HOMEWORK 8, CPSC 303, SPRING 2020
OR AN INTRODUCTION TO DIFFERENTIAL EQUATIONS

JOEL FRIEDMAN

Contents

Research supported in part by an NSERC grant.

The main goal of this homework is to give some insight into differential equations (sometimes called "diffyQ's") by exploring some of the most fundamental equations. Our focus will be on the heat equation, and then we will make some brief remarks on other ones.

This is the last homework/handout for CPSC 303 in Spring 2020; given the fall-out from COVID-19, we are likely to omit entire sections of this homework/handout.

This handout relies on the matrices studied in Homework 7; for convenience, we have included the content of Homework 7 as appendices to this article.

0. General Remarks

To understand differential equations, it is best to study individual equations and the issues that arise there. The heat equation is one of the simplest and most important differential equations, and its study reveals most of the fundamental issues in solving many types of differential equations.

Many textbooks divide differential equations into two categories, *ODE's (ordinary differential equations)* and *PDE's (partial differential equations)*, and there is nothing wrong with that. The reason one does so is that many of the same *algorithms* and *algorithmic principles* apply to all ODE's (as well as the basic existence and uniqueness theorems); on the other hand, PDE's is a vast generalization of ODE's, involving *partial derivatives* in more than one variable (typically two or more, so that PDE's are viewed as distinct from ODE's). Many classes of PDE's require algorithms of their own; i.e., there are fewer common algorithmic principles when working with PDE's.

However, to gain some understanding of ODE's and PDE's one really has to study a number of individual ODE's and PDE's.

Furthermore, since the advent of modern computers and languages such as MATLAB, it makes more sense to first study individual diffyQ's, including individual ODE's, since there are good ODE solvers that can provide crucial intuition and visualization regarding what an ODE is really saying: all ODE's and PDE's are "local" equations, such as equations in physics that tell you how a system evolves in an "infinitesimal" time step (using concepts such as *velocity* and *acceleration* which are time derivatives of *position*). However, one is really interested in knowing the "global" consequences of such equations, such as (1) whether or not there is a unique solution to the equations, and (2) what are the consequences for these systems as they evolve in time.

For example, many people are drawn to the "n-body" problem in celestial mechanics, that one can visualize by looking at the night sky and seeing the "size" of our moon (e.g., "new moon," "half moon," etc.), and by noticing the effect of seasons as the Earth rotates around the sun. In Section 9 we will give some MATLAB code to make 2-body experiments; such numerical experiments can be instructive and fun, even before you study ODE solving numerical schemes.

At this point, given the fallout from COVID-19, I suggest that if you have limited bandwith or uncertain internet access, then you should download the textbook [A&G] and the course handouts now, especially the last handout on Energy and Splines, which we refer to here as [EnergySplines]. As mentioned above, for convenience we have included Homework 7 here as appendices.

Please note: MOST OF THE HOMEWORK PROBLEMS BELOW MUST BE DONE BY HAND. You may not use MATLAB or any other computing device unless explicitly told that you may.

1. Partial Derivatives and the Heat Equation

In CPSC 303 we do not assume that you have had a third term of calculus, so in this section we briefly explain what is meant by a *partial derivative*; fortunately this is a very simple idea[1].

We need to understand only what is meant by the heat equation, which in its simplest form involves a function $u = u(x,t)$ of one "space" variable, x, and one "time" variable t, and is the equation

$$u_t = u_{xx}. \tag{1}$$

You might also see it as

$$u_t = \alpha u_{xx}, \tag{2}$$

where $\alpha > 0$ is a real constant that measures how the material involves conducts heat. However, by changing units of time appropriately, the equation (2) becomes (1).

[1] Indeed, most of a third term of calculus studies more advanced constructs based on partial derivatives (such as the *gradient* of a multivariate function), and multivariate integration, e.g., double and triple integrals.

Even if we understand the literal meaning of $u_t = u_{xx}$, it may not provide much intuition about what the equation really means; we will help develop this intuition in this article.

1.1. **Partial Derivatives.** If $u = u(x, t)$ is a function of two variables, the notation $u_t(x, t)$ means: hold x fixed, and differentiate $u(x, t)$ in t. For example:

(1) if $u(x, t) = t^2$, then $u_t(x, t) = 2t$;
(2) if $u(x, t) = x^3$, then $u_t(x, t) = 0$ (since for fixed x, $u(x, t)$ is a constant independent of t;
(3) if $u(x, t) = x^3t^2$, then $u_t(x, t) = x^3 2t$;
(4) if $u(x, t) = t^4x^3 + xe^{5t}$, then $u_t(x, t) = 4t^3x^3 + x(5e^{5t})$; and
(5) if $u(x, t) = \sin(x)e^{-t}$, then $u_t = -\sin(x)e^{-t}$;

Instead of $u_t = u_t(x, t)$ you might also see the notation

$$\frac{\partial u}{\partial t}(x, t), \quad \frac{\partial}{\partial t}u(x, t), \quad u_2(x, t)$$

(the 2 indicates differentiating with respect to the second variable); you may also see the (x, t) dropped.

One similarly defines u_x; $u_{xx} = (u_x)_x$ is the second derivative in x with t held fixed. For example:

(1) if $u(x, t) = x^4t^4$, then $u_x = 4x^3t^4$, and $u_{xx} = 12x^2t^4$;
(2) if $u(x, t) = \sin(x)e^{-t}$, then $u_x = \cos(x)e^{-t}$; and $u_{xx} = -\sin(x)e^{-t}$.

The above computations show that if $u(x, t) = \sin(x)e^{-t}$, then

$$u_t = -\sin(x)e^{-t} \quad \text{and} \quad u_{xx} = -\sin(x)e^{-t}$$

and therefore $u(x, t)$ satisfies the heat equation for all $x, t \in \mathbb{R}$.

Exercise 1.1. Which of the following functions satisfy the heat equation $u_t(x, t) = u_{xx}(x, t)$ for all $x, t \in \mathbb{R}$?

1.1(a) $u(x, t) = 3 + 4x$.
1.1(b) $u(x, t) = 3 + 4x + x^2$.
1.1(c) $u(x, t) = 3 + 4x + x^2 - t$.
1.1(d) $u(x, t) = 3 + 4x + x^2 - 2t$.
1.1(e) $u(x, t) = \sin(\omega x)e^{-t\omega^2}$ for any constant $\omega \in \mathbb{R}$.
1.1(f) $u(x, t) = \cos(\omega x)e^{-t\omega^2}$ for any constant $\omega \in \mathbb{R}$.
1.1(g) $u(x, t) = \sinh(\omega x)e^{t\omega^2}$ for any constant $\omega \in \mathbb{R}$ (where $\sinh(x) \stackrel{\text{def}}{=} (e^x - e^{-x})/2$) is the *hyperbolic sine*.

Exercise 1.2. For which $(x, t) \in \mathbb{R}^2$ is it true that $u_t(x, t) = u_{xx}(x, t)$ for the function:

1.2(a) $u(x, t) = 4x + t^2$?
1.2(b) $u(x, t) = 4x^3 + t^2$?
1.2(c) $u(x, t) = 2x^2 + 3t^3$?
1.2(d) $u(x, t) = 2x^2 - 3t^3$?

Exercise 1.3. Show that if $u(x, t), v(x, t)$ both satisfy the heat equation at some point $(x_0, t_0) \in \mathbb{R}^2$, and $\alpha, \beta \in \mathbb{R}$, then

$$w = w(x, t) \stackrel{\text{def}}{=} \alpha u(x, t) + \beta v(x, t)$$

also satisfies the heat equation at (x_0, t_0).

1.2. **The Dirichlet Problem (i.e., Initial Value Problem with Dirichlet Conditions).** Just as with splines, the heat equation holds in a certain region, and there are "boundary conditions" that determine a unique solution. Another example of this is in basketball: the trajectory of a basketball when thrown is typically modeled by an ODE, but usually depends on its initial position, velocity, and spin (angular momentum), regardless of the particular ODE used to model its trajectory (which may include gravity, air resistance, etc.).

Let us describe what is usually called the "Dirichlet problem" in the one-dimensional heat equation.

One envisions a thin, straight rod or wire that is insulated (by air or a rubber coating) along its length, and whose ends are held at a given temperature. Although the rod or wire lies in three-dimensional space, we model the rod by the interval $[0,1]$ (we assume the rod is of unit length in our units), that at time $t=0$ has a known temperature profile, and that for time $t>0$ satisfies the one-dimensional heat equation. The usual "Dirichlet problem" is the PDE that describes the temperature of the rod, $u(x,t)$, for all $x\in[0,1]$ and $t\geq 0$, given by

(1) $u(x,0)=f(x)$ for $x\in[0,1]$, where $f(x)$ is a given (usually known to us) *initial value*
(2) $u(0,t)=u(1,t)=0$ for all $t>0$, the *(zero-valued) Dirichlet problem*, and
(3) $u_t(x,t)=u_{xx}(x,t)$ for all $0<x<1$ and $t>0$ (the heat equation).

A *more general Dirichlet problem* allows for the temperature of the each endpoint of the rod to be held at some fixed real number (which are generally different at the endpoints); this is called *Dirichlet data*. The reason that we are content to take the temperature of the endpoints of the rod to be 0—which is the usual *Dirichlet problem*—is explained by the exercise below.

Exercise 1.4. Let $R\subset\mathbb{R}^2$ be a subset ("R" is to suggest the word "region"). Let $c_0,c_1\in\mathbb{R}$, and $u\colon\mathbb{R}^2\to\mathbb{R}$ be any function, and let

$$v(x,t)=u(x,t)-(c_0+c_1x).$$

1.4(a) Show that u satisfies the heat equation in R, i.e., $u_t(x,t)=u_{xx}(x,t)$, for all $(x,t)\in R$, iff v satisfies the heat equation in R.

1.4(b) Show that for any $a,b\in\mathbb{R}$ there is a unique c_0,c_1 such that $c_0=a$ and $c_0+c_1=b$.

1.4(c) Show that if we replace the Dirichlet problem by more general Dirichlet data:

$$u(0,t)=a,\quad u(1,t)=b$$

for some constants $a,b\in\mathbb{R}$, it is equivalent to substitute

$$v(x,t)=u(x,t)-(c_0+c_1x)$$

(with c_0,c_1 as above) and solve the (zero-valued) Dirichlet problem above, replacing $f(x)$ with $f(x)-c_0-c_1x$. [Remark: for simplicity we took $u\colon\mathbb{R}^2\to\mathbb{R}$ instead of $u\colon[0,1]\times[0,\infty)\to\mathbb{R}$; should this bother us?]

2. Basic Properties of the Heat Equation

In this section we will describe a number of important properties of the heat equation. We will start with its numerical solution, which is the focus of this homework set.

2.1. **The "Meaningful" Version of the Heat Equation.** If $f\colon \mathbb{R}\to\mathbb{R}$ is sufficiently differentiable, we have (see [A&G], Section 14.1, and Subsections 7.3 and 7.2 of this article):

$$\frac{f(x_0+k)-f(x_0)}{k}=f'(x_0)+O(k) \tag{3}$$

for small k, where we have used the "order" notation on page 7 of [A&G][2] and

$$f''(x_0)=\frac{f(x_0-h)-2f(x_0)+f(x_0+h)}{h^2}+O(h^2) \tag{4}$$

for small h (similarly for $O(h^2)$); note that the error term is $O(h^2)$ rather than $O(h)$, due to the fact that the numerator is the special "centred formula for the second derivative" ([A&G], middle of page 412).

Assuming that $u=u(x,t)$ is sufficiently differentiable near some point (x_0,t_0), we have that for some

$$u_t(x_0,t_0)=\frac{u(x_0,t_0+k)-u(x_0,t_0)}{k}+O(k)$$

and

$$u_{xx}(x_0,t_0)=\frac{u(x_0+h,t_0)+u(x_0-h,t_0)-2u(x_0,t_0)}{h^2}+O(h^2).$$

It follows that $u_t=u_{xx}$ at (x_0,t_0) implies that

$$\frac{u(x_0,t_0+k)-u(x_0,t_0)}{k}+O(k)=\frac{u(x_0+h,t_0)+u(x_0-h,t_0)-2u(x_0,t_0)}{h^2}+O(h^2),$$

which we may write as

$$u(x_0,t_0+k)=u(x_0,t_0)+\frac{k}{h^2}\bigl(u(x_0+h,t_0)+u(x_0-h,t_0)-2u(x_0,t_0)\bigr)+k\bigl(O(k)+O(h^2)\bigr). \tag{5}$$

We refer to (5) as the *meaningful heat equation*. The reason is that the expression

$$u(x_0+h,t_0)+u(x_0-h,t_0)-2u(x_0,t_0) \tag{6}$$

lends itself to many different "meaningful" interpretations, e.g.,

(1) it is twice
$$\frac{u(x_0+h,t_0)+u(x_0-h,t_0)}{2}-u(x_0,t_0),$$
which is the difference between $u(x_0,t_0)$ the average value of u at the "h-neighbours" (in space) of (x_0,t_0);
(2) it is $2h^2$ times the divided difference $f[x_0-h,x_0,x+h]$ where $f(x)=u(x,t_0)$;
(3) it reflects an exchange of heat between x_0 and its neighbour x_0+h, and between x_0 and its other neighbour x_0-h (for a fixed time t_0);
(4) it is a *centred second difference* in a sense that we will discuss in Section 3;
(5) etc.

These interpretations make sense for many resources that can be "shared" or "transferred" in a network or *undirected graph*.

[2] A function of k is written "$O(k)$ for small k" if for sufficiently small $|k|\neq 0$ it is bounded by a constant times $|k|$ (independent of any other variables that may appear in the equation except those variables that have been fixed).

2.2. **Numerical Solution of the Heat Equation via the Meaningful Form.** To solve the Dirichlet problem of Subsection 1.2, we use the meaningful heat equation (5) as follows: we choose integers $n, m \in \mathbb{N}$, set $h = 1/(n+1)$ and k/m (the reason for choosing h as such and not as $1/n$ will become clear in Section 4); we then set subdivide $[0,1]$ (as we might for splines) by setting

$$x_0 = 0,\ x_1 = h,\ x_2 = 2h,\ \ldots, x_n = nh,\ x_{n+1} = (n+1)h = 1$$

and we set

$$t_0 = 0,\ t_1 = k,\ t_2 = 2k,\ \ldots$$

Note that in numerical calculations we often fix a finite time to stop the calculation, say $t \le T$; if so, we can assume that T is a multiple of k and we compute up to t_m where $m = T/k$.

The basic numerical solution of the heat equation is to:

(1) set $u(x_i, t_0) = u(x_i, 0) = f(x_i)$ for $i = 1, \ldots, n$ (the values of $u(0,0)$ and $u(1,0)$ are irrelevant to our algorithm)
(2) we determine $u(x_i, k)$ via the meaningful formula, i.e.,

$$u(x_i, t_1) = u(x_i, t_0) + \frac{k}{h^2}\big(u(x_i + h, t_0) + u(x_i - h, t_0) - 2u(x_i, t_0)\big)$$

(3) and for $j = 1, 2, \ldots$ we determine all the values $u(x_i, t_{j+1})$ from all the values $u(x_i, t_j)$ as

$$u(x_i, t_{j+1}) = u(x_i, t_j) + \frac{k}{h^2}\big(u(x_i + h, t_j) + u(x_i - h, t_j) - 2u(x_i, t_j)\big).$$

There is another remarkable aspect of the "meaningful heat equation" is that one can use it to show that this scheme always works provided that $h \to 0$, and that $k/h^2 \le 1$; in fact, a standard higher order Taylor expansion our derivative approximations (see Section 7) indicates that one gets a "higher order" agreement in the approximation above when choosing $k/h^2 = 1/3$.

2.3. **Harmonics of a One-Dimensional Rod (or String).** There are some special solutions to the Dirichlet problem for the heat equation in Subsection 1.2. Namely, for each $m \in \mathbb{Z}$, we have that

$$u_m(x,t) = \sin(\pi m x) e^{-m^2 \pi^2 t}$$

solves the Dirichlet problem with initial condition

$$u_m(x, 0) = \sin(\pi m x).$$

These "sine waves" $u_m(x,0)$ are called the "harmonics" of $[0,1]$; when we study the wave equation (i.e., $u_{tt} = u_{xx}$) we see that their existence was essentially known to the ancient Greeks who experimented with string musical instruments (e.g., think of a guitar or a violin). The essential property of the function $\sin(\pi m x)$ is that it vanishes at $x = 0$ and $x = 1$ and is an "eigenvalue of the Laplacian operator" (meaning the map $u \mapsto u_{xx}$).

2.4. **Existence of Solution.** We can use *Fourier analysis* to give a "formula" for the solution ot the Dirichlet problem. Namely, if we can write the initial condition

$$u(x,0)=f(x)$$

as a sum

$$\sum_{m=1}^{\infty} a_m \sin(\pi m x)$$

for some $a_m \in \mathbb{R}$, then the solution to the Dirichlet problem is merely

$$u(x,t)=\sum_{m=1}^{\infty} a_m u_m(x,t) \sum_{m=1}^{\infty} a_m \sin(\pi m x) e^{-m^2\pi^2 t}. \tag{7}$$

The study of how to find the a_m and what one can say about the a_m under reasonable assumptions on f is called *Fourier analysis*.

Fourier analysis implies that under mild assumptions on f:

(1) if $f(x)$ is non-negative and nonzero, then $a_1 > 0$, and the dominant term in $u(x,t)$ is $a_1 \sin(\pi x) e^{-\pi^2 t}$;
(2) $u(x,t)$ is infinitely differentiable near any $t > 0$ and $x \in (0,1)$;
(3) if you touch a guitar string in the middle and strum it elsewhere, then $a_1 = 0$ and you produce a note that is one octave higher, and similarly with higher harmonics (think of "pinch harmonics" in guitar, Eddie van Halen, Billy Gibbons, and earlier guitarists, etc., or the equivalent for a violin);
(4) etc.

One also sees from (7) that if you run the heat equation backwards in time—or equivalently have a heat equation $u_t = \alpha u_{xx}$ with $\alpha < 0$ (i.e, the hot get hotter and the cold get colder)—then you get an "unbounded" operator with very "ill-behaved" properties.

2.5. **The Maximum Principle and Uniqueness.** The *weak maximum principle* for the heat equation is not hard to prove, and this proves that there is a unique continous solution to the heat equation, provided that the solution is sufficiently differentiable.

Lemma 2.1 (The Weak Maximum Principle)**.** *Let $u(x,t)$ be a continuous function on $[0,1]\times[0,T]$ for some $T\in\mathbb{R}$, and that u_x, u_{xx}, u_t exist throughout the region $R=(0,1)\times(0,T)$ and that u satisfies the heat equation $u_t = u_{xx}$ throughout R. Say that for some $M \in \mathbb{R}$ we have $u(x,t) \le M$ for (x,t) on the "heat equation boundary" of $[0,1]\times[0,T]$, i.e., for all (x,t) in*

$$\partial_{\mathrm{heat}} R \stackrel{def}{=} \{(x,t)\in[0,1]\times[0,T] \mid x = 0,1 \text{ or } t=0\}.$$

Then $u(x,t) \le M$ for all $x \in [0,1]\times[0,T]$.

Note that the "heat equation boundary" above does not include $t = T$. We remark that this lemma is intuitively believable, since if a house is at most 30°C both initially and everywhere along its boundary (and you don't have internal air conditioning), then it seems plausible that at any point in time your house will be at most 30°C at any point in time, anywhere in the house.

Proof. Otherwise $u(x_0,t_0) > M$ for some $0 < t < T$ and $x_0 \in (0,1)$, and hence $u(x_0,t_0) \geq M + \epsilon$ for some $\epsilon > 0$. Let

$$v(x,t) = M - (\epsilon/2)\frac{(x-x_0)^2 + t - t_0}{B},$$

where B is an upper bound $(x-x_0)^2 + t - t_0$ in $[0,1] \times [0,t_0]$. It follows that $v(x_0,t_0) = u(x_0,t_0)$, and $v(x,t)$ is at most $M+\epsilon/2$ on the "heat equation boundary" of $[0,1] \times [0,t]$ (which excludes $(0,1) \times \{t\}$). Hence the maximum of v is attained for some point (x_1,t_1) not on the boundary of $[0,1] \times [0,t_0]$. But in this case $v_x(x_1,t_1) = 0$, $v_{xx}(x_1,t_1) \leq 0$ and $v_t \geq 0$, and so

$$v_t(x_1,t_1) - v_{xx}(x_1,t_1) \geq 0. \tag{8}$$

But visibly

$$v_t - v_{xx} = u_t - u_{xx} - (\epsilon/2)\frac{2}{B} = -(\epsilon/2)\frac{2}{B} < 0,$$

throughout $[0,1] \times [0,t]$, which contradicts (8) □

Corollary 2.2. *If a solution exists to the Dirichlet problem of Subsection 1.2, then it is unique.*

Proof. If u, v are two solutions to the same Dirichlet problem, then $w = u - v$ is a solution to the Dirichlet problem with zero initial condition, i.e., $w(x,0) = 0$ for all $x \in (0,1)$. Then $w \leq 0$ everywhere on the "heat equation boundary" of $R = [0,1] \times [0,\infty)$ so $w \leq 0$ throughout R. Similarly $-w \leq 0$ throughout R, and hence $w = 0$ throughout R. Since $w = u - v$, it follows that $u = v$ throughout R. □

2.6. **Good News for the Heat Equation in Positive Time, Bad News in Negative Time.** There are a number of reasons why we solve the "initial value problem" for the heat equation, i.e., specify the heat profile at time $t = 0$, and consider the profile for $t > 0$: (1) we are interested in predicting the future rather than explaining the past, and (2) the profile for $t > 0$ becomes infinitely differentiable, whereas studying the profile for $t < 0$ is a poorly-behaved situation. The fact that the heat profile is infinitely differentiable allows us to use all the derivative estimates in [A&G], Chapter 14, and will lead us to the usual "optimal" choice of spacing for the x (space) grid points versus the t (time) grid points.

The only setting I know of where one studies the heat equation in backwards time is from a recent talk I heard by T. Tao (Mathematics Department, UCLA) regarding a weakened version of the Riemann Hypothesis, at the IPAM (Institute of Pure and Applied Mathematics, in UCLA); see [xxxxx] for details.

Certainly, if you run the heat equation backwards, corresponding to $u_t = \alpha u_{xx}$ with $\alpha < 0$, you are taking from your neighbours when you should be giving and vice/versa, and even for small time only "very rare" initial conditions give bounded solutions that make any classical sense.

2.7. **The Heat Equation in Higher Dimensions.** Talk a bit about

$$u_t = \Delta u,$$

where

$$\Delta u = u_{x_1 x_1} + \cdots u_{x_n x_n}$$

is the n-dimensional Laplacian.

3. Differences, Multivariate Differences, and Heat Transfer

Recall that if $U\colon \mathbb{R}\to\mathbb{R}$ is a function, then we defined the divided differences of U, and we have

$$\begin{aligned} U[x_0] &= U(x), \\ U[x_0,x_1] &= \frac{U[x_1]-U[x_0]}{x_1-x_0}, \\ U[x_0,x_1,x_2] &= \frac{U[x_2,x_1]-U[x_1,x_0]}{x_2-x_0}. \end{aligned}$$

In Chapter 10 of [A&G] we learned many remarkable properties of divided differences.

If we only take x_0,x_1,x_2 to be consecutive integers, then the formulas simplify a bit: for any $i\in\mathbb{Z}$, it is not hard to see that

$$\begin{aligned} U[i] &= U(i) \\ U[i,i+1] &= U(i+1)-U(i) \\ U[i-1,i,i+1] &= \frac{U(i+1)+U(i-1)}{2}-U(i). \end{aligned}$$

In CPSC 303 this year, we discussed *differences* without dividing, and in Homework 4, for each sequence

$$\mathbf{y}=\{y_n\}_{n\in\mathbb{Z}}=\{\ldots,y_{-1},y_0,y_1,y_2,\ldots\}$$

we defined a new sequence $D\mathbf{y}$ (known as the "(forward) difference of $\mathbf{y}$") defined by

$$(D\mathbf{y})_n = y_{n+1}-y_n;$$

D can be described more fundamentally as $D=\sigma-1$, where σ is the "shift up by one" operator

$$(\sigma\mathbf{y})_n = y_{n+1}$$

and 1 is the identity operator

$$(1\mathbf{y})_n = y_n.$$

We tend to simply write Dy and σy for $D\mathbf{y}$ and σy.

Note that $D=\sigma-1$ and $\sigma=D+1$, so that either D,σ can be expressed in terms of the other. [Later we note that this implies that these operators *commute*.]

In this article it will be helpful to introduce some generalizations of σ and D.

3.1. **Fractional Shifts and Laurant Polynomials.**

First, note that

$$(\sigma^2 y)_n = \bigl(\sigma(\sigma y)\bigr)_n = (\sigma y)_{n+1} = \sigma y_{n+2},$$

and so σ^2 is merely the operator "shift up by two." Similarly σ has an inverse operator σ^{-1}, and this is nothing more than "shift up by -1," or, equivalently "shift down by one." For this reason the following definition works.

Definition 3.1. If $h\in\mathbb{R}$ is any real number, and $f\colon\mathbb{R}\to\mathbb{R}$, we define $\sigma^h f$ to be the function

$$(\sigma^h f)(x) \stackrel{\text{def}}{=} f(x+h)$$

and refer to σ^h as "shift up by h" or "the h-th power of σ."

We see that $(\sigma^a)^k = \sigma^{ak}$ for any positive or negative integer k and real a. [There is no harm in defining $(\sigma^a)^k$ for any real k as σ^{ak}.]

We also claim that since all powers of σ commute, i.e., $\sigma^a\sigma^b = \sigma^{a+b} = \sigma^b\sigma^a$, if we have any identity in a variable x and $x^{-1} = 1/x$ (formally assuming that $x \neq 0$), then the same holds for σ replacing x. For example, since

$$x^{-1}(1+x)(1-x) = x^{-1}(1-x^2) = x^{-1} - x,$$

we have

$$\sigma^{-1}(1+\sigma)(1-\sigma) = \sigma^{-1}(1-\sigma^2) = \sigma^{-1} - \sigma.$$

3.2. **Why Differences and not Divided Differences for Heat Transfer.** The reason that we prefer to work with differences rather than divided differences is subtle; the rough idea is that divided differences are symmetric, e.g.,

$$U[i+1,i] = \frac{U(i) - U(i+1)}{i-(i+1)} = \frac{U(i+1) - U(i)}{(i+1)-i} = U[i,i+1].$$

However, if you imagine two apartments that share a wall (that is not perfectly insulated and allows for some heat transfer), then if one apartment is at 20°C and the other at 30°C, then the first apartment will gain some heat from the second, and the second will lose some heat. So operators involving heat transfer should be *antisymetric*, i.e., the heat transfer from A to B along a wall should be *minus* that from B to A.

Of course, this idea of heat transfer applies to a number of resources in networks or *graphs* where the resource can be shared.

3.3. **The Centred Second Difference Operator.** For the heat equation, both the (usual) continuous version, and the discrete version that we first study in the article, we will need the *centred second difference operator*

$$\big(D^{2,\mathrm{centre}}y\big)(i) = \big((\sigma-1)(1-\sigma^{-1})y\big)(i) = \big((\sigma+\sigma^{-1}-2)y\big)(i) = y_{i+1} + y_{i-1} - 2y_i.$$

3.4. **Multivariate Shifts and Differences.** We will need one additional piece of notation in this article: if $U = U(i,j)$ is a function of two variables, i, j, we use σ_i for the shift in the variable i, i.e.,

$$\sigma_i U(i,j) = U(i+1,j),$$

and similarly

$$\sigma_j U(i,j) = U(i,j+1), \quad D_i U(i,j) = U(i+1,j) - U(i,j), \quad D_j U(i,j) = U(i,j+1) - U(i,j).$$

Similarly we have

$$D_i^{2,\mathrm{centre}} = (\sigma_i - 1)(1-\sigma_i^{-1})U(i,j) = (\sigma_i + \sigma_i^{-1} - 2)U(i,j) = U(i+1,j) + U(i-1,j) - 2U(i,j),$$

3.5. **EXERCISES.** Recall the definitions of $\sigma, D, D^{2,\mathrm{centre}}$.

Exercise 3.1. Show that

$$D^{2,\mathrm{centre}} = (\sigma-1)(1-\sigma^{-1}) = \sigma + \sigma^{-1} - 2.$$

Exercise 3.2. Show that

$$D^{2,\mathrm{centre}} = \sigma^{-1}D^2.$$

Exercise 3.3. Say that instead of considering sequences

$$\mathbf{y} = \{y_n\}_{n\in\mathbb{Z}} = \{\ldots, y_{-1}, y_0, y_1, y_2, \ldots\}$$

we consider only the "truncated" finite sequences

$$\mathbf{y} = (y_1, \ldots, y_n)$$

(you might think of the doubly infinite sequence where we enforce that $y_i = 0$ for all $i > n$ and all $i < 1$). Say that we define σ_{trunc} to be the operator

$$\sigma_{\mathrm{trunc}}(y_1, \ldots, y_n) \stackrel{\mathrm{def}}{=} (y_2, y_3, \ldots, y_n, 0).$$

Show that in matrix form, we have

$$\sigma_{\mathrm{trunc}} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = S_{n,1} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

where $S_{n,1}$ is the matrix given in Homework 7 (and in the appendices here).

4. Discrete Heat Transfer Between Neighbours

In this section we describe the discrete version of the heat equation that we will use to solve the exact heat equation.

4.1. **Motivation of the Discrete Heat Equation.** Imagine that for some $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$ you have n "row houses" as in Figure 1.

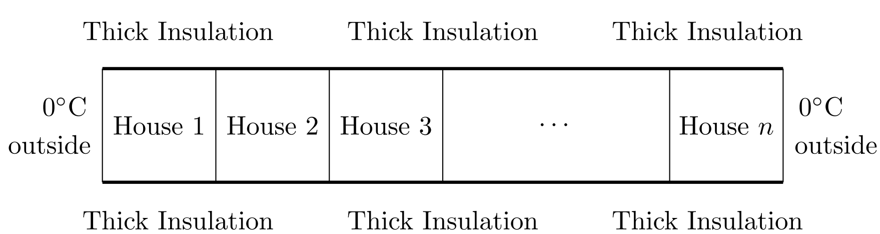

Figure 1. n row houses, insulated except at interfaces between houses.

Let us write some mathematical formulas to expres how the temperature in the n houses in Figure 1 changes over time.

For $i = 1, \ldots, n$ and $j = 0, 1, 2, \ldots$, let $\mathrm{Temp}(i, j)$ denote the average temperature in house i at time j. We will determine $\mathrm{Temp}(i, j)$ by the following assumptions:

(1) At time 0, we have

$$\mathrm{Temp}(i, 0) = f(i) \tag{9}$$

where $f(i)$ is a function that we know.

(2) The temperature outside is $0 \deg C$. From time 0 and onward, all houses have lost any source of internal heating (or cooling) that they have.

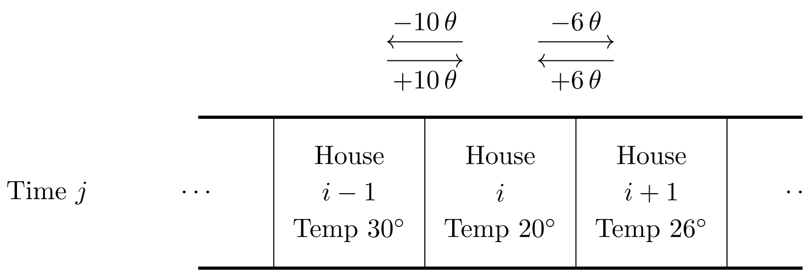

FIGURE 2. Heat Transfer at House i and time j.

(3) Each house is well-insulated, except at the interfaces between two adjacent houses, i.e., House i and House $i+1$, where each causes the other temperature to change by $\theta > 0$ times the difference; furthermore the effect on House i is the sum of the effect from House $i+1$ and House $i-1$. An example with $i \geq 2$ and $i \leq n-1$ is given in Figure 2.

(4) In addition, Houses 1 and n have an interface to the outside, which is always $0°$ centigrade/Celcius; any effect of Houses 1 and n on the temperature outside is neglible.

Now let us derive the "discrete heat equation" and its boundary conditions. First of all, (9) is a "boundary condition," specifically an "initial condition," since it specifies what the temperature profile looks like at time 0. The idea is now to determine the temperature profile for time $j = 1$ based on that at time $j = 0$, and then determine the profile at time $j = 2$ based on that at time $j = 1$, etc.

So imagine that for some j we know $\mathrm{Temp}(i, j)$ for all $i = 1, \ldots, n$, and let us determine $\mathrm{Temp}(i, j+1)$ for all i.

First, for any $i = 2, \ldots, n-2$, House i has neighbouring houses on either side, and the rule about heat transfer says:

$$\mathrm{Temp}(i, j+1) = \mathrm{Temp}(i, j) + \theta\, D_i^{2,\mathrm{centre}} \mathrm{Temp}(i, j), \tag{10}$$

where we recall that $D_i^{2,\mathrm{centre}} \mathrm{Temp}(i, j)$ is given by any of the equivalent expressions:

(1)
$$\theta\Big(\mathrm{Temp}(i+1, j) - \mathrm{Temp}(i, j)\Big) + \theta\Big(\mathrm{Temp}(i-1, j) - \mathrm{Temp}(i, j)\Big);$$

(2)
$$\theta\Big(\mathrm{Temp}(i+1, j) + \mathrm{Temp}(i-1, j) - 2\,\mathrm{Temp}(i, j)\Big);$$

(3) $2g[i-1, i, i+1]$, where g is the function defined by $g(i) = \mathrm{Temp}(i, j)$ viewing j as fixed;

(4) two times the second ("centred") divided difference of $\mathrm{Temp}(i, j)$ as a function of i, with j fixed;

(5) the second ("centred") difference of $\mathrm{Temp}(i, j)$ as a function of i, with j fixed.

From our knowledge of divided differences, this implies that the temperature second difference would be an approximate second derivative in x of $f(x) = \mathrm{Temp}(x, j)$

(with j fixed), if the temperature were defined with x taking continuous values on $[0, n] \subset \mathbb{R}$.

Second, to deal with (10) the cases $i = 1$ and $i = n$, it becomes convenient to introduce a fictitious House 0 to the left of House 1, and a fictitious House $n+1$ to the right of House n, such that these fictitious houses are always at 0°C; see Figure 3. In this case (10) also holds for $i = 1$ and $i = n$, where we set

Thick Insulation Thick Insulation Thick Insulation

House 0 Always 0°C	House 1	House 2	House 3	$\cdots$	House n	House $n+1$ Always 0°C

Thick Insulation Thick Insulation Thick Insulation

FIGURE 3. Introducing fictitious Houses 0 and $n+1$ for convenience.

$$\text{Temp}(0, j) = \text{Temp}(n, j) \quad \forall j \geq 1. \tag{11}$$

4.2. **Precise Statement of the Discrete Heat Equation.**

Definition 4.1. Let $n \in \mathbb{N} = \{1, 2, \ldots\}$ and let $\mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$ be the non-negative integers. We say that a function $U \colon \{0, 1, \ldots, n, n+1\} \times \mathbb{Z}_{\geq 0} \to \mathbb{R}$ satisfies the *discrete heat equation* if for all $i = 2, \ldots, n-1$ and $j \in \mathbb{Z}_{\geq 0}$ we have

$$U(i, j+1) = U(i, j) + \theta\, D_i^{2,\text{centre}} U(i, j), \tag{12}$$

where

$$D_i^{2,\text{centre}} U(i, j) = U(i+1, j) + U(i-1, j) - 2U(i, j),$$

i.e., the second (centred) difference of $U(i, j)$ as a function of i, with j held fixed. If $f \colon [n] \to \mathbb{R}$ is any function (here $[n] = \{1, 2, \ldots, n\}$), we say that U satisfies the *initial condition* f, if

$$U(i, 0) = f(i) \quad \text{for } i \in [n].$$

We say that U satisfies the *zero Dirichlet condition*, or simply the *Dirichlet condition*, if (12) also holds for $i = 1$ and $i = n$ (and all $j \geq 0$) provided that we have

$$U(0, j) = U(n+1, j) = 0 \quad \text{for all } j = 1, 2, \ldots.$$

4.3. **EXERCISES.**

Exercise 4.1. Consider the discrete heat equation in the case $n = 1$; therefore there is only House 1, and it is incident upon a fictitious House 0 and House 2 that are always at 0°C.

4.1(a) Show that for any $j \geq 0$, we have

$$\text{Temp}(1, j+1) = (1 - 2\theta)\text{Temp}(j, 1).$$

4.1(b) Given the initial temperature $\text{Temp}(1, 0) = f(1)$ of House 1, describe $\text{Temp}(1, j)$ as a function of $f(1)$ and j.

4.1(c) Show that for $0 < \theta \leq 1/2$, and $f(1) > 0$, $\text{Temp}(1, j)$ is positive for all j and $\text{Temp}(1, j) \to 0$ as $j \to \infty$.

4.1(d) If $1/2 < \theta < 1$ and $f(1) > 0$, how does the sign of $\text{Temp}(1, j)$ behave?

4.1(e) If $\theta < 0$ and $f(1) > 0$, what happens to $\mathrm{Temp}(1, j)$ as $j \to \infty$?

Exercise 4.2. Consider the case of $n = 1$ in Exercise 4.1, with $\mathrm{Temp}(1, 0) = f(1)$ a fixed value. Fix an integer $T > 0$. For any $m \in \mathbb{N}$, let $\theta = 1/m$ and set $g(m) = \mathrm{Temp}(1, Tm)$ (with $\theta = 1/m$). Show that

$$\lim_{m\to\infty} g(m) = e^{-2\theta T} f(1).$$

[Hint: first show that $g(m) = (1 - 2/m)^{Tm}$. Then apply ln to both sides; at this point you can use l'Hôpital's rule or Taylor's theorem.]

Exercise 4.3. Let $n = 2$ and $\theta \in \mathbb{R}$ be arbitrary.

4.3(a) Show that if for some j we have $\mathrm{Temp}(1, j) = \mathrm{Temp}(2, j)$, then

$$\mathrm{Temp}(1, j + 1) = \mathrm{Temp}(2, j + 1) = (1 - \theta)\mathrm{Temp}(1, j) = (1 - \theta)\mathrm{Temp}(2, j)$$

4.3(b) Assume that $f\colon \{1, 2\} \to \mathbb{R}$ satisfies $f(1) = f(2)$, and let $\mathrm{Temp}(i, 0) = f(i)$ for $i = 1, 2$. Show that

$$\mathrm{Temp}(1, j) = \mathrm{Temp}(2, j) = (1 - \theta)^j f(1).$$

4.3(c) Say that $f(0) = f(1) > 0$, and that $0 < \theta < 1$. What is

$$\lim_{j\to\infty} \mathrm{Temp}(1, j)$$

(and justify your answer)?

4.3(d) Say that $f(1) = f(2) > 0$, and that $\theta < 0$. What is

$$\lim_{j\to\infty} \mathrm{Temp}(1, j)$$

(and justify your answer)?

Exercise 4.4. Put something here about the discrete Fourier exact solution, when the time comes...

5. Rod Matrices in the Discrete Heat Transfer

In this section we give a simple formula—which is not necessarily so easy to use in practice—to solve the discrete heat trasfer equation.

Let us state the result; in the exercises we write out a special case which should convince you of the general result.

Exercise 5.1. Say that for all j we gather all the temperatures at time j into a vector,

$$\mathrm{TempProfile}(j) = \begin{bmatrix} \mathrm{Temp}(1, j) \\ \mathrm{Temp}(2, j) \\ \vdots \\ \mathrm{Temp}(n, j) \end{bmatrix} \tag{13}$$

5.1(a) Show that the discrete heat equation implies that for all $j \geq 0$ we have

$$\mathrm{TempProfile}(j + 1) = \Big(I + \theta\big(N_{\mathrm{rod},n} - 2I\big)\Big)\mathrm{TempProfile}(j),$$

where $N_{\mathrm{rod},n}$ is the matrix given in Appendix A.1.

5.1(b) Show that for any $j \geq 0$ we have

$$\mathrm{TempProfile}(j) = \Big(I + \theta\big(N_{\mathrm{rod},n} - 2I\big)\Big)^j \mathrm{TempProfile}(0).$$

5.1. (MORE) EXERCISES.

Exercise 5.2. Let $n = 2$ and $\theta = 1/3$, and let TempProfile be as in (13).

5.2(a) Show that for any $j \geq 0$

$$\mathrm{TempProfile}(j) = \left(\frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)^{j} \mathrm{TempProfile}(0)$$

5.2(b) Find a formula for TempProfile(j) for $j \geq 1$ in terms of $\mathrm{Temp}(1,0) + \mathrm{Temp}(2,0)$.

6. Variants of the Discrete Heat Equation

In this section we describe variants of the discrete heat equation.

Exercise 6.1. Imagine a variant of the n-House discrete heat transfer equation, where the left side of House 1 is connected to the right side of House n, instead of each connected to a $0°$C heat absorbing sink[3]. In this case the discrete heat equation is the same, except that we identify House 0 with House n and House $n+1$ with House 1, i.e., we work with house numbers "modulo n." More precisely we define

$$\mathrm{Temp}_{\mathrm{ring}}(i,j) = \left\{ \begin{array}{ll} \mathrm{Temp}(n,j) & \text{if } i = 0, \\ \mathrm{Temp}(i,j) & \text{if } 1 \leq i \leq n-1, \text{ and} \\ \mathrm{Temp}(1,j) & \text{if } i = n+1. \end{array} \right\}$$

and (10) becomes

$$\mathrm{Temp}_{\mathrm{ring}}(i,j+1) = \mathrm{Temp}_{\mathrm{ring}}(i,j) + \theta\, D_i^{2,\mathrm{centre}} \mathrm{Temp}_{\mathrm{ring}}(i,j).$$

Show that with notation as in (13), we have

$$\mathrm{TempProfile}(j) = \Big(I + \theta\big(N_{\mathrm{ring},n} - 2I\big)\Big)^{j} \mathrm{TempProfile}(0),$$

where $N_{\mathrm{ring},n}$ is the matrix given in Appendix A.2.

7. The "Meaningful" Version of The Heat Equation in One Dimension –DRAFT FORM

THIS SECTION IS IN DRAFT FORM AND NEEDS REVISION!!!

We will focus on the heat equation in one-dimension, over a one dimensional rod, $[0,1]$, whose temperature on either end is held at $0°$C. In this section we formulate this precisely, and then give a version of the heat equation—which we call the "meaningful" version—that is both (1) easier to understand intuitively, and (2) gives us a numerical solution. In fact, the numerical solution amounts to approximating the temperature at equally spaced points

$$x_0 = 0, x_1 = h, x_2 = 2h, \ldots, x_{n+1} = (n+1)h = 1,$$

where n is large (as large as we can take within our computational resources), and $h = 1/(n+1)$. This "discretization" is analogous to what we did with interpolation and splines.

---

[3] Perhaps n is very large and the houses are in a large "ring" in 2-dimensions, which is more simply modeled by a thin, 1-dimensional loop. Or maybe what seems like a flat, infinite 1-dimensional universe actually "wraps" around itself. Feel free to choose the story here.

7.1. **Executive Summary of the "Meaningful" Version.** We will mainly be concerned with the following problem: we are given a function $f\colon [0,1] \to \mathbb{R}$ (usually "well-behaved," but we will leave this vague for now). Under mild assumptions on f, we expect a unique solution $u = u(x,t)$ to the equations:

(1) u satisfies the *heat equation* on $(0,1) \times (0,\infty)$, i.e.,
$$u_t(x,t) = u_{xx}(x,t) \quad \text{for all } 0 < x < 1 \text{ and } t > 0,$$
(where u_t is the derivative of $u(x,t)$ in t with x held fixed, and u_{xx} is the second derivative of $u(x,t)$ in x with t held fixed),

(2) u has *initial value* f, i.e.,
$$u(x,0) = f(x) \quad \text{for all } 0 < x < 1;$$
and

(3) u vanishes at the boundary $x = 0$ and $x = 1$, i.e.,
$$u(0,t) = u(1,t) \quad \text{for all } t > 0.$$

To obtain the "meaningful" version of the heat equation, we can use the formulas in Chapter 14 in [A&G] to see that if u is sufficiently differentiable, then for any (x_0, t_0) we have
$$u_t(x_0,t_0) = \frac{u(x_0,t_0+k) - u(x_0,t_0)}{k} + O(k)$$
where $O(k)$ refers to a function of "order k," i.e., bounded (in absolute value) by a constant times k for small k (see also the pink box on the middle of page 410 of [A&G]). Similarly
$$u_{xx}(x_0,t_0) = \frac{u(x_0+h,t_0) + u(x_0-h,t_0) - 2u(x_0,t_0)}{h^2} + O(h^2).$$
Hence u satisfies the heat equation at some point (x_0, t_0) iff
$$u_t(x_0,t_0) = u_{xx}(x_0,t_0) \tag{14}$$
iff for h, k small we have
$$\frac{u(x_0,t_0+k) - u(x_0,t_0)}{k} + O(k) = \frac{u(x_0+h,t_0) + u(x_0-h,t_0) - 2u(x_0,t_0)}{h^2} + O(h^2) \tag{15}$$
(since this equation involves both h and k, the constant in the $O(k)$ notation has to be valid for all sufficiently small h, and similarly for the $O(h^2)$ term). This equation is then equivalent to the equation
$$\begin{aligned} u(x_0,t_0+k) &= u(x_0,t_0) + \frac{k}{h^2}\big(u(x_0+h,t_0) + u(x_0-h,t_0) - 2u(x_0,t_0)\big) + O(k/h^2) + O(h) \\ &= u(x_0,t_0) + \frac{k}{h^2} D_x^{2,\mathrm{cen}} u(x_0,t_0) + O(k/h^2) + O(h) \end{aligned}$$

———————————————————————————

The simplest form of the heat equation involves "one-dimensional rod," i.e., a closed interval on $\mathbb{R}$, which for simplicity we take to be $[0,1] \in \mathbb{R}$; the temperature $u = u(x,t)$ at each point $x \in [0,1]$ along the rod, and at each point $t \geq 0$ in time (so $t \in \mathbb{R}$) is written as:
$$u_t(x,t) = u_{xx}(x,t) \quad \text{or, more simply} \quad u_t = u_{xx}.$$
However, the equation $u_t = u_{xx}$ gives us very little intuition as to what the heat equations is really saying; furthermore it doesn't give a good approach (as is) to finding a numerical solution. We will use the formulas of Chapter 14 of [A&G] to describe the heat equation in a more "meaningful" way; this way (1) indicates how heat (and similar resources) dissipate in much more general situations, and (2) give a way to solve the heat equation numerically.

7.2. **The "Meaningful" First Derivative Formulas in Chapter 14 of [A&G].** In this section we quote a few results in Chapter 14 of [A&G], regarding how derivatives are approximated by divided differences (in the case of uniformly spaced points).

The formula at the top of page 411, Section 14.1 of [A&G] amounts to the following theorem (which we state without proof).

Theorem 7.1. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be three times differentiable in a neighbourhood of $x_0 \in \mathbb{R}$. Then for $h > 0$ there is a $\xi \in (x_0, x_0+h)$ such that*
$$f(x_0+h) = f(x_0) + hf'(x_0) + (h^2/2)f''(x_0) + (h^3/6)f'''(\xi).$$

One may write the above formula as
$$\frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) + (h/2)f''(x_0) + (h^2/6)f'''(\xi). \tag{16}$$
According to the [A&G] conventions about "order notation" (see pink box in the middle of page 410), we may also write the cruder expression
$$\frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) + (h/2)f''(x_0) + O(h^2)$$
assuming that f''' is bounded in a neighbourhood of x_0; it also follows that we can write, even more crudely,
$$\frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) + O(h) \tag{17}$$
assuming merely that f'' is bounded in a neighbourhood of x_0 (without any assumptions on f''').

Of course, (17) is the usual "meaningful" way of understanding what $f'(x_0)$ is: it is the slope of the secant line of f at x_0 and x_0+h with $h\to 0$, which (according to the $O(h)$ term above) gives the "tangent line" at x_0.

7.3. **The "Meaningful" Second Derivative Formulas in Chapter 14 of [A&G].** In this subsection we will state some analogous theorems informally and without proof.

Similarly to the previous subsection, the *centered formula for the second derivative*, middle of page 412, Section 4.1 of [A&G], shows that for a function $f\colon\mathbb{R}\to\mathbb{R}$ sufficiently differentiable at x_0 we have

$$f''(x_0)=\frac{f(x_0-h)-2f(x_0)+f(x_0+h)}{h^2}-\frac{h^2}{12}\Bigl(f''''(\xi_1)+f''''(\xi_2)\Bigr) \tag{18}$$

for some $\xi_1\in(x_0,x_0+h)$ and $\xi_2\in(x_0-h,x_0)$. Let us mention two consequences we will need.

First, if f has its fifth derivative bounded near x_0, then we may write

$$f''(x_0)=\frac{f(x_0-h)-2f(x_0)+f(x_0+h)}{h^2}-\frac{h^2}{12}f''''(x_0)+O(h^3). \tag{19}$$

Second, if f'''' exists and is bounded near x_0, then

$$f''(x_0)=\frac{f(x_0-h)-2f(x_0)+f(x_0+h)}{h^2}+O(h^2) \tag{20}$$

jjjjjjj

Move all Chap 14 formulas here.

7.4. **Great News for the Smoothness of the Heat Equation.** It turns out that in all the problems that interest us, i.e., a heat equation satisfied somewhere plus "boundary conditions" that determine a unique heat profile, the heat profile $u=u(x,t)$ (sometimes $u=u(\mathbf{x},t)$, where $\mathbf{x}$ varies in a two- or three-dimensional space), u will turn out to be infinitely differentiable near any (x,t) where the heat equation holds.

For this reason, all the assumptions in Subsection 7.2 and 7.3 regarding bounded derivatives is automatic. In fact, u turns out to be "locally analytic" which is a much stronger property. We will indicate what this means in Section **??**.

Just as in splines (Chapter 11 of [A&G]) and differential equations in general, aside from the heat equation

$$u_t=u_{xx}$$

which is meant to hold in the *interior* of some metallic (or other conducting) material, there are "boundary conditions" that we need to specify to get a unique solution.

7.5. **First Derivative Approximation in the Heat Equation.** jjjjjjj

The symbol u_t means differentiate in t holding x fixed, and hence for any x,t,

$$u_t(x,t)=\lim_{k\to 0}\frac{u(x,t+k)-u(x,t)}{k}.$$

Hence we will approximate $u_t(x,t)$ by taking k small and writing

$$u_t(x,t)\approx\frac{u(x,t+k)-u(x,t)}{k}. \tag{21}$$

Applying this to $u(x,t)$ with x fixed and t varying, we have that for any (x,t) we have

$$\frac{u(x,t+k)-u(x,t)}{k}=u_t(x,t)+(k/2)u_{tt}(x,t)+(k^2/6)u_{ttt}(x,\tau). \tag{22}$$

We will later need this more precise version of (21); but for now, (21) will suffice for our purposes.

7.6. **Second Derivative Approximation in the Heat Equation.** The symbol u_{xx} means differentiate in x twice, holding t fixed.

It follows that for any fixed (x,t) we have

$$u_{xx}(x,t)=\frac{u(x_0-h,t)-2u(x_0,t)+u(x_0+h,t)}{h^2}-\frac{h^2}{12}\Bigl(u''''(\xi_1,t)+u''''(\xi_2,t)\Bigr) \tag{23}$$

for some ξ_1,ξ_2 as before. However, for our numerical methods we will simple use

$$u_{xx}(x,t)\approx\frac{u(x_0-h,t)-2u(x_0,t)+u(x_0+h,t)}{h^2} \tag{24}$$

7.7. **The Simple Form and Basic Numerical Method for Solving the Heat Equation.** The equation $u_t=u_{xx}$ for all (x,t), along with (21) and (24) implies that for small k,h we have:

$$\frac{u(x,t+k)-u(x,t)}{k}\approx\frac{u(x_0-h,t)-2u(x_0,t)+u(x_0+h,t)}{h^2}$$

which we rearrange as

$$u(x,t+k)\approx u(x,t)+\frac{k}{h^2}\Bigl(u(x_0-h,t)-2u(x_0,t)+u(x_0+h,t)\Bigr).$$

7.8. **Motivation for the Heat Equation.** Consider a long and thin rod of some metallic material in space (i.e., $\mathbb{R}^3$, where we assume that that

(1) the rod has length 1 in some fixed units (metres, centimetres, etc.) which we model by the interal $[0,1]$;
(2) the rod is insulated (either by plastic, air, etc.)
(3)

8. The Classical Wave and Laplace Equations

There are two other classical PDE's which are extremely important in the foundations of PDE's.

First, the $t \to \infty$ limit of the heat equation $u_t = u_{xx}$, or, in n-dimensions

$$u_t = \Delta u, \quad \text{where} \quad \Delta u = \sum_{i=1}^{n} u_{x_i x_i}$$

under reasonable assumptions has a "steady-state" limit

$$\Delta u = 0$$

subject to the boundary conditions imposed on the same heat equation. In studying this equation subject to Dirichlet and Neumann boundary conditions one encounters traditional "Potential Theory" which gave rise to a wealth of mathematical techniques.

The classical n-dimensional wave equation

$$u_{tt} = \Delta u,$$

must be considered as important and interesting as the heat equation. We remark that the Dirichlet and Neumann problems for the wave equations can also be solved with harmonics in a way related to our "explicit solution" of the heat equation, and it is really in this context—that of stringed instruments such as a guitar or violin—that the ancient Greeks discovered uncovered what in modern times are called the eigenvalues and eigenfunctions of the Laplacian (on a one-dimensional rod or string).

9. ODE's and Some 2-Body Experiments

In this section we give a quick overview of the theory, some examples, and some associated numerical methods.

9.1. **ODE's in Popular Culture: "Chaos Theory" and "Catastrophe Theory".** It is interesting to note that the theory of ODE's has given rise to two rather sensationistic terms in mathemtics: *chaos theory* and *catastrophe theory*.

Chaos theory is really an umbrella term for a number of mathematics, including ergodic theory (occurring both in pure mathematics and information theory). It was certainly inspired—at least in part—by the so-called Lorenz attractor. Chaos theory has been featured in some of the *Jurassic Park* movies.

The term *catastrophe theory* is really just a "rallying cry" for the traditional part of ODE's and PDE's called *bifurcation theory* and *singularity theory*. This term was based on a manuscript of René Thom (a fields metalist with renouned contributions to topology), which is a loose aggregation of some philosophical and (partly imprecise) mathematical ideas. The term was popularized by Sir Christopher Zeeman, a uniquely gifted and pioneering expositor of mathematics. The term *catastrophe theory* was quite divisive at the time (surely a hyperextended umbrella is not a "catastrophe"). However, by the end of the last century both its critics (of the term, not the mathematics) such as Vladimir Arnold, and its proponents, such as Thom and Zeeman, seemed to be in agreement in their views.

Catastrophe Theory is the study of a parametric family of ODE's, in a neighbourhood of a point in the space of parameters which is *singular*, e.g., in the sense of minimizing dynamics, $du/dt = -(\nabla f)(u)$ where f is a critical point in the sense of Morse theory (and some much worse singularies). To date, Catastrophe Theory allows one to explain the past, but has yet to yield any true predictive power in any domain of "real world" applications. The problem is—as we understand it—there are no research articles that similutaneously (1) demonstrate an understanding of "catastrophe" theory of parametric ODE's, singularities of polynomial maps $\mathbb{R}^n \to \mathbb{R}$, and their implications, (2) demonstrate an understanding mathematical modeling, (3) use singularity (catastrophe, bifurcation, etc.) theory to formulate serious models, and (4) compare the models—in good faith—to simpler models that avoid catastrophe theory or simpler singularities (any paper using the so-called "cusp catastrophe" ought to consider the simpler and more theoretically prevalent "fold catastrophe"). With the advent of modern computers and large data sets we expect this to change.

9.2. **Examples from ODE's.**

9.3. **The Basic Existence and Uniqueness Theorems.**

9.4. **Experiments with the 2-Body Problem.** Here are some MATLAB experiments that we have created to demonstrate the numerical solution of the 2-body problem. The code is rather simplistic, and bodies of different masses still appear as points. We warn the reader that one of the `several_runs.m` file has *negative* gravity.

We invite students to play around with the initial values, and perhap modify the code to get a 3-body problem solver.

While the 2-body problem produces known results (of "celestial mechanics"), a variety of initial conditions in the 3-body problem can give "chaotic" results. It is even known that with as few as 4 bodies (?), one can get a type of singularity beyond the usual "resolvable" singularity that occurs when two bodies reach the same point.

To run this, you can save the following files in some directory and run `several_runs.m` there. This software seems to work on my laptop running MATLAB; it is definitely a "beta" version, not well tested or debugged.

`my2body.m` IS THE FILE:

```
function dy = my2body(t,y,G,m1,m2)

% y(1) = x-position mass 1
% y(2) = x-velocity mass 1
% y(3) = y-position mass 1
% y(4) = y-velocity mass 1
% similarly y(5:8) for mass 2

dy = zeros(8,1);   % make sure this is a column vector

% d/dt of change in position = velocity

dy(1) = y(2);
dy(3) = y(4);
```

```
dy(5) = y(6);
dy(7) = y(8);

% d/dt of change in velocity
%   = - G other_mass ( difference of position ) / distance^3

dist12 = sqrt( (y(1)-y(5))^2 + (y(3)-y(7))^2 );

dy(2) = ( y(5)-y(1) ) * G * m2 / dist12^3;
dy(4) = ( y(7)-y(3) ) * G * m2 / dist12^3;
dy(6) = ( y(1)-y(5) ) * G * m1 / dist12^3;
dy(8) = ( y(3)-y(7) ) * G * m1 / dist12^3;

end
```

one_run.m IS THE FILE:

```
% This program runs ode45 on my2body, which solves the two body problem.
% Now G,m1,m2 are global variables (G = Gravitation const, m1,m2 = masses)

function one_run(G,m1,m2,end_t,tstep,pause_len,yinit)

% clear   % clear all variables
close all  % close all figure windows

t = [ 0 : tstep : end_t ];
[newt,yvals] = ode45(@(t,y) my2body(t,y,G,m1,m2),t,yinit);   % this works!

dist_init = sqrt( (yinit(1)-yinit(5))^2 + (yinit(3)-yinit(7))^2) ;
mvsqu1 = m1 * (yinit(2)^2+yinit(4)^2) ;
mvsqu2 = m2 * (yinit(6)^2+yinit(8)^2) ;
energ = (mvsqu1 + mvsqu2)/2 - G * m1 * m2 / dist_init ;

ener = zeros( [ size(t,2) 1]);
for k = 1 : size(t,2)
  yinit = yvals(k,1:8);
dist_init = sqrt( (yinit(1)-yinit(5))^2 + (yinit(3)-yinit(7))^2) ;
mvsqu1 = m1 * (yinit(2)^2+yinit(4)^2) ;
mvsqu2 = m2 * (yinit(6)^2+yinit(8)^2) ;
ener(k) = (mvsqu1 + mvsqu2)/2 - G * m1 * m2 / dist_init ;
end

% ener

for k = 1 : size(t,2)
  pause(pause_len);
  hold off;
  plot(yvals(k,1),yvals(k,3),...
```

```
    'Marker','o','MarkerEdgeColor','red','MarkerFaceColor','red');
  hold on;
  plot(yvals(k,5),yvals(k,7),...
    'Marker','o','MarkerEdgeColor','black','MarkerFaceColor','black');
  set(gca,'xLim',[-2,5],'yLim',[-2,5]);
  title([ 'Two body simulation.  G = ' num2str(G) ', m1 = ' num2str(m1), ', m2 = ' num2str(m2
    ', Energy = ' num2str(energ) ] ) ;
end
```

several_runs.m IS THE FILE:

```
% several_runs : makes several two body runs, no particular pattern
%
% uses the function one_run(G,m1,m2,end_t,tstep,pause_len,yinit)

clear   % clear all variables
close all  % close all figure windows

m1=1; m2=3; end_t=15; tstep = 0.05; pause_len = 0.02;

% change gravity, everything else the same
G = 0.6;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

% change gravity, everything else the same
G = 0.3;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

% change gravity, everything else the same
G = 1;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

% change gravity, everything else the same
G = 2;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

% change gravity, everything else the same
G = 0.7;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

% negative gravity, just for fun...
G = -0.7;
one_run(G,m1,m2,end_t,tstep,pause_len,[1,-1,1,0,0,0.5,0,0]);

close all;
```

Appendix A. Some Fundamental Matrices of Interest to Us

In this section we introduce notation to describe two families of matrices of interest to us to understand certain differential equations. We have seen one of the families before when we discussed splines.

A.1. Matrices Describing Both Splines and Heat Transfer over a Rod.

For reasons that will become clear later, for any $n \in \mathbb{N} = \{1, 2, \ldots\}$ we define $N_{\mathrm{rod},n}$ to be the matrix $\mathbb{R}^{n\times n}$, i.e., the $n \times n$ square matrix

$$N_{\mathrm{rod},n} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}.$$

For example, we have

$$N_{\mathrm{rod},2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad N_{\mathrm{rod},3} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad N_{\mathrm{rod},4} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

One can alternatively describe $N_{\mathrm{rod},n}$ by its entries: for all $i, j \in [n]$, the (i,j)-th component of

$$\big(N_{\mathrm{rod},n}\big)_{i,j} = \begin{cases} 1 & \text{if } i - j = \pm 1, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

We have already seen this matrix when we discussed splines; indeed, it is nothing more than the "off diagonal part" of the $n \times n$ matrix given at the top of page 343 (Section 11.3) of [A&G] in the special case $h_1 = h_2 = \cdots = h_n = 1$. Furthermore, the full matrix in [A&G] at the top of page 343 (with $h_1 = h_2 = \cdots = h_n = 1$) can be written more simply as

$$4I_n + N_{\mathrm{rod},n},$$

where I_n dentoes the $n \times n$ identity matrix. We use the term "rod" because we think of the interval $[A, B]$ ($[a, b]$ in the textbook) with arbitrary A and $B = A + n$ as a finite, one-dimensional rod.

There are some fundamental observations in [EnergySplines], Sections 4.2-4.4, and the corresponding [A&G] Section 11.3 regarding solving the equation

$$\Big(4I_n + N_{\mathrm{rod},n}\Big)\mathbf{c} = \boldsymbol{\psi},$$

where f is a function that we are modeling, like the profile of a car, and $\boldsymbol{\psi}$ was called in class the "energy of f," since $\boldsymbol{\psi}$ is essentially comprised of second divided differences of f (again, in the special case where $h_1 = h_2 = \cdots = h_n = 1$). The observation is that

$$\Bigl(4I_n + N_{\mathrm{rod},n}\Bigr)^{-1} = (1/4)\bigl(I_n - (N/4) + (N/4)^2 - (N/4)^3 + \cdots\bigr)$$

which converges "geometrically" since $\|N/4\|_\infty \le 1/2$ (where we have written simply N for $N_{\mathrm{rod},n}$). As a consequence, the formula

$$\mathbf{c} = (1/4)\bigl(I_n - (N/4) + (N/4)^2 - (N/4)^3 + \cdots\bigr)\boldsymbol{\psi}$$

tells us that the effect of the values of $\boldsymbol{\psi}$ on $\mathbf{c}$ are "localized," in that for any $k \in \mathbb{Z}$, N^k is only nonzero on entries i, j with $|i - j| \le k$, and the term involving N^k, namely $(N/4)^k$, has ∞-norm at most $1/2^k$.

The matrix $N_{\mathrm{rod},n}$ can also be viewed as describing which integers in $[n] = \{1, \ldots, n\}$ are "immediate neighbours" or "adjacent." This matrix is useful in *graph theory*, on the graph with vertices $1, \ldots, n$ where we "joined by an edge" any two vertices differing by 1.

There are many interpretations and applications of this matrix.

Again, as in [EnergySplines] we use the (somewhat awkward choice of) letter N to suggest the notion that this matrix represents a relation between "neighbours" along a rod.

A.2. **Matrices Describing Heat Transfer over a (One-Dimensional) Ring.** Many computations turn out to be simpler when we "connect" the endpoints of the above rod, bending the rod to form a ring. This gives us the matrix

$$N_{\mathrm{ring},n} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix},$$

which is just the matrix $N_{\mathrm{rod},n}$ with a 1 added to the top right and to the bottom left corners. This matrix as each column sum and each row sum equal to 2; it also has a *cyclic symmetry* that makes it a *Toeplitz matrix* (see the Wikipedia page on *Toeplitz Matrix*); we may return to Toeplitz matrices later.

For example, we have

$$N_{\mathrm{ring},2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad N_{\mathrm{ring},3} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad N_{\mathrm{ring},4} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Appendix B. Matrix Multiplication, Polynomials, Power Series, and Exponentiation Involving the Rod and Ring Matrices

In this section we will give the tools we need to carry out calculations with the rod and ring matrices defined above.

B.1. **Rod Matrices and Shift Operators.** Let us interpret the matrix $N_{\mathrm{rod},n}$ for a fixed n as an operator.

First, for any $n \in \mathbb{N}$, note that

$$N_{\mathrm{rod},n} = S_{n,1} + S_{n,-1},$$

where

$$S_{n,1} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}, \quad S_{n,-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix};$$

hence $S_{n,1}$ is the nonzero part of $N_{\mathrm{rod},n}$ that lies above the diagonoal, and $S_{n,-1}$ the part below. However, $S_{n,1}$ has a simple interpretation as "shifting up by one," in the sense that for any $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}$ (which, as always, we think of as a column vector),

$$S_{n,1}\,\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-3} \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \\ x_5 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \\ 0 \end{bmatrix}$$

Hence the way that $S_{n,1}$ *operates* on a vector $\mathbf{x}$ is to move all its components up by one, and introduce a zero in the bottom component. Similarly we have

$$S_{n,-1}\,\mathbf{x} = \begin{bmatrix} 0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-3} \\ x_{n-2} \\ x_{n-1}, \end{bmatrix}$$

so $S_{n,-1}\mathbf{x}$ *operates* by shifting the components of $\mathbf{x}$ down by one and introduces a 0 on the top.

B.2. **Ring Matrices and Cyclic Shift Operators.** Working with ring matrices is much simpler, because they can be described as a sum of cyclic shift operators: indeed, for any $n \in \mathbb{N}$, note that

$$N_{\mathrm{ring},n} = C_{n,1} + C_{n,-1}, \tag{25}$$

where

$$C_{n,1} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_{n,-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Similarly to the previous subsection, we have

$$C_{n,1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-1} \\ x_n \\ x_1, \end{bmatrix} \tag{26}$$

and hence $C_{n,1}$ has the effect of "cyclically rotating the components of $\mathbf{x}$ up by one," taking the x_1 to be its bottom component, instead of the 0 that $S_{n,1}$ introduces. Similarly we have

$$C_{n,-1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_n \\ x_1 \\ x_2 \\ \vdots \\ x_{n-3} \\ x_{n-2} \\ x_{n-1} \end{bmatrix} \tag{27}$$

B.3. **Ring Matrices and the $C_{n,\pm1}$ are Easier to Work With than Rod Matrices and the $S_{n,\pm1}$.** For many computations, it is easier to work with the $C_{n,\pm1}$ than the $S_{n,\pm1}$, and to see how powers and polynomials of

$$N_{\mathrm{ring},n} = C_{n,1} + C_{n,-1},$$

behave as opposed to

$$N_{\mathrm{rod},n} = S_{n,1} + S_{n,-1}.$$

For example, to interpret $N_{\mathrm{ring},n}^2$, we have

$$N_{\mathrm{ring},n}^2 = \bigl(C_{n,1} + C_{n,-1}\bigr)^2$$

which equals

$$\bigl(C_{n,1} + C_{n,-1}\bigr)\bigl(C_{n,1} + C_{n,-1}\bigr) \tag{28}$$

To simiplify such an expression we note that

$$C_{n,1}^2 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ \vdots \\ x_n \\ x_1 \\ x_2, \end{bmatrix}$$

which just cyclically shifts the components of $\mathbf{x}$ by 2. More generally, if for any $k \in \mathbb{N}$, if we set

$$C_{n,k} = C_{n,1}^k,$$

then $C_{n,k}$ is the operator that cyclically rotates the components of a vector up by k; similarly for $C_{n,-k} = C_{n,-1}^k$. We similarly see that $C_{n,-1}C_{n,1}\mathbf{x}$ is just $\mathbf{x}$, and hence

$$C_{n,-1}C_{n,1} = C_{n,1}C_{n,-1} = I_n$$

the identity matrix. Hence all the $C_{n,\pm k}$ are invertible, and they all commute; setting $C_{n,0} = I_n$ (which makes sense, in that shifting by 0 does nothing to a vector), we conclude that

$$\bigl(C_{n,1} + C_{n,-1}\bigr)^2 = C_{n,2} + 2I_n + C_{n,-2}.$$

Furthermore, this can be seen as a manifestation of the identity

$$\bigl(x + x^{-1}\bigr)^2 = x^2 + 2 + x^{-2}.$$

Unfortunely, things are not as simple with the matrices or operators $S_{n,1}$ and $S_{n,-1}$. For one thing, $S_{n,1}$ is not invertible, since its bottom row consists entirely of 0's. And it is also not true that $S_{n,1}$ and $S_{n,-1}$ commute: i.e.,

$$S_{n,-1}S_{n,1} \neq S_{n,1}S_{n,-1}$$

(see the exercises). It turns out that for *positive* integers, k, one can set $S_{n,k} = S_{n,1}^k$, and this has as reasonable interpretation, just as does $S_{n,-k} = S_{n,-1}^k$, but when we multiply positive of powers of $S_{n,1}$ times those of $S_{n,-1}$, the operators we get will operate on a vector $\mathbf{x}$ like cyclic shifts except that some of the x_i are replaced by zero.

Appendix C. Review of Polynomials and Power Series in Matrices

We will use $\mathbb{R}^{n\times n}$ for a the set of $n \times n$ of matrices with real entries.

Recall that if A is an $n \times n$ matwe use the notation

$$A^0 = I, \quad A^1 = A, \quad A^2 = AA, \quad A^3 = AAA,$$

and that all these matrices *commute*, i.e.,

$$A^s A^r = A^r A^s$$

for all $r, s \in \mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$.

If $p(x) = c_0 + c_1 x + \cdots + c_m x^m$ is a polynomial, we define

$$p(A) = c_0 I + c_1 A + \cdots + c_m A^m$$

where $I = I_n \in \mathbb{R}^{n\times n}$ is the $n \times n$ identity matrix (and where $c_i \in \mathbb{R}$ or $c_i \in \mathbb{C}$, i.e., the coefficients are real-valued or complex-valued).

If p, q are any two polynomials, then we easily see that $p(A)$ and $q(A)$ commute, i.e.,

$$p(A)q(A) = q(A)p(A),$$

and, moreover, if

$$r(x) = p(x)q(x)$$

is the usual product of polynomials, then we use the shorthand $r = pq$, and we have

$$r(A) = (pq)(A) = \bigl(p(A)\bigr)\bigl(q(A)\bigr).$$

Example C.1. If $p(x) = x + 2$ and $q(x) = x + 1$, then

$$\begin{aligned}(A + 2I)(A + I) = (A + I)(A + 2I) &= A^2 + 3A + 2I \\ &= \text{“}(x^2 + 3x + 2)\,\big|_{x=A}\text{”} = \text{et cetera}\end{aligned}$$

In [EnergySplines], Section 4.2–4.4 we extended the notion of a polynomial of a matrix, A, to power series in A, such as

$$(I + A)^{-1} = I - A + A^2 - A^3 + \cdots \tag{29}$$

which is the matrix analog of

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

and where the power series in A converges whenever $\|A\|_\infty < 1$ (or similarly for other matrix norms), just as the power series in x converges whenever $|x| < 1$.

For reasons that will become clear when we work with differential equations, we will also be interested in the analog of the Taylor expansion for e^x about $x = 0$, namely

$$e^x = 1 + x + x^2/2 + x^3/3! + x^4/4! + \cdots$$

which is valid for all $x \in \mathbb{R}$, and its matrix analog

$$e^A = I + A + A^2/2 + A^3/3! + A^4/4! + \cdots$$

which converges for all $A \in \mathbb{R}^{n\times n}$. Roughly speaking, the reason is that the solution to the differential equation $dx/dt = ax$ for a constant $a \in \mathbb{R}$, and variables x, t is given by $x(t) = x(0)e^{at}$; similarly, the solution to the vector-valued differential equation $d\mathbf{x}/dt = A\mathbf{x}$ for an $A \in \mathbb{R}^{n\times n}$ and $\mathbf{x} = \mathbf{x}(t)$ taking values in $\mathbb{R}^n$ turns out to be

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0).$$

We will discuss this further in the next homework.

EXERCISES

(1) Let us make some calculations regarding the matrices $S_{n,1}$ and $S_{n,-1}$ in the case $n = 3$, where

$$S_{3,1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad S_{3,-1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

(a) Using regular matrix multiplication, compute the expressions $S_{3,1}S_{3,-1}$ and $S_{3,-1}S_{3,1}$. Are they equal?

(b) Compute the way $S_{3,1}S_{3,-1}$ operates on a vector $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ as
$$S_{3,1}\big(S_{3,-1}\mathbf{x}\big),$$
i.e., by first computing
$$S_{3,-1}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$
and then applying $S_{3,1}$ to the result.

(c) Similarly compute the way that $S_{3,-1}S_{3,1}$ operates on a vector $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

(d) Using matrix multiplications, compute the values of
$$S_{3,1}^2, \quad S_{3,1}^3.$$

(e) Compute the way that
$$S_{3,1}^2, \quad S_{3,1}^3$$
operate on a vector $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ by applying twice and three times, respectively, the operation of $S_{3,1}$ on $\mathbf{x}$.

(f) Compute the value of
$$e^{S_{3,1}t}$$
using the formula
$$e^{At} = I + (At) + (At)^2/2 + (At)^3/3! + \cdots$$
using either direct matrix calculations or operator calculations as above.

(g) Compute the value of $(I + S_{3,1})^{-1}$ using (29). Check your work by multiplying this result by $I + S_{3,1}$ to see that you get the identity matrix.

(2) (a) Assuming that $x \in \mathbb{R}$ is nonzero, simplify the expression
$$\big(x + x^{-1}\big)^3.$$

(b) Fix any value of $n \in \mathbb{Z}$; simplify the expression
$$\big(C_{n,1} + C_{n,-1}\big)^3.$$

(c) Using the binomial theorem
$$(x+y)^n = x^n + nx^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + nxy^{n-1} + y^n,$$
write an expression for
$$\big(x + x^{-1}\big)^n.$$

(d) Using the binomial theorem, write an expression for
$$\big(C_{n,1} + C_{n,-1}\big)^n.$$

(3) Let

$$A = N_{\mathrm{ring},2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

and for $n \in \mathbb{Z}_{\ge 0} = \{0, 1, 2, \ldots\}$, set $X_n = A^n$, i.e.,

$$X_0 = I_2, \quad X_1 = A, \quad X_2 = A^2, \quad \ldots$$

(a) Show that

$$X_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, X_2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

(b) Can you guess a formula for X_n based on these three examples? [There is no credit for this part, but you should take a new moments to see if there is some simple pattern to this sequence.]

(c) Show that for any $a, b \in \mathbb{R}$, we have

$$\begin{bmatrix} a & a \\ a & a \end{bmatrix} \begin{bmatrix} b & b \\ b & b \end{bmatrix} = \begin{bmatrix} 2ab & 2ab \\ 2ab & 2ab \end{bmatrix}$$

(d) Explain concisely why we may now conclude that for any $a \in \mathbb{R}$ we have

$$\begin{bmatrix} a & a \\ a & a \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2a & 2a \\ 2a & 2a \end{bmatrix}$$

(e) Show that the next few terms in the sequence $\{X_n\}$ are

$$X_3 = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}, X_4 = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}.$$

(f) Show that the $\{X_n\}$ satisfy the "two-term (matrix) recurrence equation"

$$X_{n+1} = X_n \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and also the recurrence equation

(30) $$X_{n+1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} X_n, \quad n \ge 0$$

(g) Have we seen such a "two-term (matrix) recurrence equation" related to the Fibonacci numbers? What is this "two-term (matrix) recurrence equation" related to the Fibonacci numbers?

(h) Are you surprised that, in view of the fact that

$$X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, X_2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, X_3 = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}, X_4 = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}.$$

that

$$X_0 \ne \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}?$$

[No credit for your answer to this question, just say whether or not you were surprised.] Explain why, in retrospect, you shouldn't be very surprised. [Full credit for your explanation here.]

Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, CANADA.

E-mail address: `jf@cs.ubc.ca`

URL: `http://www.cs.ubc.ca/~jf`