Classical Mathematical Methods in Engineering
Lecture 4 Monday January 26, 2026 (snow day)
Two-point boundary value problem(s)
and Sturm-Liouville Theory

John McCuan

January 26, 2026

1 Two point boundary value problems

1.1 Last time...

We considered briefly the two point boundary value problem

$$\begin{cases} x'' + t^2 x = t^3, \quad 0 \le t \le L \\ x(0) = x_0, \\ x(L) = x_L. \end{cases} \tag{1}$$

We found a basis of solutions for the homogeneous problem

$$x_h'' + t^2 x_h = 0$$

having the form

$$X_0 = 1 + \sum_{k=1}^{\infty} \alpha_k t^{4k}$$

and

$$X_1 = t + \sum_{k=1}^{\infty} \beta_k t^{4k+1}.$$

With these we attempted to solve the implied two point boundary problem

$$\begin{cases} x_h'' + t^2 x_h = 0, \qquad 0 \le t \le L \\ x_h(0) = x_0, \\ x_h(L) = x_L - L. \end{cases} \tag{2}$$

for x_h. The left boundary condition implied $a_0 = x_0$, but the second boundary condition

$$a_1 X_1(L) = x_L - L - x_0 X_0(L) \tag{3}$$

was noted to be complicated when $X_1(L) = 0$. We plotted X_1 using partial sums of the power series and produced pretty good evidence that X_1 has at least one positive zero z_1 with $z_1 \doteq 2.35834$. See Figure 1. Thus, if $0 < L < z_1$ we can solve for the coefficient a_1 and solve the problem. If $L = z_1$ things get complicated. There are probably some values of $L > z_1$ where we can solve the problem, but it is not so clear what happens after that, and definitely our ability to discern what is happening for L large using the power series seems limited.

We seek in the end to find a more general framework according to which we might know what to expect concerning the solvability of (3) and hence the solvability of (2) and hence the solvability of (1).

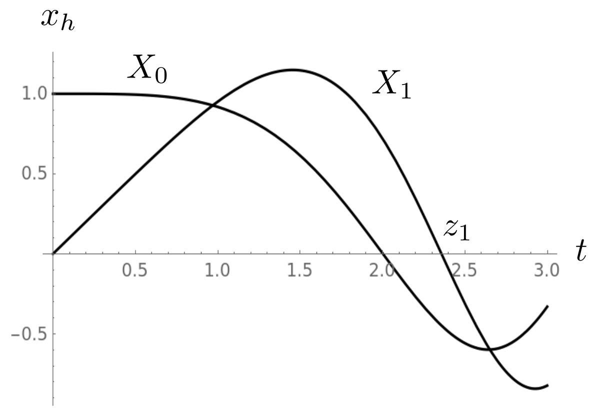

Figure 1: Plots of X_0 and X_1 from the power series.

1.2 Partial summary

Hopefully the discussion so far has served to provide something of a review of ODEs. In particular, hopefully you can recognize a second order linear operator and distinguish that operator from an inhomogeneity in the equation. There are many other kinds of ODEs and many other things it is good to know, but hopefully we've reviewed a point or two and suggested a thing or two that may be new to you:

1. In an equation like $x'' + t^2x = t^3$, the **linear operator** $L : C^2(\mathbb{R}) \to C^0(\mathbb{R})$ has values[1] given by $Lx = x'' + t^2x$. The function with values $q(t) = t^2$ here is a coefficient of the operator. The function with values $f(t) = t^3$ on the "right hand side" is the **inhomogeneity** in the equation. The inhomogeneity may also be considered a coefficient in the equation.

2. If one can find a particular solution, in this case $y_p = t$, then pretty much any questions one might have about this ODE can be reduced to corresponding questions about the **corresponding homogeneous ODE** $x_h'' + t^2x_h = 0$.

3. The **initial value problem** (IVP) has a good existence and uniqueness theory associated with it.

4. The **two point boundary value problem** (BVP) is generally more complicated.

5. In this example the solvability of the two point boundary value problem on an interval $[0, L]$ for $L > 0$ is intimately related to the positive zeros of X_1.

6. We do not have a very good handle at the moment on the situation with the positive zeros of X_1, though we can say something using **power series solutions**.

There are examples for which the two point boundary value problem always has a unique solution. You should see some problems like that in the homework assignment. I'm now going to consider a simpler second example to emphasize some of the complications that can arise and should be expected.

[1]It may be noticed that the symbol L has been used in these notes and in my lectures to mean two very different things. On the one hand, L denotes a second order linear ordinary differential operator, and on the other hand, the same symbol L denotes the right endpoint of an interval, so one might even see the clashing notation appear as $L : C^2[0, L] \to C^0[0, L]$. I'm going to keep going with this "inconsistency," and I hope you can follow the "ideas" closely enough so that such small inconsistencies in notation don't cause a problem. If you have a suggestion for better notation, I'm all ears. Obviously Haberman didn't figure out better notation as you can see on pages 36–37.

1.3 Simpler BVP

First of all I will concentrate on a simpler second order linear operator, namely

$$Lx = x'' + x.$$

I will also start immediately with the associated homogeneous ODE and have a look at the two point boundary value problem:

$$\begin{cases} x'' + x = 0, \quad 0 \leq t \leq L \\ x(0) = x_0, \\ x(L) = x_L. \end{cases} \tag{4}$$

Here the general solution also has the form $x = a_0 X_0(t) + a_1 X_1(t)$. You can find it/them using power series. You may also remember other approaches. I'll use one you should remember using a technique called "familiarity with functions." For this technique we look at the ODE and interpret it in words:

$$x'' = -x$$

means "find a function whose second derivative is minus the function." The functions $\cos t$ and $\sin t$ have this property, and $\{\cos t, \sin t\}$ is a basis of solutions for this homogeneous problem. Notice that taking $X_0(t) = \cos t$ and $X_1(t) = \sin t$ we have

$$\begin{aligned} X_0(0) &= 1 \\ X_0'(0) &= 0 \\ X_1(0) &= 0 \\ X_1'(0) &= 1 \end{aligned}$$

which should look familiar. These conditions are extremely well-suited for the IVP.

Exercise 1 *Pose the IVP at $t = 0$ associated with $x'' + x = 0$ and find the unique solution.*

The two point boundary value problem is another story. Only the zero order conditions $X_0(0) = \cos 0 = 1$ and $X_1(0) = 0$ are of immediate use yielding

$$x(0) = a_0 = x_0$$

as you should find for the initial value problem. After that we are faced with

$$x(L) = x_0 cos(L) + a_1 \sin(L) = x_L.$$

Thus the version of (3) arising here is

$$a_1 \sin(L) = x_L - x_0 \cos(L). \tag{5}$$

We know exactly what happens in this case: If $L \notin \{\pi, 2\pi, 3\pi, \ldots\} = \{k\pi : k = 1, 2, 3, \ldots\}$, then $\sin(L) \neq 0$, and we can solve (5) uniquely:

$$a_1 = \frac{x_L - x_0 \cos(L)}{\sin(L)}.$$

Consequently, we can solve the two point boundary value problem (4) uniquely:

$$x(t) = x_0 \cos t + \frac{x_L - x_0 \cos(L)}{\sin(L)} \sin t.$$

If $L = k\pi$ for some $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$ (the natural numbers), then either k is odd and has the form $k = 2m + 1$ for some $m \in \mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ (the natural numbers with zero) or k is even and has the form $k = 2m$ for some $m \in \mathbb{N}_0$. If $k = 2m + 1$ is odd and $L = (2m + 1)\pi$, then $\cos(L) = -1$ and

the problem (4) has no solution if $x_L \neq -x_0$, and
the problem (4) has infinitely many solutions $x(t) = x_0 \cos t + a_1 \sin t$ if $x_L = -x_0$.

Similarly, if $k = 2m$ is even and $L = 2m\pi$, then $\cos(L) = 1$ and

the problem (4) has no solution if $x_L \neq x_0$, and
the problem (4) has infinitely many solutions $x(t) = x_0 \cos t + a_1 \sin t$ if $x_L = x_0$.

The point is the following: One can see exactly what happens because one knows the function $\sin t$ and exactly where the zeros of $\sin t$ happen to fall.

1.4 A different problem

You should get the idea here that what happens with two point boundary value problems for ODEs can be somewhat complicated, especially when one isn't immediately familiar with the detailed properties of some basis solution for a homogeneous ODE. Mathematicians have responded to this situation in a couple different ways. One thing that has been done is that certain "special functions" that come up in a variety of applications and are basis solutions of particular homogeneous ODEs have been studied carefully and hard. Many things are known about these solutions, and maybe you know some of their names. Obviously the most famous are sine and cosine. There are also Bessel functions, Legendre polynomials, Airy functions, Hermite polynomials, and others. These functions will have standard implementations in mathematical software, and many properties are known about them.

The other response is a more general one, and also perhaps a somewhat unexpected one, that can be thought of in this way: Set aside the two point boundary value problem, and consider a different kind of problem called a Sturm-Liouville problem. We were able to solve (4) but let us begin by casting that problem into the Sturm-Liouville framework as a first example.

Instead of trying to solve

$$\begin{cases} x'' + x = 0, \quad 0 \leq t \leq L \\ x(0) = x_0, \\ x(L) = x_L \end{cases} \tag{6}$$

directly, introduce an additional parameter and make finding that parameter part of the problem. The Sturm-Liouville version of (6) is

$$\begin{cases} x'' + x + \lambda x = 0, \quad 0 \leq t \leq L \\ x(0) = 0, \\ x(L) = 0. \end{cases} \tag{7}$$

If you think about it, the introduction of an extra parameter—or an extra degree of freedom if you like—should make the problem easier to solve. On the other hand, one now has more things to solve for. On the the third hand (if you happen to have three hands) we have simplified the boundary values which have a couple interesting consequences. One of those consequences is that the problem always has a "trivial solution" $x(t) \equiv 0$. We'll discuss this more later.

In any case, the real advantage is that there is a general theory of existence and uniqueness that goes along with such problems and, in a certain sense, captures a consistent vision of what we (and others) found complicated about two point boundary value problems.

Let's try to solve this strange first Sturm-Liouville problem (7) without really knowing what we should expect for a solution. The first thing we might observe is that the general form of the solution of the ODE

$$x'' + (1+\lambda)x = 0,$$

assuming λ is some fixed real constant, depends on λ. If $1+\lambda > 0$, i.e., $\lambda > -1$, then we again get sines and cosines, but not exactly necessarily $\cos t$ and $\sin t$—those are only for $\lambda = 0$. On the other hand if $\lambda < -1$ something different happens. There are cases.

CASE 1: $\lambda < -1$.

Just to be clear, we are starting here with the assumption that λ is a fixed real number with $\lambda < -1$, and we want to then consider the two point boundary value problem

$$\begin{cases} x'' + (1+\lambda)x = 0, \quad 0 \le t \le L \\ x(0) = 0, \\ x(L) = 0 \end{cases}$$

involving the modified operator $Lx = x'' + (1+\lambda)x$. We could use "familiarity with functions" to find the general solution of this ODE, but as this is still something of a review of ODEs, let me recall a different approach with which you are probably familiar. This approach which is known to work, in some way, shape, or form, for constant coefficient linear ODEs is called "try an exponential." Accordingly, we write

$$x(t) = e^{\alpha t}$$

where the coefficient α is a real number to be determined. In this case writing $-(\lambda+1) = \mu^2 > 0$ because $\lambda + 1 < 0$ we find

$$x'' - \mu^2 x = \alpha^2 e^{\alpha t} - \mu^2 e^{\alpha t} = (\alpha^2 - \mu)^2 e^{\alpha t}.$$

The exponential function, in case you have forgotten, never vanishes—no matter what. Therefore, if we have $x'' - mu^2 x = 0$ in this case, then we must have

$$\alpha^2 - \mu^2 = (\alpha - \mu)(\alpha + \mu) = 0.$$

This is a polynomial equation for α having two real solutions corresponding to two basis functions for the ODE and a general solution of the form

$$x(t) = a_0 e^{\mu t} + a_1 e^{-\mu t} = a_0 e^{-(1+\lambda)t} + a_1 e^{(1+\lambda)t}.$$

This is a perfectly good general solution we have here. Some experience (and some linear algebra based on the fact that the set of solutions for a homogeneous linear ODE is a vector space) suggests we might trade in the original exponential basis $\{e^{-\mu t}, e^{\mu t}\}$ for a more convenient one as follows: Take as X_0 the linear combination

$$X_0(t) = \frac{1}{2}e^{\mu t} + \frac{1}{2}e^{-\mu t}.$$

This is a function known as $\cosh(\mu t)$ pronounced sort of like "gosh!" and also is called the hyperbolic cosine. You can check that in this case we have

$$X_0(0) = 1 \qquad \text{and} \qquad X_0'(0) = 0.$$

Again, many properties of this "special function" are known. For example cosh is even and convex. Also,

$$X_0'(t) = \mu \sinh(\mu t) = \mu\left(\frac{1}{2}e^{\mu t} - \frac{1}{2}e^{-\mu t}\right).$$

Appearing here is the hyperbolic sine, which is odd, and if we set $X_1(t) = \sinh(\mu t)$ we get a different basis of solutions $\{X_0, X_1\}$ for the solution space of our ODE with the second basis element satisfying

$$X_1(0) = 0 \qquad \text{and} \qquad X_1'(0) = \mu.$$

This sort of change of basis is convenient. You'll notice for example that with the exponential basis we get

$$e^{\mu t}\Big|_{t=0} = 1 = e^{-\mu t}\Big|_{t=0} \qquad \text{and} \qquad \frac{d}{dt}e^{\mu t}\Big|_{t=0} = \mu = -\frac{d}{dt}e^{-\mu t}\Big|_{t=0}.$$

The convenience really comes out when we take the boundary values into account: Taking

$$x(t) = a_0X_0(t) + a_1X_1(t) = a_0\cos(\mu t) + a_1\sinh(\mu t) = a_0\cos[(1+\lambda)t] - a_1\sinh[(1+\lambda)t]$$

we need

$$a_0\cosh(0) + a_1\sinh(0) = 0 = a_0\cosh(L) - a_1\sinh(L).$$

This means $a_0 = 0$ (left equation) and $a_1\sinh(L) = 0$ (right equation). Since the lone zero of $\sinh t$ is at $t = 0$, we know $\sinh(L) \neq 0$, and $a_1 = 0$. Thus we have found the unique solution: $x(t) \equiv 0$.

This is a solution we expected to have. It's a little disappointing, but everything we have done is very determinate. There are no questions: In **CASE 1** we only get the zero or "trivial" solution.

CASE 2: $1 + \lambda = 0$.

In this case, the ODE becomes $x'' = 0$. This has general solution $x(t) = a_0 + a_1t$ with basis of solutions $\{1, t\}$. The boundary condition $x(0) = 0$ forces $a_0 = 0$ again. And then the boundary condition $x(L) = 0$ forces $a_1L = 0$ or $a_1 = 0$.

Again there is only the trivial solution.

CASE 3: $1 + \lambda = \mu^2 > 0$.

Now we get the ODE $x'' = -\mu^2x$. As mentioned, we'll get sines and cosines here. Let's recall how to do that using the "try an exponential solution" approach, which should hold for us some fond memories and some insightful observations.

Setting $x(t) = e^{\alpha t}$ yields

$$\alpha^2e^{\alpha t} = -\mu^2e^{\alpha t}.$$

Again since $e^{\alpha t} \neq 0$ we get $\alpha^2 = -\mu^2$ and $\alpha = \pm\mu i$. Here i is the complex number $\sqrt{-1}$. You may not have thought about it before, but the existence and uniqueness theory for initial value problems, and for the consideration of ODEs in general, works perfectly well for solutions $x : (0, L) \to \mathbb{C}$ which are complex valued. Thus, considering the ode $x'' + \mu^2x = 0$ as an ODE for a complex valued function, there is indeed a basis of solutions for all the solutions. The basis still has exactly two elements, and it is

$$\left\{e^{i\mu t}, e^{-i\mu t}\right\}. \tag{8}$$

These functions may be somewhat less familiar to you, but here is a nice formula that should bring them (at least somewhat) down to earth:

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

This is called Euler's formula, and it expresses a complex exponential in terms of a "standard form" complex number. Thus, our basis for all complex solutions can also be written as

$$\left\{\cos(\mu t) + i\sin(\mu t), \cos(\mu t) - i\sin(\mu t)\right\}.$$

We have taken account here of the fact that sin is odd so $\sin(-\mu t) = -\sin(\mu t)$. And also cos is even.

In any case, these are complex solutions, but there are real solutions. If we squint a little, we can see the real solutions—the ones in which we were originally interested—as linear combinations of the complex solutions we've got. This is rather like what we did in **CASE 1** exchanging the real exponentials for cosh and sinh. Here's what I get when I squint:

$$\frac{1}{2}[\cos(\mu t) + i\sin(\mu t)] + \frac{1}{2}[\cos(\mu t) - i\sin(\mu t)] = \cos(\mu t).$$

Thus $X_0(t) = \cos(\mu t)$ is a real basis solution. Also,

$$\frac{1}{2i}[\cos(\mu t) + i\sin(\mu t)] - \frac{1}{2i}[\cos(\mu t) - i\sin(\mu t)] = \sin(\mu t).$$

Thus, we can trade in the complex basis for a basis of real valued functions

$$\{\cos(\mu t), \sin(\mu t)\}.$$

These functions span exactly the same space of complex solutions as long as we use complex numbers for coefficients in our linear combinations. If we just use real coefficients however, we get a two-dimensional real subspace of real solutions. And that's nice. We can turn again to the boundary values (for the third time on this problem):

$$a_0 \cos(\mu t)\Big|_{t=0} + a_1 \sin(\mu t)\Big|_{t=0} = 0$$

means of course $a_0 = 0$, and this is looking like a story we've heard before. Finally then we consider

$$a_1 \sin(\mu t)\Big|_{t=L} = a_1 \sin(\mu L) = 0.$$

Notice that the trouble with a vanishing coefficient in our quest to solve for a_1 has become curiously reversed. If $\sin(\mu L) \neq 0$, then we do get a unique solution, but it is only the trivial solution with $x(t) \equiv 0$ again. But now the vanishing of the coefficient can lead to infinitely many solutions of a nice variety. That is, precisely when

$$\mu L = k\pi \qquad \text{for some } k \in \mathbb{N}$$

we get that $a_1 \sin(\mu L)$ solves the two point boundary value problem (7) for every $a_1 \in \mathbb{R}$. Translating back into terms of the original Sturm-Liouville parameter/eigenvalue λ we get a sequence of values

$$\lambda_k = \frac{k^2\pi^2}{L^2} - 1$$

because $1 + \lambda = \mu^2$. Some of these numbers might be negative if L is large, but eventually for k large the numbers λ_k will be positive and tend toward positive infinity. This kind of behavior captures the general situation with Sturm-Liouville problems.

2 Sturm-Liouville theory

I'm going to change notation. Instead of $x = x(t)$ for solutions of our ODEs suggesting time as the independent variable and x as who knows what, I'm going to use x suggesting space as the independent variable—though that doesn't always have to be the case, but that is the tradition in Sturm-Liouville theory. Maybe that's the notation used by Sturm and Liouville. In any case, we'll now have a function $\phi = \phi(x)$.

The general **Sturm-Liouville problem** is the following: Find all real values λ and functions $\phi \in C^2[a, b]\backslash\{0\}$ satisfying

$$\begin{cases} (p\phi')' + q\phi + \lambda\sigma\phi = 0 \quad a < x < b \\ \alpha_0\phi(a) + \beta_0\phi'(a) = 0 \\ \alpha_1\phi(b) + \beta_1\phi'(b) = 0. \end{cases} \tag{9}$$

In our problem (7) above we have p and q and σ are the constant functions with value identically 1, but in general p and q and σ can be some functions of x. We had $a = 0$ and $b = L$ with $\alpha_0 = \alpha_1 = 1$ and $\beta_0 = \beta_1 = 0$.

Here the second order linear operator $L\phi = (p\phi')' + q\phi = p\phi'' + p'\phi' + q\phi$ is called the **Sturm-Liouville operator**. The equation

$$(p\phi')' + q\phi + \lambda\sigma\phi = 0 \tag{10}$$

is called the Sturm-Liouville ODE. Note, however, it is not really just an ODE, but rather an ODE with an unknown real parameter λ. Furthermore, the equation is

$$L\phi = -\lambda\sigma\phi.$$

This looks a little bit like an eigenvalue problem from linear algebra ($A\mathbf{v} = \lambda\mathbf{v}$). For this reason the equation (10) is called a **Sturm-Liouville eigenvalue problem**. Like an eigenvalue problem from linear algebra, zero eigenvalues λ are okay, but zero eigenvectors don't count. A function ϕ along with a **Sturm-Liouville eigenvalue** λ satisfying (9) is typically called a (Sturm-Liouville) **eigenfunction**. If you find a solution (λ, ϕ) then every nonzero multiple of ϕ is also an eigenfunction.[2]

Sometimes the entire Sturm-Liouville problem (9) is also called a Sturm-Liouville eigenvalue problem.

In any case, those are pretty much the general outlines of the players in the problem and how such a problem works.

There are many theorems and related results about Sturm-Liouville problems. The result I will state below to give a general flavor of what to expect applies to Sturm-Liouville problems which are said to be **regular**, and it is somewhat important to know what to look for in order to see if a particular Sturm-Liouville problem is regular or not. Many Sturm-Liouville problems are not regular and some or all of the assertions in the theorem below still hold for many of those problems. However, there are a few aspects of a regular problem that are crucial...or should not be violated in a way that is too serious.

The basic idea is that the coefficient functions p and σ should be positive. The coefficient p is the leading coefficient in the Sturm-Liouville operator and for a regular problem one should have

$$p(x) > 0 \qquad \text{for} \qquad a \le x \le b.$$

The function σ is called the $\mathcal{L}^2$ **weight** and for a regular problem this one should also satisfy

$$\sigma(x) > 0 \qquad \text{for} \qquad a \le x \le b.$$

All three coefficient functions p, q, and σ should be continuous:

$$p, q, \sigma \in C^0[a, b].$$

Finally each of the boundary conditions should be saying something. For a regular problem it is required that

$$|\alpha_0| + |\beta_0| > 0 \qquad \text{and} \qquad |\alpha_1| + |\beta_1| > 0.$$

The coefficients α_0, α_1, β_0 and β_1 are just real constants. I will come back to the boundary conditions a little later.

[2]The notion of an eigenfunction being "zero" might deserve some comment. The exclusion $\phi \in C^2[a, b]\backslash\{0\}$ excludes the function with values constant zero. This does not mean the function ϕ does not admit certain arguments $x \in [a, b]$ for which $\phi(x) = 0$, i.e., zeros of the function, but it just means the function cannot be identically zero. One might ask if this makes sense in the sense of vectors in a vector space $C^2[a, b]$, but it does.

Here is the main result:

Theorem 1 *A regular Sturm-Liouville problem has a sequence of solutions, i.e., eigenvalue/eigenfunction pairs* (λ_j, ϕ_j)*,* $j = 1, 2, 3, \ldots$*, and the following hold:*

1. *(eigenvalue properties)*

   (a) *All eigenvalues are real numbers.*

   (b) *The eigenvalues/eigenfunctions may be ordered so that*

   $$\lambda_1 < \lambda_2 < \lambda_3 < \cdots$$

   and λ_j *tends to* $+\infty$ *as* j *tends to* ∞*:*

   $$\lim_{j\to\infty} \lambda_j = +\infty.$$

   (There is a unique smallest eigenvalue.)

2. *(basic eigenfunction properties)*

   (a) *Each eigenspace is one-dimensional:*

   $$\{\phi \in C^2[a,b] : L\phi + \lambda\sigma\phi = 0\} = \text{span}\{\phi_j\}.$$

   (b) *The eigenfunction* ϕ_j *has exactly* $j-1$ *zeros on* (a,b)*.*

3. *($\mathcal{L}^2$ eigenfunction properties)*

   (a) $\{\phi_j\}_{j=1}^{\infty}$ *is a "complete" subset of (weighted)* $\mathcal{L}^2(a,b)$ *space: This means in particular that one can write any function* $\phi \in \mathcal{L}^2(a,b)$ *as a series*[3]

   $$\phi = \sum_{j=1}^{\infty} a_j \phi_j.$$

   This is called an **eigenfunction expansion***.*

   (b) $\{\phi_j\}_{j=1}^{\infty}$ *is an "orthogonal" sequence in* $\mathcal{L}^2(a,b) \cap C^2[a,b]$*:*

   $$\int_{(a,b)} \phi_i(x)\ \phi_j(x)\ \sigma(x) = 0, \qquad i \neq j.$$

   (This is why σ *is called the weight function.)*

[3]The actual meaning of the equality here and exactly which functions are included, i.e., what is the space of $\mathcal{L}^2$ functions, require some explanation, but you can think of power series or Fourier series to perhaps get some idea. We should come back to this property in more detail later.

2.1 Brief comment on $\mathcal{L}^2(a, b)$

You may recall our discussion of regularity in which the continuous functions $C^0[a, b]$ or $C^0(a, b)$ comprise a relatively large set compared to various functions with derivatives:

$$C^0(a, b) \supset C^1(a, b) \supset C^2(a, b) \supset \cdots \supset C^\infty(a, b) \supset C^\omega(a, b)$$

with $C^\omega(a, b)$ denoting the real analytic functions on the open interval (a, b). It is difficult to make a rigourous comparison specifying the interval (a, b) or $[a, b]$ precisely, but let me write in some vague schematic fashion

$$\mathcal{L}^2 \supset C^0 \supset C^1 \supset C^2 \supset \cdots \supset C^\infty \supset C^\omega.$$

The idea this is supposed to suggest is that $\mathcal{L}^2$ is a vastly larger vector space of functions than the functions which are continuous or differentiable. On the far right with the real analytic functions one has power series expansion.

On the far left one has eigenfunction expansion which is much more general.

One other comparison/comment: $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^n$ are finite dimensional normed vector spaces with

$$|(x_1, x_2, \dots, x_n)| = \sqrt{\sum_{j=1}^n x_j^2}$$

and these are also **inner product spaces** with inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^n x_j y_j.$$

Note $|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

The spaces $C^k[a, b]$ are also normed spaces, but they are not inner product spaces.

The space

$$\mathcal{L}^2(a, b) = \left\{ \phi : \int_{(a,b)} \phi^2 < \infty \right\}$$

is a normed vector space with norm

$$\|\phi\|_{\mathcal{L}^2} = \sqrt{\int_{(a,b)} \phi^2}.$$

Also $\mathcal{L}^2(a, b)$ is an inner product space with

$$\langle \phi, \psi \rangle_{\mathcal{L}^2} = \int_{(a,b)} \phi\psi.$$

this is how we tell functions in $\mathcal{L}^2$ are orthogonal. What I have introduced above is the "standard" version of $\mathcal{L}^2$. There is a similar space called "weighted" $\mathcal{L}^2$. For weighted $\mathcal{L}^2$

$$\|\phi\|_{\mathcal{L}^2} = \sqrt{\int_{(a,b)} \phi^2 \ \sigma}$$

and

$$\langle \phi, \psi \rangle_{\mathcal{L}^2} = \int_{(a,b)} \phi\psi \ \sigma.$$

2.2 Boundary conditions

Some of the regular boundary conditions associated with a Sturm-Liouville problem have names. The condition

$$\phi(a) = \phi(b) = 0$$

is called the **Dirichlet boundary condition**. The condition

$$\phi'(a) = 0 = \phi'(b)$$

is called the **Neumann boundary condition**. The conditions

$$\phi(a) = 0 = \phi'(b) \qquad \text{and/or} \qquad \phi'(a) = 0 = \phi(b)$$

are called **Robin or mixed boundary conditions**. More generally, one can consider

$$\begin{cases} \alpha_0\phi(a) + \beta_0\phi'(a) = 0 \\ \alpha_1\phi(b) + \beta_1\phi'(b) = 0 \end{cases}$$

with

$$\det\begin{pmatrix} \alpha_0 & \beta_0 \\ \alpha_1 & \beta_1 \end{pmatrix} \neq 0.$$

These are all regular conditions and they are all **homogeneous boundary conditions** (due to the zeros involved).

There are boundary conditions which are not regular but are very common and have a name. These are **periodic boundary conditions**:

$$\begin{cases} \phi(a) = \phi(b) \\ \phi'(a) = \phi'(b). \end{cases}$$

Exercise 2 *Solve the Sturm-Liouville problem associated with the operator* $L\phi = \phi''$*, with weight* $\sigma \equiv 1$ *on* $[0, L]$ *and with periodic boundary conditions.*

3 Application to the two point boundary value problem

I return now to the two point boundary value problem (1)

$$\begin{cases} x'' + t^2x = t^3, \quad 0 \le t \le L \\ x(0) = x_0, \\ x(L) = x_L \end{cases} \tag{11}$$

and

$$\begin{cases} x_h'' + t^2x_h = 0, \qquad 0 \le t \le L \\ x_h(0) = x_0, \\ x_h(L) = x_L - L. \end{cases} \tag{12}$$

Looking at this problem, I am motivated to consider the Sturm-Liouville ODE

$$\phi'' + x^2\phi + \lambda\sigma\phi = 0 \tag{13}$$

involving the Sturm-Liouville operator

$$L\phi = \phi'' + x^2\phi.$$

This tells me the leading coefficient $p \equiv 1$ and a middle coefficient $q(x) = x^2$. I don't know the weight function $\sigma = \sigma(x)$ at this point.

Notice that a function ϕ satisfying $L\phi = 0$ like X_1 with a zero at some value $L = z$ actually satisfies a two point boundary value problem with homogeneous Dirichlet boundary values:

$$X_1(0) = 0 = X_1(L).$$

The idea here is that I want to somehow "normalize" so that a natural version of (13) arises. Here is an idea: Let $\phi(x) = X_1(zx)$ where z is a positive zero of X_1. Then $\phi''(x) = z^2 X_1''(zx)$, and

$$L\phi(x) = z^2 X_1''(zx) + x^2 X_1(zx).$$

But remember X_1 is a solution of the homogeneous ODE $x_h'' + t^2 x_h = 0$. This implies

$$X_1''(zx) + (zx)^2 X_1(zx) = 0 \qquad \text{or} \qquad X_1''(zx) = -(zx)^2 X_1(zx).$$

From this I find

$$L\phi(x) = -z^4 x^2 X_1(zx) + x^2 X_1(zx) = (1 - z^4)x^2\phi$$

or

$$L\phi + \lambda x^2\phi = \phi'' + x^2\phi + \lambda x^2\phi = 0$$

where $\lambda = z^4 - 1$. This strongly suggests (or at least somewhat suggests) consideration of the Sturm-Liouville problem

$$\begin{cases} \phi'' + x^2\phi + \lambda x^2\phi = 0, & 0 < x < 1 \\ \phi(0) = 0 \\ \phi(1) = 0 \end{cases} \tag{14}$$

may tell me interesting things about my original problem.

In this problem, the leading coefficient is $p \equiv 1$, the middle coefficient is $q(x) = x^2$, and the weight function is $\sigma(x) = x^2$. I have Dirichlet boundary values on $[0, 1]$.

Everything looks good with this problem. It is almost regular. The only failing is that the weight $\sigma(x) = x^2$ vanishes at $x = 0$. Let me assume for a moment however that the conclusions of the Sturm-Liouville theory/theorem hold, then I expect a sequence of eigenvalues

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots.$$

If this is true, this tells me something rather important. There is a potential relation

$$z^4 = \lambda + 1 \tag{15}$$

between the positive zeros z of X_1 and the Sturm-Liouville eigenvalues of (14). In fact we know every positive zero does definitely correspond to a Sturm-Liouville eigenvalue for this problem. Furthermore, assuming the eigenvalues are increasing and become arbitrarily large, we might imagine, at least at first, there are some eigenvalues $\lambda < -1$ for which (15) has no real roots, but eventually there will be a first eigenvalue λ_k with

$$\lambda_k + 1 > 0.$$

Then we can write

$$z_1 = \sqrt[4]{\lambda_k + 1} \qquad \text{and} \qquad z_{1+j} = \sqrt[4]{\lambda_{k+j} + 1}$$

for every $j = 1, 2, 3, \ldots$.

If we go further and assume each eigenfunction ϕ_j corresponding to λ_j has exactly $j - 1$ zeros on $(0, 1)$, then the first eigenfunction ϕ_1 will not vanish on $(0, 1)$ at all, and we can assume $\phi_1 > 0$. The natural question is the following: Can we find or say something about ϕ_1?

Here is a key idea: The eigenfunction ϕ_1 will (more or less) be a solution to the **IVP**

$$\begin{cases} \phi'' + x^2\phi + \lambda x^2\phi = 0, & 0 < x < 1 \\ \phi(0) = 0 \\ \phi'(0) = 1. \end{cases} \tag{16}$$

Why is this true, and what do I mean by "more or less?" I've already mentioned that ϕ_1 is not supposed to vanish. If we can find a non-vanishing solution, then that solution is either all positive valued or all negative valued. If we end up with $\phi_1(x) < 0$ for $0 < x < 1$, then $-\phi_1$ is also a solution with $-\phi_1(x) > 0$ for $0 < x < 1$.

Assuming now $\phi_1 > 0$ we know, $\phi_1'(0) \geq 0$, but if $\phi_1'(0) = 0$, then ϕ_1 is a solution of the IVP

$$\begin{cases} \phi'' + x^2\phi + \lambda x^2\phi = 0, & 0 < x < 1 \\ \phi(0) = 0 \\ \phi'(0) = 0. \end{cases}$$

But this IVP is subject to the existence and uniqueness theorem for initial value problems which tells us, since $\phi \equiv 0$ is a solution that ϕ_1 must also be this constant zero solution. But that would rule ϕ_1 out as an eigenfunction, or alternatively we have already noted $\phi_1(x) > 0$ for $0 < x < 1$, so this is not what happens. It must be the case that $\phi_1'(0) > 0$.

If $\phi_1'(0) \neq 1$, then $\phi_1/\phi_1'(0)$ is definitely a solution of (16) and also an eigenfunction. We have established that there is a first eigenfunction ϕ_1 which is non-vanishing, positive and is a solution of the initial value problem (16) on the interval $[0, 1]$. This makes ϕ_1 and consequently λ_1 relatively easy to find in several ways.

As a first approach we can go back and find the unique solution of (16) as a function of λ in terms of a power series. This series will have coefficients in terms of λ. Shall we do that?

The ODE is $\phi'' + (1 + \lambda)x^2\phi = 0$, and we try

$$\phi(x) = \sum_{j=0}^{\infty} a_j x^j.$$

We note that

$$\phi'(x) = \sum_{j=1}^{\infty} j a_j x^{j-1}$$

and

$$\phi''(x) + (1 + \lambda)x^2\phi(x) = \sum_{j=2}^{\infty} j(j-1)a_j x^{j-2} + \sum_{j=0}^{\infty}(1 + \lambda)a_j x^{j+2} = 0.$$

After isolating the first few terms and shifting some indices we have

$$2a_2 + 6a_3 x + \sum_{j=2}^{\infty}[(j+2)(j+1)a_{j+2} + (1 + \lambda)a_{j-2}]x^j = 0.$$

Equating coefficients we see $a_2 = 0 = a_3$ and we have the recurrence relation

$$a_{j+2} = -(1 + \lambda)\frac{1}{(j+2)(j+1)}\, a_{j-2} \qquad \text{for} \qquad j = 2, 3, 4, \ldots.$$

Shifting indices again

$$a_{j+4} = -\frac{1 + \lambda}{(j+4)(j+3)}\, a_j \qquad \text{for} \qquad j = 0, 1, 2, 3, \ldots.$$

By the initial condition $\phi(0) = 0$, we know $a_0 = 0$, and by the recursion relation $a_4 = 0 = a_8 = a_{12} = a_{4k}$ for $k = 0, 1, 2, 3, \ldots$.

By the initial condition $\phi'(0) = 1$, we get $a_1 = 1$. The recursion relation then gives $a_5 = -(1+\lambda)a_1/20$ and

$$a_{4k+1} = \frac{(-1)^k(1+\lambda)^k}{(4k+1)4k(4k-3)4(k+1)\cdots(5)(4)}\, a_1 = \frac{(-1)^k(1+\lambda)^k}{4^k k! \prod_{m=1}^{k}[4(k-m)+5]}.$$

The rest of the coefficients are zero, and

$$\begin{aligned}\phi(x) &= x + \sum_{j=1}^{\infty} \frac{(-1)^k(1+\lambda)^k}{4^k k! \prod_{m=1}^{k}[4(k-m)+5]} x^{4k+1} \\ &= x + \sum_{j=1}^{\infty} \beta_k (1+\lambda)^k x^{4k+1}.\end{aligned}$$

In particular, $\phi(1)$ as a function of λ is given by

$$\rho(\lambda) = 1 + \sum_{j=1}^{\infty} \beta_k (1+\lambda)^k.$$

We may note something useful at this point based on the power series. Recall that β_k is alternating with sign $(-1)^k$. If $\lambda < -1$ so that $1+\lambda < 0$, then $(1+\lambda)$ is also alternating with sign $(-1)^k$. Thus, the product is positive and $\rho(\lambda) > 1$. There can be no eigenvalues with $\rho < -1$. In fact $\rho(-1) = 1 > 0$, so there are no eigenvalues for (14) with $\lambda + 1 \le 0$. Every eigenvalue λ corresponds to a zero of X_1 with

$$z_j = \sqrt[4]{\lambda + 1}, \qquad j = 1, 2, 3, \ldots.$$

Plotting a partial sum

$$1 + \sum_{j=1}^{10} \beta_k (1+\lambda)^k$$

for this function for $-1 < \lambda < 45$ we obtain the plot shown in Figure 2. A first positive root λ_1 is indicated

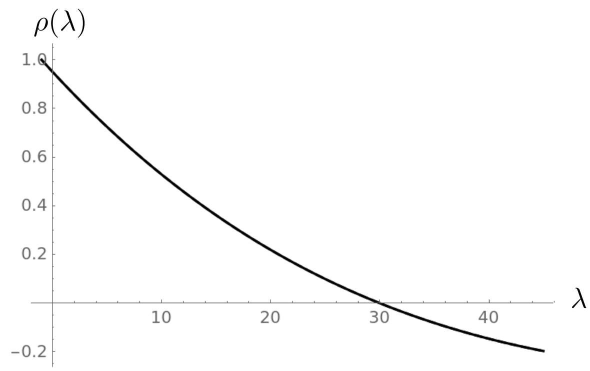

Figure 2: Plot of $\rho = \phi_1(1)$ as a function of λ from the power series for ϕ_1.

around $\lambda = 30$ and a numerical root find algorithm indicates $\lambda_1 \doteq 29.9333$ giving

$$z_1 = \sqrt[4]{\lambda_1 + 1} \doteq 2.35834$$

confirming the first value we found above.

4 ODE solver approximation and other zeros

It should not be missed that the general situation, though we do not necessarily know the precise locations of the zeros, can be guessed or predicted. Assuming the Sturm-Liouville problem (14) behaves like a regular Sturm-Liouville problem, there is a sequence of eigenvalues

$$\lambda_1 \doteq 29.9333 < \lambda_2 < \lambda_3 < \cdots$$

with the j-th positive zero of X_1 given by the formula

$$z_j = \sqrt[4]{\lambda_j + 1}.$$

I'm going to compute the value $\lambda_1 \doteq 29.333$ using a different method and indicate how to nicely compute many more zeros. Mathematical software has very nice (adaptive fourth and fifth order Runge-Kutta methods) implementations for solving initial value problems for ordinary differential equations. One can count on much greater accuracy than when using partial sums for power series "by hand."

It is also relatively easy to obtain an approximation for the solution of an IVP like (16) with the numerical solution given as a function of λ. Figure 3 shows plots of the solutions on $[0, 1]$ for various values of λ. Applying a root find algorithm to the equation $\phi(1; \lambda) = 0$ to determine λ_1 results in an extremely

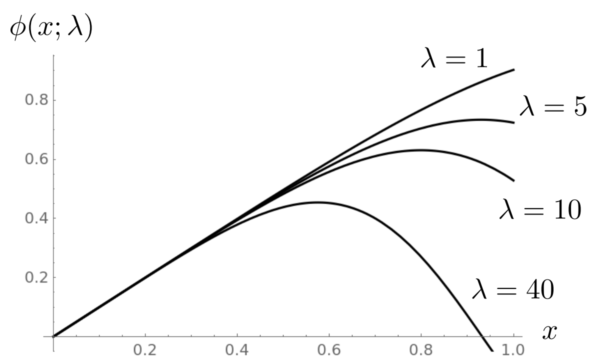

Figure 3: Plots of solutions $\phi = \phi(x)$ of (16) for $0 \leq x \leq 1$ and λ taking various values near λ_1. These are computed with an ordinary differential equations numerical "solver" for initial value problems. It is clear the first eigenvalue occurs for λ somewhere between $\lambda = 10$ and $\lambda = 40$.

accurate value $\lambda_1 \doteq 29.333$ as we have computed before.

It is relatively easy to repeat this procedure looking for solutions with exactly one zero on $(0, 1)$, that is for the second Sturm-Liouville eigenfunction. A series of solutions of the initial value problem are again shown on the left in Figure 4. Solving the same equation $\rho(\lambda) = \phi(1) = 0$ for λ between $\lambda = 130$ and $\lambda = 160$ gives the value $\lambda_2 \doteq 138.53$ and the second zero at $z_2 = \sqrt[4]{\lambda_2 + 1} \doteq 3.4369$.

In order to show the relative power of the method, I've computed the sixteenth zero of X_1. See Figure 5. I suspect this sixteenth zero would be pretty difficult to calculate accurately using power series at $t = 0$ for $X_1 = X_1(t)$ or at $x = 0$ for the solution of (16) with coefficients given as a function of λ.

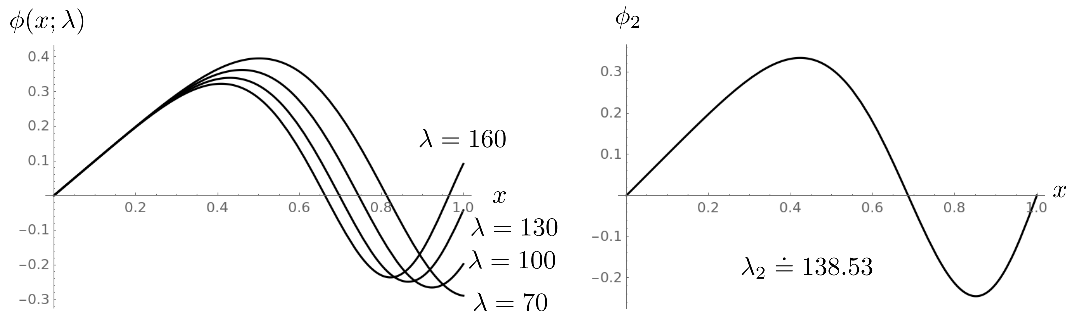

Figure 4: Plots of solutions $\phi = \phi(x)$ of (16) for $0 \leq x \leq 1$ and λ taking various values. The second eigenvalue occurs with $\lambda_2 \doteq 138.53$

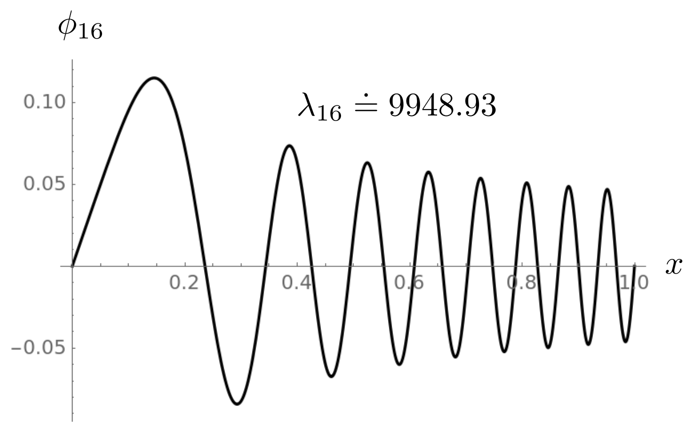

Figure 5: The sixteenth eigenfunction corresponding to $\lambda_{16} \doteq 9948.93$ and the sixteenth zero of X_1 at $z_{16} = \sqrt[4]{\lambda_{16}} \doteq 9.98746$.