MATH343 Project Report
Solving Heat Equations Using Numerical Methods

Georgia Corbett, Warren Wang

December 11, 2023

1 Introduction

Modeling of physical systems in engineering and physical sciences often require the usage of differential equations. They quickly become complicated enough to no longer be approachable analytically to derive a particular solution. This is where numerical methods come into play. Generally, there are two classes of differential equation problems: initial value problems and boundary value problems. In the context of this project, we will focus on boundary value problems (BVP). BVPs impose a set of boundary conditions on the dependent variable. This means the solution should be equal to some set of predefined values at some distinct set of points for the independent variable(s). This is different from initial value problems (IVP) in that IVPs require the solution to be equal to some values at the same dependent variable points at different degrees of differentiation.

One can imagine that it would be very useful to solve BVPs in general, because they can be used to model many real world situations such as the stress of high winds against the circular edge of a radio telescope. We may be interested in knowing how much stress is accumulated on the fringes of the dishbowl, which provides some set of boundary conditions. Moreover, the point is to demonstrate the usefulness and generality of BVPs.

Our project is motivated by the 2-dimensional Heat Equation for steady conduction given by:

$$\frac{\partial u}{\partial t} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)$$

As the 2-dimensional case is quite complicated, we turn to the 1-dimensional metal rod with certain boundary conditions. Following common conventions, we will define $u(x,t)$ as the temperature of the rod at position x at time t which must satisfy the one-dimensional heat equation:

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$$

assuming ideal conditions [Dai14]. This equation is also quite complicated so we will first start with a simpler boundary value problem. Namely, we wish to solve

$$\frac{\partial u^2}{\partial x^2} = f(x) \tag{1}$$

for some function f and boundary conditions $u(1) = a$ and $u(n) = b$.

2 Finite Element Method

Despite the simplicity of the 1-dimensional differential equations, they are often unsolvable using conventional differential equation methods. Instead, we turn to the **Finite Element Method** to solve this differential equation. The main idea of the Finite Element Method is to split a complicated system into finite simpler components that are easier to compute. Now the question becomes *How do we find the finite simpler components?* It turns out to be much more complicated that just discretizing the interval we are integrating over. The first step is to find the **weak form** of the boundary problem.

2.1 Strong form and Weak form

A boundary-value problem has a "weak and strong" form. The strong form is the Canonical form often seen in textbooks:

$$\frac{\partial^2 u}{\partial x^2} = f(x), u(a) = v, u(b) = w \quad \text{for } v, w \in \mathbb{R}.$$

For simplicity's sake, we will assume the boundary conditions are $u(a) = 0, u(b) = 0$. Notice the strong form includes a second partial derivative that restricts our choices for a function $u(x)$. To get around this, we can consider the weak form. The weak form uses integration by parts and ultimately lets us use functions that are not defined for higher derivatives. In Proposition on p.5 in [Hug87], we know that if u is a solution of the strong form, it is also a solution of the weak form and vice versa. We will use the Finite Element Method to find approximate solutions to the weak form, which in turn are solutions to the strong form. The derivation of the weak form from the strong form is below.

2.1.1 Derivation of Weak Form

Define $h = \frac{b-a}{n}$ where n is the number of intervals and denote the nodes $a = x_0, x_1, x_2, \ldots, x_n = b$. Then we define functions $\Phi_i : [a, b] \to [0, 1]$ such that $\Phi_i(x) = \begin{cases} L_1^i & x_i - h \leq x \leq x_i \\ L_0^{i+1} & x_i \leq x \leq x_i + h \\ 0 & otherwise \end{cases}$ where L_1^i is the Lagrange polynomial of $(x_i - h, 0)$ and $(x_i, 1)$ and L_0^{i+1} is the Lagrange polynomial of $(x_i, 1)$ and $(x_i + h, 0)$.

Note that the solution $u(x)$ will be of the form $u(x) = u_1\phi_1(x) + \cdots + u_n\phi_n(x)$ where $u_i = u(x_i)$.

Now we will multiply Φ_i to the differential equation on both sides.

$$\Phi_i(x)\frac{\partial^2 u}{\partial x^2} + \Phi_i(x)f(x) = 0.$$

We take the integral over the rod length:

$$\int_a^b \Phi_i(x)\left(\frac{\partial^2 u}{\partial x^2} - f(x)\right)dx = 0. \tag{2}$$

Now we integrate $\int_a^b \Phi_i(x)\frac{\partial^2 u}{\partial x^2}dx$ by parts:

$$\begin{aligned}\int_a^b \Phi_i(x)\frac{\partial^2 u}{\partial x^2}dx &= \Phi_i(x)\frac{\partial u}{\partial x}\Big|_a^b - \int_a^b \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x}dx & (3)\\ &= \Phi_i(b)\frac{\partial u}{\partial x}(b) - \Phi_i(a)\frac{\partial u}{\partial x}(a) - \int_a^b \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x}dx & (4)\\ &= -\int_a^b \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x}dx & (5)\end{aligned}$$

as $\Phi_i(a) = \Phi_i(b) = 0$.

So Eq. (2) becomes the weak form:

$$\int_a^b \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x} + \Phi_i(x)f(x)dx = 0. \tag{6}$$

2.1.2 Solving the Weak Form

Note that Eq. (6) must hold for any Φ_i so this general equation can be broken into n equations for each Φ_i. Furthermore, since Φ_i is 0 for $x < x_i - h$ and $x > x_i + h$, we only have to integrate over $x_i - h \leq x \leq x_i + h$:

$$\int_a^b \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x} + \Phi_i(x)f(x)dx = \int_{x_i-h}^{x_i+h} \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x} + \Phi_i(x)f(x)dx = 0. \tag{7}$$

Recall that $\Phi_i(x)$ is defined using Lagrange polynomials and $u(x) = u_1\Phi_1(x) + \dots u_n\Phi_n(x)$. We use properties of the Lagrange polynomials and $u(x)$ to simplify Eq. (7). Then we have

$$\int_{x_i-h}^{x_i+h} \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x} + \Phi_i(x)f(x)dx = \int_{x_i-h}^{x_i} \left(u_{i-1}(L_0^i)'(x) + u_i(L_1^i)'(x)\right)(L_1^i)'(x) + L_1^i(x)f(x)dx$$
$$+ \int_{x_i}^{x_i+h} \left(u_i(L_0^{i+1})'(x) + u_{i+1}(L_1^{i+1})'(x)\right)(L_0^{i+1})'(x) + L_0^{i+1}(x)f(x)dx. \quad (8)$$

Then we have a system of equations where each of these $\int_{x_i-h}^{x_i+h} \frac{\partial \Phi_i}{\partial x}\frac{\partial u}{\partial x} + \Phi_i(x)f(x)dx = 0$ for each $i = 1, 2, \dots, n-1$. We then group by coefficients u_i, use a Gaussian Quadrature Rule to replace the integrals, rewrite the system as a matrix, and solve for u_i. Note that this matrix is "nice" matrix as it is a "banded" matrix, so there is a band of nonzero elements along the main diagonal with the remaining entries being 0. This provides easier and faster computation to solve for the coefficients and thus, solve for $u(x)$.

3 Convergence of Finite Element Methods

Theorem 3.1 *Denote $u^h(x)$ the numerical approximation of $u(x)$ using step size $h > 0$. Then $u^h(x_A) = u(x_A)$ for $A = 1, 2, \dots, n+1$.*

This theorem is stating that at the nodes of FEM, the numerical approximation $u^h(x)$ and exact solution $u(x)$ are the same value. It is not obvious that this theorem shows the convergence of the finite element method. However, as $n \to \infty$, the number of points agreed upon by $u^h(x)$ and $u(x)$ also heads to infinity.

Note that the proof is not imperative to the understanding of the project and one can skip to the next section. Additionally, the convergence of the method is depicted in Figure 2.

Before we continue to the proof of Theorem 3.1, we must first state some definitions. Let H^1 be the set of continuous functions such that the derivative is square integrable $H^1 = \left\{ f : \int_0^1 \left(\frac{\partial f}{\partial x}\right)^2 dx < \infty \right\}$. These functions are nice as they form a complete inner metric space under an inner product and are used in several applications. Then the set $\mathcal{V} = \{w : w \in H^1, w(1) = 0\}$. These functions are often referred to as "weighting functions" or "variations." Further, we denote a finite dimensional collection of functions in $\mathcal{V}$, to be $\mathcal{V}^h$. Consider the Dirac delta function at $y \in [0,1]$, $\delta_y(x) = \delta(x-y)$ defined so that for any continuous function w on $[0,1]$ we have the following property:

$$(w, \delta_y(x)) = \int_0^1 w(x)\delta(x-y)dx = w(y).$$

A more visual way to think of $\delta_y(x)$ is that it is a concentrated pulse force at y and 0 everywhere else. Using the Dirac delta function, we consider a related problem to the strong form, Eq. (1):

$$\frac{\partial^2 g}{\partial x^2} + \delta_y = 0 \quad \text{on } [0,1] \text{ with } g(1) = 0, \frac{\partial g}{\partial x}(0) = 0 \quad (9)$$

where g is the Green's function. Similarly, there is variation of the weak form:

$$\int_0^1 \frac{\partial w}{\partial x}\frac{\partial g}{\partial x}dx = (w, \delta_y) = w(y). \quad (10)$$

We will denote $a(w, g) = \int_0^1 \frac{\partial w}{\partial x}\frac{\partial g}{\partial x}dx$ for the remainder of the report.

To prove Theorem 3.1, we introduce two lemmas.

Lemma 3.2 *$a(u - u^h, w^h) = 0$ for all $w^h \in \mathcal{V}^h$.*

To prove this, we use the weak form (Eq. (6)) where $a = 0, b = 1$ and the **Galerkin Equation:** given functions f and h, find $u^h = v^h + g^h$ such that for all $w^h \in \mathcal{V}^h$ we have

$$a(w^h, v^h) = (w^h, f). \quad (11)$$

Then

$$a(w^h, u) - a(w^h, u^h) = (w^h, f) - \big((w^h, f)\big) = 0. \tag{12}$$

Notice that $a(u - u^h, w^h) = a(u, w^h) - a(u^h, w^h)$ as $\int_0^1 \frac{\partial(u-u^h)}{\partial x}\frac{\partial g}{\partial x}dx = \int_0^1 \frac{\partial u}{\partial x}\frac{\partial g}{\partial x}dx - \int_0^1 \frac{\partial u^h}{\partial x}\frac{\partial g}{\partial x}dx$. Furthermore, $a(w, v) = a(v, w)$ as order of the product of the functions in the integral do not matter. Thus, $a(u - u^h, w^h) = 0$ as required.

We continue to the second lemma.

Lemma 3.3 *$u(y) - u^h(y) = a(u - u^h, g)$ where g is Green's function.*

To prove this, we first use the definition of δ_y. Then we have $(u - u^h, \delta_y) = \int_0^1 (u - u^h)(x)\delta(x - y)dx = (u - u^h)(y) = u(y) - u^h(y)$. Then using Eq. (10), we have $(u - u^h, \delta_y) = a(u - u^h, g)$ since $u - u^h \in \mathcal{V}$. With these two lemmas, we can now prove Theorem 3.1.

Proof: Let $y = x_A$ for $A = 1, 2, \ldots, n$ and $g \in \mathcal{V}^h$. Then by Lemma 3.3 we have $u(x_A) - u^h(x_A) = a(u - u^h, g)$. By Lemma 3.2, we have $u(x_A) - u^h(x_A) = a(u - u^h, g) = 0$ as required. □

4 Methods

We observe that in the theoretical work that we will need to approximate many integrals to use as values in solving a system of linear equations. For this, we will use a Gaussian Quadrature Rule of arbitrary order $n = 2$. We also recognize that the "hat" or "tent" functions are made of up linear Lagrange Interpolation polynomials of degree 1 and due to the nature of their structure, we can generalize functions for the `ith` element for nodes 1 and 2.

Using these methods, we are able to compute the values that fall into a matrix A, of the form of a tri-diagonal matrix, that satisfies the equation $Ax = b$ where each row of A and b can be computed using the following generalization of the linear equation from the weak form:

$$\begin{aligned} u_{i-1}\int_{(i-1)h}^{ih} L_0^{i'} L_1^{i'} + u_i\left(\int_{(i-1)h}^{ih} L_1^{i'} L_1^{i'} + \int_{ih}^{(i+1)h} L_0^{i+1'} L_0^{i+1'}\right) \\ + u_{i+1}\int_{ih}^{(i+1)h} L_1^{i+1'} L_0^{i+1'} = -\int_{(i-1)h}^{ih} L_1^{i} f(x) - \int_{ih}^{(i+1)h} L_0^{i+1} f(x) \end{aligned}$$

We use this generalized form to compute the tri-diagonal entries in A and the column vector entries in b to solve for x in $Ax = b$ using MATLAB's built in linear equation solver. Using the resulting coefficients u_i for $i = 1, 2, \ldots$ we can construct the approximated solution

$$u(x) = u_1 * \Phi_0(x) + u_2 * \Phi_1(x) + \ldots$$

$$u(x) = \sum_{i=1}^{n} u_i \Phi_{i-1}(x)$$

5 Conclusion

We sought out to solve the differential equation $\frac{\partial u^2}{\partial x^2} = f(x)$ for some function f and boundary conditions $u(1) = a$ and $u(n) = b$. We have chosen the simple function $f(x) = x^2 - 3x$ with boundary conditions $u(0) = u(1) = 0$. Using the approach discussed in the Methods Section 4, we are able to approximate f with an arbitrary number of n "tent" functions, denoted $\Phi_i(x)$ in this report.

Here we plot the numerical approximation using $n = 20$ against the actual solution; we observe that our approximation does in fact approach the actual solution.

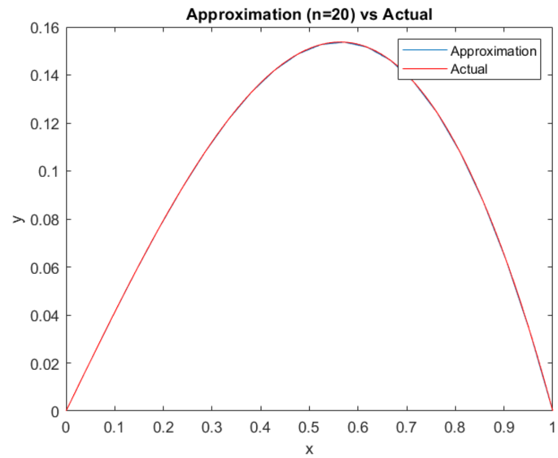

Figure 1: Plotting numerical approximation against the actual solution

We can also experimentally show the convergence theorem by plotting multiple approximations with increasing n:

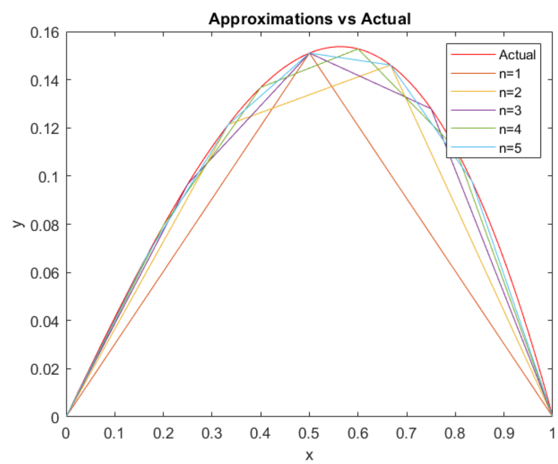

Figure 2: Showing convergence theorem experimentally

Code to reproduce our results can be found at this Github link.

References

[Dai14] R. C. Daileda. The one-dimensional heat equation, February 2014.

[Hug87] Thomas J. R. Hughes. *The finite element method*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1987. Linear static and dynamic finite element analysis, With the collaboration of Robert M. Ferencz and Arthur M. Raefsky.