. To do this, define $T : [-1, 1] \to \mathbb{R}$ by

$$T(x) := \frac{1}{2}x + \frac{3}{2}. \tag{153}$$

Then the inverse formula would be

$$T^{-1}(x) = 2x - 3. \tag{154}$$

Then under this transformation we obtain the collection $\{w_1, w_2, w_3, w_4\} \subset L^2(1, 2)$ where $w_i(x) := v_i(T^{-1}(x))$ for $i = 1, 2, 3, 4$.

b) Below we give the desired plots.

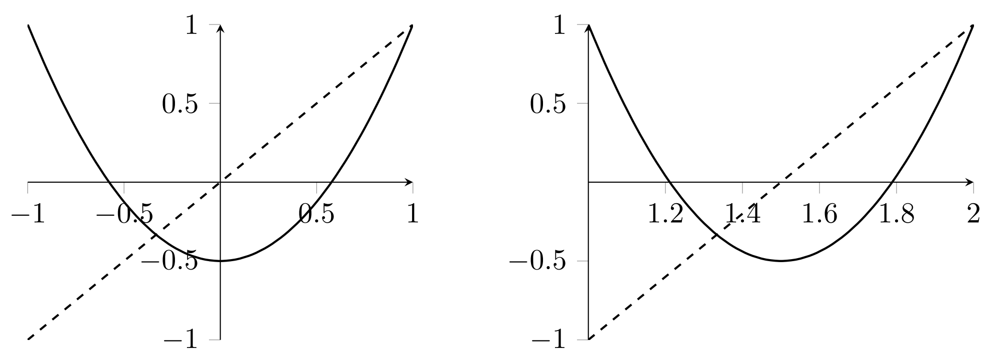

Figure 5: Plot of v_2 and v_3 on the left and w_2 and w_3 on the right.

□

REMARK 32: Suppose we wanted to repeat the previous exercise, but going from a general interval $[a, b]$ to another interval $[c, d]$, i.e., suppose we seek a linear transformation $T : [a, b] \to \mathbb{R}$ such $T([a, b]) = [c, d]$. Since T is linear, this is actually given by the first order Lagrange interpolating polynomial with the data $T(a) = c$ and $T(b) = d$. That is,

$$T(x) = \frac{x-a}{b-a} \cdot T(b) + \frac{x-b}{a-b} \cdot T(a). \tag{155}$$

Consequently, upon rearranging we obtain

$$T(x) = \frac{T(b) - T(a)}{b-a} \cdot x + \frac{bT(a) - aT(b)}{b-a} = \frac{d-c}{b-a} \cdot x + \frac{bc - ad}{b-a}. \tag{156}$$

Returning the above example, there we have $[a, b] = [-1, 1]$ and $[c, d] = [1, 2]$. Then (156) implies

$$T(x) = \frac{2-1}{1-(-1)} \cdot x + \frac{1(1) - (-1)2}{1-(-1)} = \frac{1}{2} \cdot x + \frac{3}{2}, \tag{157}$$

which is precisely what was obtained in (153). $\diamond$

Example 30: Consider the eigenvalue problem

$$-x^2y'' - 2xy' + \frac{27}{4}y = \lambda y \quad \text{in} \quad [1,2], \quad \text{subject to} \quad y(1) = 1, \quad y(2) = 2. \tag{158}$$

a) Is this a Sturm-Liouville problem? If so, is it regular?

b) Find the form of the candidate eigenfunctions when $\lambda < 9$.

c) Find the eigenfunction corresponding to $\lambda = 5$.

Solution:

a) Observe the ODE in (158) may be written as

$$-(py')' + qy = \lambda \quad \text{in} \quad [1,2] \quad \text{subject to} \quad y(1) = 1, \quad y(2) = 2. \tag{159}$$

where $p(x) = x^2$ and $q(x) = 27/4$, which is precisely the form of a Sturm-Liouville problem. Moreover, because p, p', and q are continuous on $[1,2]$ and p is never zero in $[1,2]$, the problem is a regular Sturm-Liouville problem.

b) The ODE in (158) is a Cauchy-Euler equation, which has solutions of the form x^m. Suppose $y = x^m$. Then for $x \in [1,2]$

$$\begin{aligned} 0 &= -x^2y'' - 2xy' + (27/4 - \lambda)y \\ &= -m(m-1)x^m - 2mx^m + (27/4 - \lambda)x^m \\ &= \left[-m^2 - 3m + (27/4 - \lambda)\right] x^m. \end{aligned} \tag{160}$$

This yields the characteristic equation

$$0 = -m^2 - 3m + (27/4 - \lambda) \quad \Rightarrow \quad m = \frac{3 \pm \sqrt{3^2 + 4(27/4 - \lambda)}}{-2} = -\frac{3}{2} \mp \sqrt{9 - \lambda}. \tag{161}$$

Consequently, for $\lambda < 9$ we have eigenfunctions of the form

$$y(x) = c_1 x^{-3/2 - \sqrt{9-\lambda}} + c_2 x^{-3/2 + \sqrt{9-\lambda}}. \tag{162}$$

Applying our boundary condition, we see

$$1 = y(1) = c_1 \cdot 1 + c_2 \cdot 1 \quad \Rightarrow \quad c_1 = 1 - c_2. \tag{163}$$

Given λ, the second boundary condition $y(2) = 2$ can be used to find c_1 and c_2 explicitly.

c) For $\lambda = 5$ we see $\sqrt{9-\lambda} = \sqrt{9-5} = \sqrt{4} = 2$, which implies the formula in (162) becomes

$$y(x) = c_1 x^{-3/2-2} + c_2 x^{-3/2+2} = c_1 x^{-7/2} + c_2 x^{1/2} = (1-c_2)x^{-7/2} + c_2 x^{1/2}. \tag{164}$$

Then the second boundary condition yields

$$2 = y(2) = (1-c_2)2^{-7/2} + c_2 x^{1/2} \quad \Rightarrow \quad c_2 = \frac{16\sqrt{2}-1}{15}, \tag{165}$$

where the algebraic steps were completed using Wolfram. Thus the eigenfunction is

$$\boxed{y(x) = \left(1 - \frac{16\sqrt{2}-1}{15}\right) x^{-7/2} + \frac{16\sqrt{2}-1}{15} x^{1/2}.} \tag{166}$$

□

Remark 33: The following example briefly discusses how to obtain a solution to a particular type of ODE. Howevever, I have received word from Professor Meszaros that he wants this type of problem to be completed using material from 326-327 and 348-349 in our text, **not** as was proposed below. The idea is to expand the right hand side in terms of the eigenfunctions of the differential operator on the left hand side. One can do a correspondence for the coefficients. The solution will then be given in terms of an eigenfunction expansion. ⋄

Example 31: Define the differential operator $L : C^2[0, 1] \to \mathbb{R}$ by $Ly := y'' - 2y' + 2y$. and consider the ODE

$$Ly = f \quad \text{subject to} \quad y(0) = y(1) = 0. \tag{167}$$

If f is given to be identically zero, then does a solution to the ODE exist? If so, give a solution and describe how it could be represented in terms of a cosine series expansion and in terms of a sine series expansion.

Solution:

As in the previous example, the ODE

$$y'' - 2y' = 0 \quad \text{in } [0, 1] \tag{168}$$

is a Cauchy-Euler equation. Here solutions are of the form $y = x^m$. Plugging this in, we obtain the characteristic equation

$$0 = m(m - 1) - 2m + 2 = m^2 - 3m + 2 = (m - 1)(m - 2), \tag{169}$$

which implies the general solution is

$$y(x) = c_1 x + c_2 x^2. \tag{170}$$

Applying the second boundary condition we see

$$0 = y(1) = c_1 + c_2 \quad \Rightarrow \quad c_1 = -c_2. \tag{171}$$

The first boundary condition does not give an additional info, and so we see

$$y(x) = \alpha x(1 - x) \tag{172}$$

is a solution for each $\alpha \in \mathbb{R}$. In particular $y(x) = x(1-x) = x - x^2$ is a solution.

We leave the full details of the expansion of $y(x)$ in terms of sines and cosines to the reader as these are quite similar to previous results. Note here we may use either the family

$$\{\sin(n\pi x) \ : \ n \in \mathbb{N}\} \quad \text{or} \quad \{\cos(n\pi x) \ : \ n \in \mathbb{N} \cup \{0\}\} \tag{173}$$

for the expansion. For more explicit explanation, confer the following theorem. □

Half-Range Expansions Theorem: Suppose f is a smooth function defined on $(0, p)$. Then f has a *cosine series expansion*

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{p}x\right) \quad \text{for } x \in (0, p), \tag{174}$$

where

$$a_0 = \frac{1}{p}\int_0^p f(x)\ \mathrm{d}x \quad \text{and} \quad a_n = \frac{2}{p}\int_0^p f(x)\cos\left(\frac{n\pi}{p}x\right)\ \mathrm{d}x \quad \text{for } n \in \mathbb{N}. \tag{175}$$

Similarly, f has a *sine series expansion*

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{p}x\right)\ \mathrm{d}x \quad \text{for } x \in (0, p), \tag{176}$$

where

$$b_n = \frac{2}{p}\int_0^p f(x)\sin\left(\frac{n\pi}{p}x\right)\ \mathrm{d}x \quad \text{for } n \in \mathbb{N}. \tag{177}$$

That is, on $(0, p)$ each of the series converge to $f(x)$. If f is only piecewise smooth, then at points where f is discontinuous the series converge to $(f(x^+) + f(x^-))/2$. △

Section 6: Summary

Herein we summarize the material covered thus far in these discussion notes and what a student might be expected to know.

1. Be able to define a vector space and determine whether some set with associated addition and scalar multiplication forms a set (e.g., see the second problem in Homework 1).

2. Be able to define and distinguish and relationship between the terms *minimum, lower bound,* and *minimizer* in the context of this course. Also given a simple functional $J : V \to \mathbb{R}$ and an admissibility class $\mathcal{A}$, be able to find each of these corresponding things.

3. Be able to distinguish between infimum and minimum.

4. Given a functional J and admissibility class $\mathcal{A}$, be able to determine an appropriate vector space V that contains (i.e., so that $\mathcal{A} \subseteq V$).

5. Be able to find lower bounds for a functional J over an admissibility class $\mathcal{A}$ using basic methods (e.g., see Example 7).

6. Be able to compute the Gâteaux derivative of a functional J, denoted $DJ(y)h$. (This has been heavily emphasized so far in the quarter.)

7. Understand the Euler-Lagrange equation and how to use it to determine differential equations satisfied by extremals of a functional J of the form

$$J(y) = \int_a^b L(x, y, y') \ \mathrm{d}x. \tag{178}$$

Remark 34: Students please note these discussion notes have not yet discussed natural boundary conditions and the slight difference this creates in the problem solution. I highly recommend reading Section 4.4.3 of our text. ⋄