An Introduction to Partial Differential Equations in the Undergraduate Curriculum

Katherine Socha & David Arnold

LECTURE 8
Sturm-Liouville Theory—Part I

1.1. Outline of Lecture

- Example of a non-homogeneous boundary value problem
- The Ten-Step Program

1.2. The heat equation with a radiation boundary condition

In this lecture, we consider the initial boundary value problem (IBVP) with nonhomogeneous boundary data,

$$\begin{aligned} u_t = K u_{xx} &: \quad 0 < x < L, \quad t > 0 \\ \left.\begin{aligned} u(0,t) &= T_1 \\ u_x(L,t) &= -h\big[u(L,t) - T_2\big] \end{aligned}\right\} &: \quad t > 0, \\ u(x,0) = T_3 &: \quad 0 < x < L, \end{aligned}$$

where h, T_1, T_2, T_3, and K are (strictly) positive constants. We have seen all of these expressions except for the boundary condition at $x = L$: this is called a **radiation** or **Robin** condition. It describes how heat radiates radiates from the end into the surrounding medium. The general form of a homogeneous Robin condition at $x = a$ is

$$-k u_x(a,t) + \beta u(a,t) = 0 \quad t > 0$$

where k is the thermal conductivity of the bar (introduced in Lecture 2 (Bernoff)), a positive constant. If $\beta < 0$ (and, of course, $u(a,t) > 0$), then heat flows into the bar, an *absorption* condition. If, as in our IBVP, $\beta > 0$, then heat flows out of the bar, a radiation condition. Rewriting our IBVP so that the radiation condition is more readable, we see

$$\begin{aligned} u_t = Ku_{xx} &: \quad 0 < x < L, \quad t > 0, \\ \left.\begin{array}{l} u(0,t) = T_1 \\ u_x(L,t) + hu(L,t) = hT_2 \end{array}\right\} &: \quad t > 0, \\ u(x,0) = T_3 &: \quad 0 < x < L. \end{aligned} \tag{1.1}$$

A straightforward application of the separation of variables technique that worked so well for the heat equation with homogeneous Dirichlet or Neumann boundary data leads to a hard problem (see Problem 1), so we will try to transform the problem into one we actually *can* handle. Motivated by the examples shown in Lectures 2, 5 (Wittwer), and 8 (Vajiac) we will begin by seeking a steady-state solution.

1.2.1. The steady state solution

Finding a steady state solution means that we seek a solution that is independent of time. That is, find a function $u_s(x)$ that satisfies $u_t = Ku_{xx}$ on $0 < x < L$, together with the boundary data $u_s(0) = T_1$ and $u_s'(L) + hu_s(L) = hT_2$. Because $\{d/dt\}(u_s(x)) = 0$, we have

$$u_s''(x) = 0.$$

Integrate twice to find $u_s(x) = ax + b$. Requiring that this line satisfies the boundary data allows us to compute the integration constants a and b. After this work, we find the steady-state solution:

$$u_s(x) = \frac{h(T_2 - T_1)}{1 + hL} x + T_1. \tag{1.2}$$

Exercise 1. Do the algebra to prove that this is the correct form of the steady solution.

1.2.2. Homogeneous boundary conditions

Now we transform our problem by setting $v(x,t) = u(x,t) - u_s(x)$. A straightforward calculation (Problem 2) demonstrates that v must satisfy the new IBVP

$$v_t = Kv_{xx} : \quad 0 < x < L, \quad t > 0$$

$$\left.\begin{array}{l} v(0,t) = 0 \\ v_x(L,t) + hv(L,t) = 0 \end{array}\right\} : \quad t > 0,$$

$$v(x,0) = T_3 - u_s(x) : \quad 0 < x < L.$$

How is this an improvement? Notice that both boundary conditions are now homogeneous and it is only the initial condition that varies in x. Guided by the previous lectures, we expect that this non-constant $v(x,0)$ will not pose any difficulties.

The goal now is to find a solution v by separation of variables, and then to find a solution $u(x,t) = v(x,t) + u_s(x)$ of (1.1).

1.2.3. Separation of variables

Applying the standard separation of variables argument leads us to two ordinary differential equations. For $v(x,t) = X(x)T(t)$, we find, as in Lecture 5,

$$\frac{T'(t)}{KT(t)} = \frac{X''(x)}{X(x)} \tag{1.3}$$

which we set equal to $-\lambda$ to find

$$T'(t) = -\lambda K T(t)$$
$$X''(x) = -\lambda X(x).$$

Again, the general solution for the time dependence has the form

$$T(t) = Ce^{-\lambda K t}.$$

1.2.4. A Sturm-Liouville Equation

From this point forth, we seek only nontrivial (nonzero) solutions $v(x,t)$.

From the separation of variables argument $T(t) = Ce^{-\lambda K t}$. Note that T is nonzero for all t. We now must find nontrivial solutions of the equation

$$-X'' = \lambda X. \tag{1.4}$$

Recall the boundary conditions, for all $t > 0$

$$v(0,t) = 0 \qquad \text{and} \qquad v_x(L,t) + hv(L,t) = 0.$$

Substituting $v(x,t) = X(x)T(t)$ in each of these conditions provides

$$X(0)T(t) = 0 \qquad \text{and} \qquad X'(L)T(t) + hX(L)T(t) = 0$$

for all $t > 0$. Since $T(t)$ is nonzero, divide these equations by $T(t)$ to capture

$$X(0) = 0 \qquad \text{and} \qquad X'(L) + hX(L) = 0.$$

Putting all of this together leads to a **Sturm-Liouville problem**,

$$-X'' = \lambda X, \quad X(0) = 0, \quad X'(L) + hX(L) = 0. \tag{1.5}$$

If a nontrivial solution of equation (1.5) exists, then the constant λ is called an **eigenvalue** and the solution X is called its associated **eigenfunction**.

Postponing until Lecture 9 a discussion of the possibility of complex eigenvalues, we consider three distinct cases:

i. $\lambda < 0$,
ii. $\lambda = 0$, or
iii. $\lambda > 0$.

Let's first consider the possibility that λ is negative.

1.2.4.1. *Case i:* $\lambda < 0$. A purely analytical approach is described in Appendix 1.5, but here we introduce a useful and more general approach called an **energy argument**. First, multiply both sides of $-X'' = \lambda X$ by X and integrate both sides with respect to x on the interval $[0, L]$:

$$-\int_0^L X(x)X''(x)\,dx = \lambda \int_0^L X^2(x)\,dx.$$

Then integrate by parts to change the form of the left-hand side of this equation:

$$-X(x)X'(x)\Big|_0^L + \int_0^L \Big[X'(x)\Big]^2 dx = \lambda \int_0^L \Big[X(x)\Big]^2 dx. \tag{1.6}$$

However,

$$-X(x)X'(x)\Big|_0^L = -X(L)X'(L) + X(0)X'(0).$$

The boundary conditions require $X(0) = 0$ and $X'(L) = -hX(L)$. Thus,

$$-X(x)X'(x)\Big|_0^L = -X(L)\big[-hX(L)\big] + 0 \cdot X'(0) = hX(L)^2 \geq 0.$$

Now equation (1.6) becomes

$$\frac{hX^2(L) + \int_0^L (X'(x))^2\,dx}{\int_0^L X^2(x)dx} = \lambda \tag{1.7}$$

from which it follows that $\lambda \geq 0$.

1.2.4.2. *Case ii:* $\lambda = 0$. If $\lambda = 0$, then $-X'' = \lambda X$ becomes $X'' = 0$, which has general solution

$$X(x) = A + Bx.$$

The boundary condition $X(0) = 0$ implies that $A = 0$, so the solution is now $X(x) = Bx$, with derivative $X'(x) = B$. The second boundary condition provides

$$0 = X'(L) + hX(L) = B + h(BL) = B(1 + hL).$$

Because both $h > 0$ and $L > 0$, it follows that $B = 0$. Hence, $X(x) = 0$ for all x. Since $\lambda = 0$ has only the trivial solution, we conclude that $\lambda = 0$ is not an eigenvalue.

1.2.4.3. *Case iii:* $\lambda > 0$. Since $\lambda > 0$, we now know

$$X(x) = A\cos\left(x\sqrt{\lambda}\right) + B\sin\left(x\sqrt{\lambda}\right).$$

Applying the first boundary condition gives

$$0 = X(0) = A\cos 0 + B\sin 0 = A.$$

Thus, $X(x) = B\sin\left(x\sqrt{\lambda}\right)$. Notice that we must have $B \neq 0$ to obtain a nontrivial solution. The second boundary condition requires that

$$0 = X'(L) + hX(L) = B\sqrt{\lambda}\cos\left(L\sqrt{\lambda}\right) + hB\sin\left(L\sqrt{\lambda}\right).$$

Dividing through by $B\cos\left(L\sqrt{\lambda}\right)$ shows

$$0 = \sqrt{\lambda} + h\tan\left(L\sqrt{\lambda}\right).$$

To obtain nontrivial solutions $X(x)$, we must hope that this trigonometric equation has at least one solution for λ. This is a perfect opportunity to use a computer application such as Maple. In fact, we will see that there are infinitely many solutions! Write these eigenvalues of the BVP as λ_j such that

$$-\tan\left(L\sqrt{\lambda_j}\right) = \frac{\sqrt{\lambda_j}}{h}, \tag{1.8}$$

with corresponding eigenfunctions

$$X_j(x) = \sin\left(x\sqrt{\lambda_j}\right).$$

1.2.5. Orthogonality

The eigenvalues of the Sturm-Liouville problem (1.5) are the positive solutions of

$$-\tan\left(L\sqrt{\lambda}\right) = \frac{\sqrt{\lambda}}{h} \tag{1.9}$$

and have associated eigenfunctions

$$X_j(x) = \sin\left(x\sqrt{\lambda_j}\,\right). \tag{1.10}$$

In the Fourier theory developed in Lecture 6, the eigenfunctions had the remarkable property of orthogonality. Because we can find only approximate (numerical) solutions of (1.9), it will be difficult to show that (for $j \neq k$)

$$\int_0^L \sin\left(x\sqrt{\lambda_j}\right)\,\sin\left(x\sqrt{\lambda_k}\right)dx = 0,$$

for our eigenfunctions (1.10) . Returning to the differential equation itself may provide more insight into the nature of these eigenfunctions.

Suppose that $\lambda_1 \neq \lambda_2$ are two distinct eigenvalues with associated eigenfunctions X_1 and X_2 that both satisfy the Sturm-Liouville problem

$$-X'' = \lambda X, \qquad X(0) = 0, \qquad X'(L) + hX(L) = 0.$$

Thus,

$$\begin{aligned} -X_1'' &= \lambda_1 X_1 \\ -X_2'' &= \lambda_2 X_2. \end{aligned}$$

Multiply the first equation by X_2 and the second by X_1 and subtract:

$$-X_2X_1'' + X_1X_2'' = (\lambda_1 - \lambda_2)X_1X_2.$$

Integrate this expression with respect to x on the interval $[0, L]$ to find

$$\int_0^L \left(-X_2X_1'' + X_1X_2''\right)dx = (\lambda_1 - \lambda_2)\int_0^L X_1X_2\,dx.$$

Note that the integrand on the left is exact, so

$$\int_0^L \frac{d}{dx}\Big[X_1X_2' - X_2X_1'\Big]\,dx = (\lambda_1 - \lambda_2)\int_0^L X_1X_2\,dx.$$

Thus,

$$X_1X_2' - X_2X_1'\Big|_0^L = (\lambda_1 - \lambda_2)\int_0^L X_1X_2\,dx. \tag{1.11}$$

Because $X_1(0) = 0$ and $X_2(0) = 0$,

$$X_1 X_2' - X_2 X_1' \Big|_0^L = X_1(L)X_2'(L) - X_2(L)X_1'(L). \tag{1.12}$$

The next step is to argue that the right-hand side of equation (1.12) also equals zero.

Both X_1 and X_2 must satisfy the second boundary condition of the Sturm-Liouville equation. Thus,

$$X_1'(L) + hX_1(L) = 0$$
$$X_2'(L) + hX_2(L) = 0.$$

This system can be written in matrix form:

$$\begin{bmatrix} X_1'(L) & X_1(L) \\ X_2'(L) & X_2(L) \end{bmatrix} \begin{bmatrix} 1 \\ h \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The fact that this matrix equation has a nontrivial solution demands that the determinant of the coefficient matrix equal zero. Thus,

$$0 = \begin{vmatrix} X_1'(L) & X_1(L) \\ X_2'(L) & X_2(L) \end{vmatrix} = X_1'(L)X_2(L) - X_2'(L)X_1(L).$$

This last result, along with equations (1.11) and (1.12), provide

$$0 = (\lambda_1 - \lambda_2) \int_0^L X_1 X_2 \, dx.$$

Finally, because λ_1 and λ_2 are distinct eigenvalues,

$$0 = \int_0^L X_1 X_2 \, dx,$$

and X_1 and X_2 are orthogonal.

1.2.6. A formal solution and orthogonality

We may construct a formal solution for v by superposition:

$$v(x,t) = \sum_{j=1}^{\infty} C_j e^{-\lambda_j K t} \sin\left(x\sqrt{\lambda_j}\,\right) \tag{1.13}$$

We want this formal solution to satisfy the initial condition. This requires that

$$T_3 - u_s(x) = v(x,0) = \sum_{j=1}^{\infty} C_j \sin(x\sqrt{\lambda_j}). \tag{1.14}$$

How do we know that such C_j exist?

We have shown that the eigenfunctions

$$\left\{ \sin\left(x\sqrt{\lambda_j}\,\right) \right\}_{j=1}^{\infty} \tag{1.15}$$

are pairwise orthogonal. (Refer to Problem 3 in which you will demonstrate via Maple that these functions are orthogonal.) Further, this set forms a basis for L^2 in which we may write our linear function $v(x,0)$ as a linear combination of elements of (1.15). (This idea will be explored further in Lecture 9.)

We may obtain the "Fourier" coefficients C_j just as the Fourier coefficients were constructed in Lecture 6 (Aarao). Multiply both sides of (1.14) by $\sin(x\sqrt{\lambda_k})$ and integrate over the interval $[0, L]$. This gives

$$\int_0^L [T_3 - u_s(x)] \sin\left(x\sqrt{\lambda_k}\right) dx = \int_0^L \sum_{j=1}^{\infty} C_j \sin\left(x\sqrt{\lambda_k}\right) \sin\left(x\sqrt{\lambda_j}\right) dx.$$

Assuming uniform convergence, we may interchange the order of integration and summation. By orthogonality, the right-hand side is non-zero only when $j = k$. Hence,

$$\int_0^L [T_3 - u_s(x)] \sin\left(x\sqrt{\lambda_k}\right) dx = \int_0^L C_k \sin^2\left(x\sqrt{\lambda_k}\right) dx. \tag{1.16}$$

We may now write

$$C_k = \frac{\int_0^L [T_3 - u_s(x)] \sin\left(x\sqrt{\lambda_k}\right) dx}{\int_0^L \sin^2\left(x\sqrt{\lambda_k}\right) dx} \tag{1.17}$$

for integers $k \geq 1$. [1]

Now we have completely determined the series solution for v, and hence for $u(x,t) = v(x,t) + u_s(x)$:

$$u(x,t) = \sum_{j=1}^{\infty} C_j e^{-\lambda_j K t} \sin\left(x\sqrt{\lambda_j}\right) + \frac{h(T_2 - T_1)}{1 + hL} x + T_1 \tag{1.18}$$

with (λ_j, C_j) solutions of equations (1.8) and (1.17), respectively.

[1] Using the neat inner product and norm notations introduced in Lecture 6, we may write equation (1.17) above as

$$C_k = \frac{\left(v(x,0), \sin\left(x\sqrt{\lambda_k}\right)\right)}{\|\sin\left(x\sqrt{\lambda_k}\right)\|^2}.$$

1.2.7. Convergence

How do we know that the formal series solution (1.18) converges uniformly on $[0, L]$? Further, the series formulations of u_t, u_x, u_{xx} obtained through term-by-term differentiation also ought to converge uniformly. After all, in order to claim that a certain function u satisfies the IBVP, it had better be true that u is actually a function, differentiable once in t and twice in x!

Lemma 1.19. *For all $j \geq 1$, the coefficients C_j given by (1.17) satisfy $|C_j| \leq M$ for some constant M.*

Exercise 2. Prove the lemma.

This claim allows us to prove that all of the series formulations for u, u_t, u_x, and u_{xx} are uniformly convergent on $[0, L]$.

1.2.8. Long-term behavior (Asymptotics)

Finally, what happens to the series solution for $u(x, t)$ as $t \to \infty$? Since $\lambda_j K > 0$ we can use the Lemma to show that

$$\left| C_j e^{-\lambda_j K t} \sin\left(x\sqrt{\lambda_j}\right) \right| \leq |C_j| e^{-\lambda_j K t} \leq M e^{-\lambda_j K t} \to 0 \tag{1.20}$$

as $t \to \infty$. By uniform convergence, then,

$$\sum_{j=1}^{\infty} C_j e^{-\lambda_j K t} \sin\left(x\sqrt{\lambda_j}\right) \longrightarrow 0 \quad \text{as } t \to \infty. \tag{1.21}$$

Thus, $u(x, t) \to u_s(x)$, the steady state solution!

1.3. Summary: The Ten-Step Program

At this point we review the procedure discussed in the previous pages.

1. Find a steady-state solution u_s.
2. Change variables by $v = u - u_s$, transforming our IBVP with inhomogeneous boundary conditions to an IBVP with homogeneous boundary conditions.
3. Apply separation of variables, using $v(x, t) = X(x)T(t)$ to obtain two ODEs, in x and in t, from
$$\frac{X''(x)}{X(x)} = \frac{T'(t)}{kT(t)} = -\lambda.$$
4. Solve for the t-dependence: $T(t) = Ce^{-\lambda K t}$.
5. Use an energy argument to determine sign of λ.

6. Solve the x-ODE to determine pairs (λ_j, X_j).
7. Create a formal solution $v(x,t)$ by superposition and use orthogonality to determine the series coefficients.
8. Determine a formal solution for $u = v + u_s$.
9. Prove that the series expressions for u, u_t, u_x, and u_{xx} converge uniformly. Thus, u in series form is genuinely a solution.
10. Study long-term behavior. Show that

$$\lim_{t\to\infty} u(x,t) = u_s(x).$$

1.4. Challenge Problems for Lecture 8

Problem 1. Use separation of variables to try to solve the IBVP

$$\begin{aligned} u_t = Ku_{xx} :&\quad 0 < x < L, \quad t > 0 \\ \left.\begin{aligned} u(0,t) &= T_1 \\ u_x(L,t) + hu(L,t) &= hT_2 \end{aligned}\right\} :&\quad t > 0, \\ u(x,0) = T_3 :&\quad 0 < x < L. \end{aligned}$$

(Hint: it won't be pretty. Where do you run into trouble?)

Problem 2. Show that $v(x,t) = u(x,t) - u_s(x)$ satisfies the IBVP

$$\begin{aligned} v_t = Kv_{xx} :&\quad 0 < x < L, \quad t > 0 \\ \left.\begin{aligned} v(0,t) &= 0 \\ v_x(L,t) + hv(L,t) &= 0 \end{aligned}\right\} :&\quad t > 0, \\ v(x,0) = T_3 - u_s(x) :&\quad 0 < x < L. \end{aligned}$$

when $u(x,t)$ satisifies

$$\begin{aligned} u_t = Ku_{xx} :&\quad 0 < x < L, \quad t > 0 \\ \left.\begin{aligned} u(0,t) &= T_1 \\ u_x(L,t) + hu(L,t) &= hT_2 \end{aligned}\right\} :&\quad t > 0, \\ u(x,0) = T_3 :&\quad 0 < x < L. \end{aligned}$$

and $u_s(x)$ is its steady state solution.

Problem 3. Use Maple to demonstrate that functions $\sin\left(x\sqrt{\lambda_j}\right)$ are orthogonal when λ_j are the solutions of the transcendental equation (1.8).

Problem 4. (Logan, p. 131) Consider a large, circular, tubular ring of circumference $2L$ that contains a chemical of concentration $c(x,t)$ dissolved in water. Let x be the arc-length parameter with $0 < x < 2L$.

If the concentration of the chemical is initially given by $c_0(x)$, then $c(x,t)$ satisfies the IBVP

$$\begin{aligned} c_t = Dc_{xx} :&\quad 0 < x < 2L, \quad t > 0 \\ \left.\begin{aligned} c(0,t) &= c(2L,t) \\ c_x(0,t) &= c_x(2L,t) \end{aligned}\right\} :&\quad t > 0, \\ c(x,0) = f(x) :&\quad 0 < x < 2L. \end{aligned}$$

These boundary conditions are called *periodic boundary conditions*, and D is the diffusion constant. Apply the separation of variables method and show that the associated Sturm-Liouville problem has eigenvalues $\lambda_n = (n\pi/L)^2$ for $n = 0, 1, 2, \ldots$ and eigenfunctions $X_n(x) = A_n \cos(n\pi x/L) + B_n \sin(n\pi x/L)$ for $n = 1, 2, \ldots$. Show that the concentration is given by

$$c(x,t) = \frac{A_0}{2} + \sum_{n=0}^{\infty} \left(A_n \cos(n\pi x/L) + B_n \sin(n\pi x/L)\right) e^{-n^2\pi^2 Dt/L^2}$$

and find the formulae for the A_n and B_n.

1.5. Appendix

Here is the analytical argument that the Sturm-Liouville problem

$$-X'' = \lambda X, \quad X(0) = 0, \quad X'(L) + hX(L) = 0 \tag{1.22}$$

has only nonnegative eigenvalues. Suppose that $\lambda < 0$ and, so, write $\lambda = -\omega^2$, where $\omega > 0$. Then equation (1.22) becomes

$$X'' - \omega^2 X = 0.$$

It is easily seen that the general solution is[2]

$$X = A \cosh \omega x + B \sinh \omega x.$$

The boundary condition $X(0) = 0$ forces

$$0 = X(0) = A \cosh 0 + B \sinh 0 = A,$$

and $X(x) = B \sinh \omega x$, which has derivative $X'(x) = B\omega \cosh \omega x$. The second boundary condition now provides

$$0 = X'(L) + hX(L) = B\omega \cosh \omega L + hB \sinh \omega L.$$

Dividing by B,

$$0 = \omega \cosh \omega L + h \sinh \omega L,$$

[2]The hyperbolic sine and cosine are defined as follows: $\cosh x = (e^x + e^{-x})/2$ and $\sinh x = (e^x - e^{-x})/2$.

or equivalently,

$$\tanh \omega L = -\frac{\omega}{h}.$$

It is not difficult to show that the only solution of this equation is $\omega = 0$, contradicting our assumption that $\omega > 0$.

Therefore, it cannot be the case that λ is negative. That is, λ must be nonnegative ($\lambda \geq 0$).