DISCRETE APPROXIMATION TO PDEs

0. Many PDE problems cannot be solved "in closed form"; that is, although one can show that they have solutions and can determine some properties of the solutions, one cannot write down formulas for the solutions. Thus information about these problems only becomes "useful" when some means of approximating the solutions can be found, and one needs a computational approach to these problems. While this course will offer very little material on the numerical solution of partial differential equations,[1] there are a few small topics that we should look at because they help one to understand the behavior of solutions of the equations from a purely mathematical standpoint.

1. Approximate Derivatives. {Some of this material can be found in Strauss's book, §8.1, in considerably more condensed form. Notational differences: He's calling the function $u(x)$ instead of $f(x)$, and his "small increment" is named Δx instead of h. The content is about the same. The reader should note, however, that the pathology Strauss describes in his §8.2 will not occur for us, because we are "discretizing in the space variable only," not in both the "space" and "time" variables.} Both integration and differentiation of "arbitrary" functions involve limit processes, and therefore cannot be carried out in a "finite" way on a computer.[2] Let's look at differentiation. The definition

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

suggests that if h is sufficiently small, then $\dfrac{f(x+h) - f(x)}{h}$ can be used to approximate $f'(x)$. This is true but too naïve—and inaccurate—to work very well. Some insight into improving this approximation can be gained by looking at Taylor series, or—better—at Taylor polynomials with error term.[3] At a fixed point x in the domain of f, we have for variable h

$$f(x+h) = f(x) + \frac{f'(x)}{1!}\,h + \frac{f''(z_h)}{2!}\,h^2$$
$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{f''(z_h)}{2!}\,h$$

and the last term on the r. h. s., $\dfrac{f''(z_h)}{2!}\,h$, is an **error term** that tells us precisely how much our approximation differs from the true value of $f'(x)$. Clearly the error will be worse the bigger f'' is near x; that is geometrically reasonable, because f'' measures curvature of the graph of f (if $f''(x) \equiv 0$, then the function $f(x) \equiv ax + b$—its graph is a straight line—and our "approximate" formula for the derivatives becomes exact), and the more curved the graph is, the worse becomes the "secant line approximation to the slope of the tangent line" that our approximate-derivative formula represents.

When the error term has the form of a bounded function times a power h^k of h—as, for example, our error term $\dfrac{f''(z_h)}{2!}\,h$ does, with $k = 1$—it is customary to say that the error term is "$O(h^k)$". Sometimes one will also abbreviate the error term by writing "$O(h^k)$" instead of the error term itself: thus the last set-off line above might have been written as

$$\frac{f(x+h) - f(x)}{h} = f'(x) + O(h)\ .$$

---

(1) This is not to underemphasize the importance of the topic! It is literally true that *all* the research interests of the numerical analysts in this mathematics department center on the numerical treatment of problems in partial differential equations.

(2) Of course it is possible to differentiate and integrate some elementary functions (where "elementary" has a precise meaning) symbolically, as users of Maple and MATLAB know. But if one wants to get a number out of integrating $\exp(-x^2/2)$ over a finite interval, one will have to approximate it somehow.

(3) For information on this subject, see Stewart's calculus book, used in the first three semesters of calculus at RU–NB. In the 3rd edition, see sections 10.10 and 10.12, p. 633 ff.; in the 4th edition, more-or-less the same material is found in sections 11.10 and 11.12, p. 751 ff. The approach to the remainder term is different in the two editions: the version we're using, which is the standard one known to generations of mathematicians, is emphasized in the 3rd edition but finessed in favor of an estimate in the 4th edition.

This is a rather informal definition of the notion of **big-oh of** h^k, but it will do for the time being. In this case, the fact that the error is $O(h^1)$ tells us that the approximation is not very good: if $h = .01 = 10^{-2}$, for example, then the error is also of the size 10^{-2}, and two-digits-after-the-decimal-point accuracy isn't very accurate. Can we find a better approximation? Let's try comparing the difference-quotient approximations for two secant lines, one on one side of x and one on the other, and let's write more terms of the Taylor polynomial. Using both h and $-h$ as increments, we have

$$\begin{aligned} f(x+h) &= f(x) + \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 + \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_h)}{4!}\,h^4 \\ f(x-h) &= f(x) - \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 - \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_{-h})}{4!}\,h^4 \end{aligned}$$

where the second line exhibits alternation of signs corresponding to the fact that even powers of h and $-h$ are the same, but odd powers undergo a sign change. The number $f'(x)$ occurs with different sign on the r. h. sides of the two equations. If we **subtract** the second equation from the first the terms containing $f'(x)$ will reinforce, but many other terms will cancel:

$$\begin{aligned} f(x+h) &= f(x) + \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 + \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_h)}{4!}\,h^4 \\ f(x-h) &= f(x) - \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 - \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_{-h})}{4!}\,h^4 \\ f(x+h) - f(x-h) &= 2\cdot\frac{f'(x)}{1!}\,h + 2\cdot\frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_h) - f^{(4)}(\xi_{-h})}{4!}\,h^4 \end{aligned}$$

and dividing both sides by $2h$ results in

$$\begin{aligned} \frac{f(x+h) - f(x-h)}{2h} &= f'(x) + \frac{f^{(3)}(x)}{3!}\,h^2 + \frac{f^{(4)}(\xi_h) - f^{(4)}(\xi_{-h})}{2}\,\frac{h^3}{4!} \\ \frac{f(x+h) - f(x-h)}{2h} &= f'(x) + O(h^2) \end{aligned}$$

where again we have followed the convention of writing "$O(h^2)$" to indicate an error term—whose exact form does not concern us at this point—that can be estimated by a bounded function of h times h^2. Note that this formula **doubles the accuracy obtained by simply taking difference quotients**: if we use this "symmetric difference quotient" to approximate $f'(x)$, then for $h = 10^{-2}$ we have an error that is measured in $h^2 = (10^{-2})^2 = 10^{-4}$-ths—we have doubled (in terms of accurate decimal digits) the accuracy that the simple one-sided difference quotient gives for approximating the derivative of f. Moreover, this approximate formula is **exact** on quadratic functions like $f(x) = ax^2 + bx + c$, while the one-sided approximate formula that we started with was only exact on linear functions.

In a similar manner, we can derive an approximate expression for $f''(x)$: indeed, if we had added the two Taylor polynomials with increments h and $-h$ above, we would have gotten

$$\begin{aligned} f(x+h) &= f(x) + \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 + \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_h)}{4!}\,h^4 \\ f(x-h) &= f(x) - \frac{f'(x)}{1!}\,h + \frac{f''(x)}{2!}\,h^2 - \frac{f^{(3)}(x)}{3!}\,h^3 + \frac{f^{(4)}(\xi_{-h})}{4!}\,h^4 \\ f(x+h) + f(x-h) &= 2\cdot f(x) + 2\cdot\frac{f''(x)}{2!}\,h^2 + \frac{f^{(4)}(\xi_h) + f^{(4)}(\xi_{-h})}{4!}\,h^4 \\ \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} &= f''(x) + \frac{f^{(4)}(\xi_h) + f^{(4)}(\xi_{-h})}{4!}\,h^2, \text{ or less exactly} \\ \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} &= f''(x) + O(h^2)\,. \end{aligned}$$

The way we use these expressions in seeking to approximate a PDE will be discussed below.

2. Discrete Approximations of PDEs—the Heat Equation. Let's begin by considering the "diffusion" or "heat" PDE, Strauss's (8) on p. 15 or (10) on p. 16. With "an appropriate choice of units" the constant k of (8) or the function κ of (10) can be taken as equal (identically) to 1, and we get the **normalized heat equation** (where t is the "time' variable and x is the "space" variable in $u(x,t)$)

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

To **discretize** this equation **in the space variable(s)**, we imagine the x-axis "marked out in units of length h," so that $x_i = i * h$ for $i = \dots, -2, -1, 0, 1, 2, \dots$. We consider the values of $u(x,t)$ only for $x = x_i$ for some index i. This gives us a sequence of functions of the one variable t indexed by i, which we obtain by setting

$$u_i(t) = u(x_i, t) \ .$$

Next, we replace the second-partial-with-respect-to-x operation by its discrete approximation found above ($f''(x)$ for the function $f(x)$):

$$u_{xx}(x_i,t) \approx \frac{u(x_{i-1},t) - 2u(x_i,t) + u(x_{i+1},t)}{h^2} \ .$$

This approximation gives us a system of ordinary differential equations, indexed by i, as an approximation to the heat PDE with which we started:

$$\begin{gathered} \cdots \\ \frac{du_i(t)}{dt} = \frac{1}{h^2} \cdot [u(x_{i-1},t) - 2u(x_i,t) + u(x_{i+1},t)] \\ \cdots \end{gathered}$$

or in matrix form

$$\frac{d}{dt}\begin{bmatrix} \vdots \\ \vdots \\ u_i(t) \\ \vdots \\ \vdots \end{bmatrix} = \frac{1}{h^2}\begin{bmatrix} \ddots & & & & \\ & \ddots & & & \\ \cdots & 1 & -2 & 1 & \cdots \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}\begin{bmatrix} \vdots \\ \vdots \\ u_i(t) \\ \vdots \\ \vdots \end{bmatrix} \ .$$

We have deliberately left the dimension of these vectors and matrices undefined for now, since this discussion is heuristic rather than rigorous—we're trying to get some insight into how the solutions of the heat equation might behave, based on our knowledge of the behavior of solutions of linear systems of ODEs. However, matrices of this "shape" share some common properties, regardless of dimension. For one thing, a matrix of the form indicated on the last set-off line above is a **symmetric** matrix, and therefore **all its eigenvalues are real**. It is not difficult to show, using the "pattern of signs" of the entries in this matrix, that **all its eigenvalues are nonpositive**. Just as a computational example, let's use `Matlab` to find decimal approximations for the eigenvalues of a 11×11 matrix of the form

$$A = \frac{1}{h^2}\begin{bmatrix} -1 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ 0 & 1 & -2 & 1 & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ & & & & 1 & -2 & 1 \\ & & & & & 1 & -1 \end{bmatrix}$$

(in which the 0th and 10-th rows require, approximately, that $\dfrac{\partial u}{\partial x}$ be zero at $x = 0$ and $x = 10h$—this is approximately the "insulated endpoints" or **Neumann boundary condition**, about which more later in the course). We can think of x ranging between 0 and 1 in steps of $h = 0.1$. For this choice of h the eigenvalues are negative numbers ranging between ≈ -391.89 and 0—it's easy to see that 0 is an eigenvalue, since this matrix sends $[1, \ldots, 1]^T$ to the zero vector. This tells us that up to a change of basis, or replacement of A by a diagonal matrix(4)

$$S^{-1}AS = \begin{bmatrix} 0 & & & & \\ & \lambda_1 & & & \\ & & \lambda_2 & & \\ & & & \ddots & \\ & & & & -391.89 \end{bmatrix} = \Lambda$$

—the vector-matrix differential equation $\dfrac{d\mathbf{u}(t)}{dt} = A\mathbf{u}(t)$ has the solution

$$\mathbf{u}(t) = \exp(tA)\mathbf{u}(0) = S\exp(t\Lambda)S^{-1}\mathbf{u}(0) = S \begin{bmatrix} e^{0t} & & & & \\ & e^{\lambda_1 t} & & & \\ & & e^{\lambda_2 t} & & \\ & & & \ddots & \\ & & & & e^{-391.89t} \end{bmatrix} S^{-1}\mathbf{u}(0) \qquad (*)$$

in terms of the "initial position" $\mathbf{u}(0)$—which here should be thought of as the discretized initial heat distribution, or initial distribution of mass before diffusion:

$$\mathbf{u}(0) = \begin{bmatrix} u(0,0) \\ u(0.1,0) \\ u(0.2,0) \\ \vdots \\ u(1,0) \end{bmatrix}.$$

Another way of writing the solution $(*)$ would be(5) to write $\mathbf{u}(0)$ as a linear combination of eigenvectors of A and then use the fact that if the initial position vector for $\mathbf{u}' = A\mathbf{u}$ is an eigenvector $\mathbf{v}$ of A belonging to the eigenvalue λ, then the solution of the initial-value problem is $\mathbf{u}(t) = e^{\lambda t}\mathbf{v}$. If the eigenvectors of this A are $\mathbf{v}_0, \ldots, \mathbf{v}_{10}$, where the corresponding eigenvalues $\lambda_0, \ldots, \lambda_{10}$ are numbered in increasing order of size, so $\lambda_0 = 0$ and $\lambda_{10} = -391.89$, then from the expansion

$$\mathbf{u}(0) = a_0\mathbf{v}_0 + \cdots + a_{10}\mathbf{v}_{10}$$

of the initial heat distribution, we would get

$$\mathbf{u}(t) = a_0 e^{0t}\mathbf{v}_0 + \cdots + a_{10}e^{-391.89t}\mathbf{v}_{10} = a_0\mathbf{v}_0 + \cdots + a_{10}e^{-391.89t}\mathbf{v}_{10}$$

(4) Here S is the matrix whose columns are the eigenvectors of A. In this case, because A is symmetric, we know that S can be taken to be an orthogonal matrix (whose inverse is its transpose); this would simplify the calculation if we were actually doing it concretely.

(5) If you took Math 244 instead of Math 252, this way of writing the solution might look more familiar.

as the solution of $\mathbf{u}'(t) = A\mathbf{v}(t)$ with the correct initial value. This solution exhibits some of the kind of behavior that we expect to see of heat: as $t \to +\infty$, all the terms of this expansion decay rapidly to $\mathbf{0}$ with the exception of the term $a_0\mathbf{v}_0$, which has no t-dependence and thus simply sits still. We have already identified the eigenvector of A belonging to $\lambda_0 = 0$ as $[1, \ldots, 1]^T$, so the solution of the discretized heat equation tends to a "solution of constant temperature" as $t \to +\infty$—exactly what physical intuition would demand. Moreover, because A is symmetric so its eigenvectors are orthogonal, we can pick off the $\mathbf{v}_0$ term in the expansion of $\mathbf{u}(0)$ simply by dotting with $[1, \ldots, 1]^T$:

$$\begin{aligned}
\mathbf{u}(0) &= a_0\mathbf{v}_0 + \cdots + a_{10}\mathbf{v}_{10} \\
[1, \ldots, 1]^T \bullet \mathbf{u}(0) &= a_0[1, \ldots, 1]^T \bullet \mathbf{v}_0 + \cdots + a_{10}[1, \ldots, 1]^T \bullet \mathbf{v}_{10} = a_0[1, \ldots, 1]^T \bullet \mathbf{v}_0 \\
u_0(0) + \cdots + u_{10}(0) &= a_0 \cdot 11 \quad (\text{since } [1, \ldots, 1]^T \bullet \mathbf{v}_0 = [1, \ldots, 1]^T \bullet [1, \ldots, 1]^T) \\
a_0 &= \frac{u_0(0) + \cdots + u_{10}(0)}{11}, \quad \text{the "average initial temperature"}
\end{aligned}$$

so $\lim_{t\to+\infty} \mathbf{u}(t) = \frac{u_0(0) + \cdots + u_{10}(0)}{11} [1, \ldots, 1]^T$, the constant solution of the (discretized) heat equation whose temperature is the average of the (discretized) initial temperature—again, perfectly in accord with physical intuition (remember, the ends of the bar are insulated). And as a final example of how well the heat equation (in this discretized form) corresponds to physical intuition: if we write

$$A = 100 \begin{bmatrix} -1 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ 0 & 1 & -2 & 1 & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ & & & & 1 & -2 & 1 \\ & & & & & 1 & -1 \end{bmatrix}$$

$$= 100 \begin{bmatrix} -2 & 0 & & & & & \\ 0 & -2 & 0 & & & & \\ 0 & 0 & -2 & 0 & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ & & & & 0 & -2 & 0 \\ & & & & & 0 & -2 \end{bmatrix} + 100 \begin{bmatrix} 1 & 1 & & & & & \\ 1 & 0 & 1 & & & & \\ 0 & 1 & 0 & 1 & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ & & & & 1 & 0 & 1 \\ & & & & & 1 & 1 \end{bmatrix} = -200I_{11} + B$$

where I_{11} is (of course) the 11×11 identity matrix and "B" is the second matrix on the second set-off line, then because the identity matrix commutes with all matrices we have

$$e^{tA} = e^{-200tI}e^{tB} = e^{-200t}e^{tB} = e^{-200t}\sum_{k=0}^{\infty} \frac{t^k}{k!} B^k$$

in which all the matrices in the series for e^{tB} have only nonnegative numbers as their entries. It follows that for all values of t, all the entries in the matrix e^{tA} must be nonnegative numbers. Thus if $\mathbf{u}(0)$ is a vector with all nonnegative coördinates—*i.e.*, the temperature is nonnegative at time zero—then $\mathbf{u}(t) = e^{tA}\mathbf{u}(0)$ will have all nonnegative coördinates for every value of t: in the course of redistribution of heat, nonnegative temperatures never go negative. In this situation it is not difficult to show, more generally, that if $m \le u_i(0) \le M$ for $i = 0, \ldots, 10$—*i.e.*, the initial temperature is bounded below by m and above by M—then $m \le u_i(t) \le M$ for $i = 0, \ldots, 10$ holds for all $t \ge 0$. Just consider the vector $\mathbf{u}(0) - m[1, \ldots, 1]^T$, all of whose coördinates are nonnegative, and apply e^{tA} to it, obtaining $\mathbf{u}(t) - m[1,\ldots,1]^T = e^{tA}[\mathbf{u}(0) - m[1,\ldots,1]^T]$—remember, $[1,\ldots,1]^T$ is the eigenvector of A corresponding to the eigenvalue zero—in which all the coördinates of the vector on the r. h. s. are nonnegative, which fact shows that $u_i(t) \geq m$ for all $t \geq 0$. A similar argument employing $M[1,\ldots,1]^T - \mathbf{u}(0)$ shows that $u_i(t) \leq M$ for all $t \geq 0$. Again, the fact that initial bounds on the temperature are never exceeded is strongly in accord with physical intuition.

3. Discrete Approximations of PDEs—the Heat Equation again. There is a slightly different way to view the discretized heat equation that gives more insight into what can be expected from the "continuous heat equation." Instead of cutting off the r. h. s. of the formally-infinite matrix-vector differential equation

$$\frac{1}{h^2}\begin{bmatrix} \ddots & & & & \\ & \ddots & & & \\ \cdots & 1 & -2 & 1 & \cdots \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}\begin{bmatrix} \vdots \\ \vdots \\ u_i(t) \\ \vdots \\ \vdots \end{bmatrix}$$

at both ends, "making endpoint adjustments" to give systems like the 11×11 system considered in the preceding §, and looking for eigenvectors and -values, let us consider that formal infinite matrix as defining a "discrete linear operator" that sends sequences to sequences, and try to find its eigenvectors and -values. Thus we would be looking for (infinite-in-both-directions) sequences $(\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ and eigenvalues λ with the properties[(6)] that for every index i,

$$x_{i-1} - 2x_i + x_{i+1} = \lambda x_i \ .$$

There are very systematic ways to hunt for such sequences, but for the time being we can be guided by the fact that the addition formulas for the sine and cosine produce any number of these sequences, because for any number ω (even if it's complex)

$$\begin{aligned} \sin\omega(x+h) &= \sin\omega x\cos h + \cos\omega x\sin h \\ \sin\omega(x-h) &= \sin\omega x\cos h - \cos\omega x\sin h \\ \sin\omega(x-h) - 2\sin\omega x + \sin\omega(x+h) &= 2\sin\omega x\cos\omega h - 2\sin\omega x \\ &= -2(1-\cos\omega h)\sin\omega x = -4\sin^2(\omega h/2)\sin\omega x \\ \frac{\sin\omega(x-h) - 2\sin\omega x + \sin\omega(x+h)}{h^2} &= -\left(\frac{\sin(\omega h/2)}{h/2}\right)^2 \sin\omega x \ , \end{aligned}$$

and similarly[(7)]

$$\begin{aligned} \cos\omega(x-h) &= \cos\omega x\cos h + \sin\omega x\sin h \\ \cos\omega(x-h) &= \cos\omega x\cos h - \sin\omega x\sin h \\ \cos\omega(x-h) - 2\cos\omega x + \cos\omega(x+h) &= 2\cos\omega x\cos\omega h - 2\cos\omega x \\ &= -2(1-\cos\omega h)\cos\omega x = -4\sin^2(\omega h/2)\cos\omega x \\ \frac{\cos\omega(x-h) - 2\cos\omega x + \cos\omega(x+h)}{h^2} &= -\left(\frac{\sin(\omega h/2)}{h/2}\right)^2 \cos\omega x \ . \end{aligned}$$

(6) Of course there should be a division by h^2 in there, but we'll deal with that in a moment.

(7) We could as easily have got the cosine relation by remembering that $\cos x = \sin(x+\pi/2)$, but redundancy may bring conviction.

Then for every choice of ω we have

$$\frac{1}{h^2}\begin{bmatrix} \ddots & & & & & & \\ & \ddots & & & & & \\ & & \ddots & & & & \\ \cdots & \cdots & 1 & -2 & 1 & \cdots & \cdots \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & \ddots \end{bmatrix}\begin{bmatrix} \vdots \\ \vdots \\ \sin(\omega(i-1)h) \\ \sin\omega(ih) \\ \sin(\omega(i+1)h) \\ \vdots \\ \vdots \end{bmatrix} = -\left(\frac{\sin(\omega h/2)}{h/2}\right)^2 \begin{bmatrix} \vdots \\ \vdots \\ \sin(\omega(i-1)h) \\ \sin\omega(ih) \\ \sin(\omega(i+1)h) \\ \vdots \\ \vdots \end{bmatrix}$$

and

$$\frac{1}{h^2}\begin{bmatrix} \ddots & & & & & & \\ & \ddots & & & & & \\ & & \ddots & & & & \\ \cdots & \cdots & 1 & -2 & 1 & \cdots & \cdots \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & \ddots \end{bmatrix}\begin{bmatrix} \vdots \\ \vdots \\ \cos(\omega(i-1)h) \\ \cos\omega(ih) \\ \cos(\omega(i+1)h) \\ \vdots \\ \vdots \end{bmatrix} = -\left(\frac{\sin(\omega h/2)}{h/2}\right)^2 \begin{bmatrix} \vdots \\ \vdots \\ \cos(\omega(i-1)h) \\ \cos\omega(ih) \\ \cos(\omega(i+1)h) \\ \vdots \\ \vdots \end{bmatrix} .$$

Moreover—and this may not have escaped the reader's notice—if we consider what happens when $h \to 0^+$, we get (in the cosine case—the sines behave the same way)

$$\lim_{h\to 0^+} \frac{\cos\omega(x-h) - 2\cos\omega x + \cos\omega(x+h)}{h^2} = \left\{-\lim_{h\to 0^+}\left(\frac{\sin(\omega h/2)}{h/2}\right)^2\right\}\cos\omega x = -\omega^2\cos x$$

where $-\omega^2$ is the eigenvalue of the linear mapping of functions $u \mapsto u''$ that corresponds[(8)] to the "eigenvector" (or **eigenfunction**) $\cos\omega x$.

But we now have a problem: we have too many eigenvectors! We get two for every complex number ω. However, let us look at what happens for our discretized 11-point "unit interval" 0.0,0.1,...,0.9,1.0 that we got when we took $h = 0.1$ in §2 above. If we insulate the endpoints, then we are specifying that there be no heat flux at $x = 0.0$ and $x = 1.0$. Heat flux is proportional to the space derivative (or gradient in dimensions > 1), so in the discretized heat equation it would be reasonable to use only the vectors of cosines: $\cos\omega x$ automatically has derivative zero at the left endpoint $x = 0.0$. However, $\cos\omega x$ has derivative $-\omega\sin\omega$ at the right endpoint $x = 1.0$, and this will be zero if and only if $\omega = j\pi$ for some integer j. Making those choices of ω will take care of the "boundary condition." So we now have 11 "basic eigenvectors" of the form

$$\mathbf{C}_j = \begin{bmatrix} \cos(j\pi\cdot 0h) \\ \cos(j\pi\cdot 1h) \\ \vdots \\ \cos(j\pi\cdot 9h) \\ \cos(j\pi\cdot 10h) \end{bmatrix} \quad \text{where } h = 0.1 \text{ and } j = 0,\ldots,10$$

(8) As we shall see, these eigenfunctions play the same rôle for the continuous heat equation as the vectors that we shall now construct will play for the discretized heat equation.

where—according to the computation on pp. 6–7 above—the eigenvalues associated with $\mathbf{C}_j$ for $j = 0, \ldots, 10$ are the numbers $-\omega_j^2 = -\left(\dfrac{\sin(j\pi/20)}{1/20}\right)^2 = -400\sin^2(j\pi/20)$ respectively, and thus range between 0 and -400. These eigenvectors are orthogonal[9] with respect to an inner product slightly different from the usual one, namely the inner product given by

$$[u_0, \ldots, u_{10}]^T \bullet [v_0, \ldots, v_{10}]^T = \frac{1}{2}\, u_0 v_0 + u_1 v_1 + \cdots + u_9 v_9 + \frac{1}{2}\, u_{10} v_{10} \ ;$$

the reason for the "half-weights" on the endpoints $0 = 0h$ and $1 = 10h$ has to do with certain properties of the Fourier transform (discrete or continuous) and does not require an explanation now. In any event, since there are 11 of these vectors and they are orthogonal and nonzero they form a basis for $\mathbb{R}^{11}$, and thus any initial temperature distribution $[U(0.0), \ldots, U(1.0)]^T$ on the "discrete unit interval" 0.0,0.1,. . . ,0.9,1.0 can be represented in some (unique) form

$$\begin{bmatrix} U(0.0) \\ U(0.1) \\ \vdots \\ U(1.0) \end{bmatrix} = c_0 C_0 + c_1 C_1 + \cdots + c_{10} C_{10} \ .$$

It then follows by the same argument as in §2 above that if we multiply the eigenvector in each term by the exponential of t times its corresponding eigenvalue, we obtain a vector-valued function

$$\mathbf{u}(t) = \begin{bmatrix} u_0(t) \\ u_1(t) \\ \vdots \\ u_{10}(t) \end{bmatrix} = c_0 e^{0t} C_0 + c_1 e^{-\omega_1^2 t} C_1 + \cdots + c_{10} e^{-\omega_{10}^2 t} C_{10} \qquad (\$)$$

that satisfies both the initial condition $\mathbf{u}(t) = \begin{bmatrix} U(0.0) \\ U(0.1) \\ \vdots \\ U(1.0) \end{bmatrix}$—because the coefficients c_j were taken to insure this—and the discretized heat equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{u(x-h) - 2u(x) + u(x+h)}{h^2}$$

at each $x_i = ih$, $i = 0, \ldots, 10$. This will appear more striking if we write out the equation ($) in functions instead of vectors: the equation ($) says that the function of (x, t) given by

$$u(x,t) = c_0 e^{0t} \cos(0\pi x) + c_1 e^{-\omega_1^2 t} \cos(1\pi x) + \cdots + c_{10} e^{-\omega_{10}^2 t} \cos(10\pi x)$$

satisfies the discretized heat equation $\dfrac{\partial u(x,t)}{\partial t} = \dfrac{u(x-h) - 2u(x) + u(x+h)}{h^2}$ at each $x = x_i$, $i = 0, \ldots, 10$, because it satisfies it term-by-term. And of course—because the derivatives of these cosines are zero at $x = 0$ and $x = 1$—this function also satisfies the "insulated endpoints" boundary conditions, with no heat flux at the endpoints.

(9) They are not orthonormal, however; the 0-th and 10-th ones have length-squared 10 and the others have length-squared 5. The reader familiar with Maple or MATLAB might enjoy checking these assertions about orthogonality and length; I checked them before asserting them here, although there are good abstract reasons for knowing them to be true.

4. Discrete Approximations of PDEs—the Wave Equation. We shall eventually "take the limit as $h \to 0^+$" and get exact solutions of the heat and wave equations on finite intervals. Meanwhile, let's look at the analogue of what we did with the heat equation in §3 above, but this time let's consider the wave equation in the following situation:

(1) We look at the string only for $0 \le x \le 1$, and we assume that

(2) the string is fixed at both endpoints, so $u(0,t) \equiv 0$ and $u(1,t) \equiv 0$ for all $t \ge 0$;

(3) the string has initial position $\varphi(x)$ for $0 \le x \le 1$ and $t = 0$, but its initial velocity $u_t(x,t) \equiv 0$; the string is "pulled into its initial position and then released."

Condition (2) tells us that we need "basis vectors" indexed by 0.0,0.1,...,0.9,1.0 whose coördinate (or value) at 0.0 and 1.0 is zero. A moment's thought will convince the reader that instead of the vectors $\mathbf{C}_j$ we used in considering the heat equation with insulated endpoints, we should try the vectors

$$\mathbf{S}_j = \begin{bmatrix} \sin(j\pi \cdot 0h) \\ \sin(j\pi \cdot 1h) \\ \vdots \\ \sin(j\pi \cdot 9h) \\ \sin(j\pi \cdot 10h) \end{bmatrix} \quad \text{where } h = 0.1,$$

since these have the correct value at the endpoints. However, the "dimension of the problem" will be 9 rather than 10 in this case: since $\sin(j\pi \cdot 0h) = 0 = \sin(j\pi \cdot 10h)$ holds for all $j = 0, \ldots, 10$, we can hope to get only 9 independent vectors and must thus let $j = 1, \ldots, 9$. These 9 vectors **are** independent: it is easy to check[10] that with respect to the usual inner product

$$[u_1, \ldots, u_9]^T \bullet [v_1, \ldots, v_9]^T = u_1 v_1 + \cdots + u_9 v_9$$

the vectors $\{\mathbf{S}_1, \ldots, \mathbf{S}_9\}$ are orthogonal and have length-squared equal to 5. The eigenvalues of the "discretized second space derivative" that go with the respective $\mathbf{S}_j$'s (but only for $j = 1, \ldots, 9$) are again the numbers $-\omega_j^2 = -\left(\dfrac{\sin(j\pi/20)}{1/20}\right)^2 = -400\sin^2(j\pi/20)$ respectively, as before. We can thus get solutions of the wave equation $u_{tt} = c^2 u_{xx}$ discretized into

$$\frac{d^2 u(x_i,t)}{dt^2} = c^2\, \frac{u(x_i - h) - 2u(x_i) + u(x_i + h)}{h^2}\ , i = 1, \ldots, 9$$

with initial position given by the eigenvector

$$\begin{bmatrix} u(x_0,0) \\ u(x_1,0) \\ \vdots \\ u(x_9,0) \\ u(x_{10},0) \end{bmatrix} = \mathbf{S}_j$$

by solving the single second-order scalar ordinary differential equations

$$\frac{d^2 a_j(t)}{dt^2} = c^2\,(-\omega_j^2)\, a_j(t) \qquad (\#)$$

(10) It is not even too difficult to use trig identities and check this by hand, but the verification with Maple or MATLAB is very easy.

for $j = 1, \ldots, 9$. {Recall the general setup for solving second-order homogeneous linear systems of ordinary differential equations: given the system

$$\frac{d^2\mathbf{u}(t)}{dt^2} = A\mathbf{u}(t)$$

and an eigenvector $\mathbf{v}$ of the matrix A belonging to an eigenvalue $-\mu^2$, the vector-valued function $\mathbf{u}(t) = a(t)\mathbf{v}$ will be a solution of the system if and only if the scalar function $a(t)$ satisfies $\dfrac{d^2a(t)}{dt^2} = -\mu^2\, a(t)$. This, of course, happens if and only if $a(t)$ has the form $a(t) = A\cos\mu t + B\sin\mu t$ for some constant coefficients A and B, and those can be determined from the initial conditions on $\mathbf{u}(0)$ and $\mathbf{u}'(0)$. If there are enough eigenvectors of A to form a basis of the vector space, then a complete solution of the vector initial-value problem can be determined in this way by expanding the initial position and velocity vectors in the eigenvector basis and "working term-by-term."} In this situation, we see that the general solution of (#) is

$$a_j(t) = A_j\cos(c\omega_j t) + B_j\sin(c\omega_j t)\ .$$

The initial position and velocity of this function equal A_j and $c\omega_j B_j$ respectively, so we will want to take $B_j = 0$ for all $j = 1\ldots, 9$. But now the method of solution for the discretized wave equation is clear: *mutatis mutandis*, it works just like the heat equation. Expand the initial position into a linear combination of the $\mathbf{S}_j$'s:

$$\begin{bmatrix} \varphi(x_0,0) \\ \varphi(x_1,0) \\ \vdots \\ \varphi(x_9,0) \\ \varphi(x_{10},0) \end{bmatrix} = A_1\mathbf{S}_1 + \cdots + A_9\mathbf{S}_9\ .$$

Insert the factors $\cos(c\omega_j t)$ into each term to obtain

$$\begin{bmatrix} u(x_0,t) \\ u(x_1,t) \\ \vdots \\ u(x_9,t) \\ u(x_{10},t) \end{bmatrix} = A_1\cos(c\omega_1 t)\mathbf{S}_1 + \cdots + A_9\cos(c\omega_9 t)\mathbf{S}_9\ ,$$

or—writing things as functions—

$$u(x,t) = A_1\cos(c\omega_1 t)\sin(1\pi x) + \cdots + A_9\cos(c\omega_9 t)\sin(9\pi x)\ .$$

The resulting $u(x,t)$ has initial velocity $u_t(x,0) = 0$ for all real x (a consequence of our using cosines) and satisfies the initial condition $u(x,0) = \varphi(x)$ at $x = 0.0, 0.1.\ldots, 0.9, 1.0$ but also—most importantly—satisfies the discretized wave equation

$$\frac{\partial^2 u(x,t)}{\partial t^2} = c^2\cdot\frac{u(x-h) - 2u(x) + u(x+h)}{h^2} \quad \text{at } x = 0.0, 0.1.\ldots, 0.9, 1.0.$$

One can have some additional fun with the expansion

$$u(x,t) = A_1\cos(c\omega_1 t)\sin(1\pi x) + \cdots + A_9\cos(c\omega_9 t)\sin(9\pi x)\ .$$

Products of sines and cosines are expressible in terms of sines:

$$\begin{aligned}\sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta\\ \sin(\alpha-\beta) &= \sin\alpha\cos\beta - \cos\alpha\sin\beta\\ \frac{\sin(\alpha-\beta)+\sin(\alpha+\beta)}{2} &= \sin\alpha\cos\beta\ ,\end{aligned}$$

and thus for each $j = 1, \ldots, 9$

$$\cos(c\omega_j t)\sin(j\pi x) = \frac{\sin(j\pi x - c\omega_j t) + \sin(j\pi x + c\omega_j t)}{2}\ .$$

We can thus write $u(x,t)$ in the form

$$\begin{aligned}&A_1\cos(c\omega_1 t)\sin(1\pi x) + \cdots + A_9\cos(c\omega_9 t)\sin(9\pi x) =\\ &A_1\,\frac{\sin(1\pi x - c\omega_1 t)+\sin(1\pi x + c\omega_1 t)}{2} + \cdots + A_9\,\frac{\sin(9\pi x - c\omega_9 t)+\sin(j\pi x + c\omega_9 t)}{2} =\\ &\frac{[A_1\sin(1\pi x - c\omega_1 t)+\cdots+A_9\sin(9\pi x - c\omega_9 t)] + [A_1\sin(1\pi x + c\omega_1 t)+\cdots+A_9\sin(9\pi x + c\omega_9 t)]}{2}\ .\end{aligned}$$

Term-by-term, this expression is interesting: each term $A_j\sin(j\pi x \mp c\omega_j t)$ is a sine wave which, with the passage of time, is translated up or down the x-axis at velocity $c\omega_j$. Unfortunately, the various terms travel at various speeds! However, if we were to take the limit as $h \to 0^+$ in the calculation of the ω's on pp. 6–7, we would find that

$$\lim_{h\to 0^+}\frac{\sin j\pi(x-h) - 2\sin j\pi x + \sin j\pi(x+h)}{h^2} = \left\{-\lim_{h\to 0^+}\left(\frac{\sin(j\pi h/2)}{h/2}\right)^2\right\}\sin j\pi x = -(j\pi)^2\sin x\ .$$

If we formally take this limit and consider the function

$$U(x,t) = A_1\cos(c1\pi t)\sin(1\pi x) + \cdots + A_9\cos(c9\pi t)\sin(9\pi x)$$

we will find, to our pleasure, that each term satisfies the (**un**discretized) wave equation. One can either verify this directly or observe that each term in

$$\cos(cj\pi t)\sin(j\pi x) = \frac{\sin(j\pi x - cj\pi t)+\sin(j\pi x + cj\pi t)}{2} = \frac{\sin j\pi(x-ct) + \sin j\pi(x+ct)}{2}$$

has the form of a D'Alembert solution $u(x,t) = f(x \mp ct)$. Thus in the special case of initial velocity zero and initial position given by a function of the special form[(11)]

$$\varphi(x) = A_1\sin(1\pi x) + \cdots + A_9\sin(9\pi x)$$

we have been able to return to the D'Alembert solution of the wave equation, term by term, and write

$$\begin{aligned}U(x,t) &= \sum_{j=1}^{9} A_j\,\frac{\sin j\pi(x-ct)+\sin j\pi(x+ct)}{2}\\ &= \frac{\varphi(x-ct)+\varphi(x+ct)}{2}\end{aligned}$$

as the solution of the wave equation on the finite interval $[0,1]$, with end conditions and initial conditions. We shall be able to do this for general $\varphi(x)$, but we shall have to develop the methods of **eigenfunction expansions** and more specifically **Fourier series** and **Fourier transforms** in order to do it in full generality.

(11) The fact that this sum—called a *trigonometric polynomial*—has 9 terms is of no significance; any finite linear combination of terms $\sin j\pi x$, where each j is an integer, could be handled in the same way.