11-2022

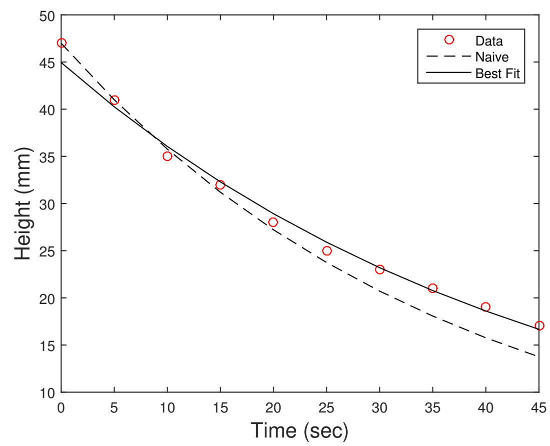
Math 57: Applied Differential Equations I

John Mayberry
University of the Pacific, jmayberry@pacific.edu

Math 57: Applied Differential Equations I

John Mayberry

November 18, 2022



Notes to the user:

- The one who does the work does the learning. This is a story about *my* relationship with differential equations. Read my words for what they're worth, but more importantly do problems, ask questions, and write your own story[1].
- Edwards and Penney's "Differential Equations and Boundary Value Problems" is a more rigorous and thorough coverage of differential equations than this one and is referenced in various places throughout the following pages when we want to avoid getting "lost in the sauce" of technical details. You can easily find free pdfs of older editions if you want to follow up on these side quests.
- Many of the typos originally present in these notes have been caught by astute students in past semesters[2], but I recently re-edited some sections so please let me know if you find new ones.

Without further adieu, let the games begin!

[1]The late Joe Strummer liked to challenge his audiences with the statement "If you think you can do better, go start your own fucking band." I'll challenge the same here...with a book instead of a band.

[2]Special thanks to Shyla Solis for correcting so many.

Contents

Prologue: Modeling with Differential Equations

After three semesters of calculus, you are probably well acquainted with derivatives. In this course, we will be using equations involving derivatives to model systems in motion and describe the "rates of change" for various physical quantities as they evolve over time or space. An equation involving one or more derivatives of an unknown function is called a ***differential equation*** (hence the title of this course). To illustrate one of the simplest possible scenarios, we turn to an early motivation for the study of differential equations (albeit in a modern form); namely the calculation of velocity and position from acceleration.

When I was an undergraduate, my father purchased a Suzuki Hayabusa, a ridiculously fast motorcycle with the claimed capability of accelerating from 0-60 mph in 2.6 seconds. One day he posed to me the following question:

> Assuming a rider is initially stopped at a traffic light and accelerates at a constant rate, how far could they travel during 2.6 seconds?

To answer this question, let $y(t)$ denote the distance (in miles) traveled by the rider after t seconds. If we reformulate the question above into a mathematical equation involving our newly defined variables t, y, then our goal is to find $y(2.6)$. Although there are quicker ways of answering this question, we will take a route which foreshadows future course developments and find a general formula for $y(t)$.

We start out by examining the "dynamic" assumption of constant acceleration. Formulating constant acceleration as a mathematical equation leads to our first example of a differential equation:

$$\frac{d^2y}{dt^2} = a \tag{1}$$

where $a > 0$ is the currently unknown constant acceleration. Equation (1) is called a differential equation[1] because it involves derivatives of the unknown function $y(t)$. In our example, we can further classify Equation (1) as a ***second order*** differential equation because the highest order derivative which appears is the second derivative of y with respect to t. Our goal is to find a formula for the function y in terms of t; that is, we want to find a function $y(t)$ whose second derivative with respect to t is the constant a. A function $y(t)$ with this property is called a ***solution*** to the differential equation.

We will learn many different techniques for obtaining solutions to differential equations throughout this course, but in this example, we can just go back to Calculus 1 and integrate both sides of Equation (1) with respect to t to obtain a new ***first order*** differential equation

$$\frac{dy}{dt} = at + c_1 \tag{2}$$

where c_1 is an unknown integration constant. This equation tells us the velocity, dy/dt, of the motorcycle at time t. We then integrate a second time to obtain the expression

$$y(t) = \frac{a}{2}t^2 + c_1 t + c_2 \tag{3}$$

where c_2 is another integration constant. This expression for $y(t)$ is called the ***general solution*** to Equation (1): if we differentiate both sides of this expression twice with respect

[1] We shall often abbreviate differential equation with DE throughout the rest of the text.

to t, we see that (1) is satisfied and conversely, any function $y(t)$ whose second derivative is a must have the form given in Equation (3) for some constants c_1, c_2.

To determine the values of a, c_1, and c_2, we need to use additional "static" information that is given by the problem:

- the initial velocity is 0 mph.
- the initial distance traveled is 0 miles.
- the velocity after 2.6 seconds is 60 mph.

Translating these statements into mathematical equations yields:

$$\begin{aligned} y(0) &= 0\ mi \\ y'(0) &= 0\ mi/s \\ y'(2.6) &= 60\ mi/h = 60\ mi/(3600s) = \frac{1}{60}\ mi/s \end{aligned} \tag{4}$$

where we use y' as shorthand for the derivative of y with respect to t. Substituting in the values from the second line in (4) into equation (2), we can see that

$$0 = y'(0) = a(0) + c_1.$$

Thus, $c_1 = 0$. Using this fact along with $y(0) = 0$ in the expression $y(t) = at^2/2 + c_1 t + c_2$, we obtain

$$0 = \frac{a}{2}(0) + c_2$$

so that $c_2 = 0$ as well. Finally, using the third condition from (4) in (2)

$$\frac{1}{60} = a(2.6)$$

or $a = 1/156$. Having nailed down our three unknowns, we are finally in a position to answer our initial question: the rider will have traveled:

$$y(2.6) = \frac{1}{2(156)}(2.6)^2 = 0.022 \quad \text{miles}$$

after 2.6 seconds (or about 114.4 feet).

Let's review the steps we went through to arrive at this solution:

(I) We translated "dynamic" information provided by the application (constant acceleration) into a differential equation model (1) which described how the unknown function of interest $y(t)$ evolved over time.

(II) We determined the general solution to the differential equation (Equation (3)).

(III) We plugged "static" information about the position and velocity of the motorcycle into our general solution to solve for the unknowns a, c_1, and c_2 in our general solution.

(IV) We analyzed the behavior of our equation, in this case using it to answer a question of practical interest.

And that's all there is to this course. At least, that's what you're hoping to hear at this point, right? If only life were so simple! We will see that while these four steps do describe an ideal theoretical procedure for modeling with differential equations, the reality is much more complicated. Even over the course of the first week, we will see that

(i) We can never fully encapsulate reality into a mathematical equation; even getting "close" is challenging.

(ii) We cannot always find a nice formula for the general solution to a differential equation, if it has a solution at all.

(iii) We do not always have enough data (or good enough data) to accurately estimate unknown parameters.

(iv) We need to develop new perspectives to make sense out of solutions once we have them.

The good news is that despite the fact that no model is perfect, some models are useful[2] and that will be the focus of our short time together: developing useful examples of models which one can use to simulate, investigate, and answer questions involving real world phenomenon. We will start out with first order equations (like (2)). This will comprise Chapter 1 of the course. Chapter 2 will focus on second order equations (like (1)) while Chapter 3 will focus on systems of equations which involve several unknown functions. There will be practice problems, Canvas quizzes, and modeling assignments (often involving the scientific computing software Matlab) dispersed throughout the modules to get you actively engaged in the learning process. Also, if you're the type who enjoys doing math more than reading math (like me), I would encourage you to read these notes with a sheet of scratch paper and try to do the calculations on your own before you read example solutions. The details will be here to check, but you'll feel better if you figure them out on your own first. So with all that being said, happy modeling!

Exercises

(1) A body is dropped from the top of Burns tower (approximately 160 feet high) at midnight. Assuming a constant acceleration of 32 ft/sec^2 and zero initial velocity, estimate the time at which the body will hit the ground.

(2) Use Equations (2) and (3) to derive the following five "kinematic equations" for an object moving under constant acceleration a.

   (a) $v = at + v_0$ where v_0 is the initial velocity of the object and v is the velocity after t seconds

   (b) $d = \frac{a}{2}t^2 + v_0 t + d_0$ where d_0 is the initial distance traveled (usually 0 unless we measuring "distance" as a displacement or position) and d is the distance traveled after t seconds.

   (c) $(v_0 + v)/2 = d/t$ where v_0, v, d are all defined as in (a),(b).

   (d) $a = (v - v_0)/t$.

   (e) $v^2 = v_0^2 + 2ad$.

[2]Box, George (1979) *Robustness in the Strategy of Scientific Model Building.* Robustness in Statistics: Proceedings of a Workshop. edited by RL Launer and GN Wilkinson

Part I
First Order Differential Equations

Section 1.1: Exponential Growth and Decay

Section objectives:

- Translate statements about proportional rates of change into first order differential equations
- Solve equations of the form $y' = ay$ using separation of variables
- Estimate parameters in exponential models from data

In the prologue, we mentioned four algorithmic steps which we would like, but are often difficult, to perform in modeling with differential equations. This section will take a look at a slightly more complicated example which demonstrates ambiguity in the modeling process. Exploring this problem will also introduce a new technique for solving differential equations that goes a step beyond direct integration.

A Model for Beer Froth

Suppose that beer is poured quickly into a large cylindrical glass. A layer of froth (usually called head) will form at the surface of the beer, but will gradually subside over time. Suppose we don't want to start drinking the beer until the head subsides to a height of 1 cm, but also don't want to stand around waiting for that to happen. Can we set up a simple model to predict the amount of time we will need to wait?

When developing a model, we usually want to have some data for calibration and validation. A video describing a simple experiment that can be used to generate such data can be found here: `http://www.youtube.com/watch?v=r5RMNBVojgg`. If you plan on reproducing this experiment (I recommend using a solution of dish soap and water instead if you are under 21!), you should also view the follow up video at: `http://www.youtube.com/watch?v=68PEGUYKi_Q`. For the purposes on these notes, we will use the data shown in Table 1 which was obtained from one such experiment. A plot of this data is also included in Figure 2.

Setting up a mathematical model for this data involves determining an equation which relates the height of beer froth to time in a manner that matches the experimental data. Before reading on, we encourage you to jot down a few observations about the data table and the corresponding figure. Here are some questions you might want to consider:

- Does the shape of the graph resemble any familiar families of functions?
- Why does the rate of change of the height decrease over time?
- Is there a simple equation we could use to relate the height of the beer foam to time?
- How would you estimate the time until the foam reaches a height of 1 cm (10 mm)?

Time	**Height**
0	47
5	41
10	35
15	32
20	28
25	25
30	23
35	21
40	19
45	17

Table 1: Time (sec) and Height (mm) from a beer pouring experiment involving Sierra Nevada's KellerWeis Beer.

Guessing a particular family of functions from a graph is one commonly used method of mathematical modeling, particularly in statistics. For example, many students look at the graph in Figure 2 and say that it looks "exponential". However, consider the graphs shown in Figure 1. One of these two is exponential and the other is quadratic. Can you guess which is which?

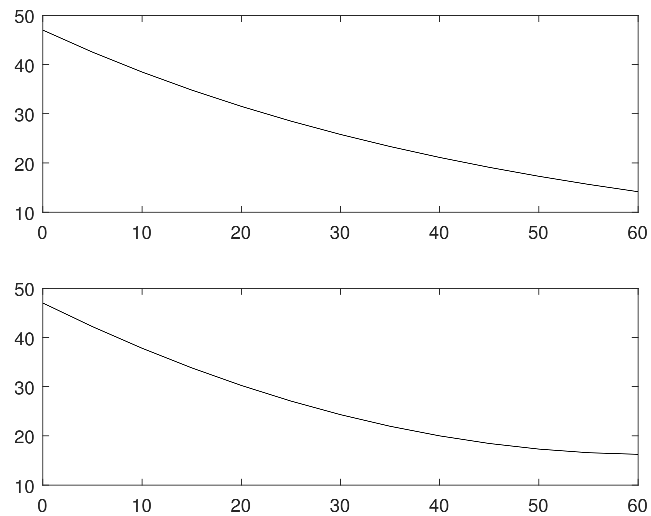


Figure 1: A comparison of an exponential and quadratic function.

If you guessed the top one was quadratic and the bottom one was exponential, then you were wrong! So while the guess and check method of model fitting is certainly used and useful in certain situations, in this course, we will focus more on the scientific approach of building up models from core assumptions. To do so in this particular situation, we turn to the work of physicist Dr. Arnd Leike of the University of Munich who won an "Ig Nobel Prize" in 2002 for developing a model for head inspired by the decay of radioactive isotopes[1]. The model

[1]Leike, A. (2002)*Demonstration of the Exponential Decay Law Using Beer Froth.* European Journal of

was based on the following assumption:

The volume of head decreases at a rate proportional to the current volume

We are ultimately interested in the height $h(t)$ of the beer froth after t seconds (say in mm), but the statement above involves volume so we first define $V(t)$ as the volume of froth after t seconds (in mm^3). To translate Dr. Leike's assumption above into an equation, we will appeal to a trivially sounding, but reoccurring idea in this course which we call the **balancing principle**:

The rate of change of a quantity is equal to the difference between the rate at which new quantity is created and the rate at which the quantity is destroyed.

An easy way of remembering this is to think

$$\text{Rate of Change} = \text{Rate in} - \text{Rate out}.$$

In our current application, the quantity is $V(t)$ and the rate of change of $V(t)$ is its derivative dV/dt. After the initial pour, no new froth is being created, but our assumption implies that it is being destroyed at a rate proportional to V. "Proportional to V" means that the rate out is λV for some constant λ and therefore, $V(t)$ will evolve according to the differential equation

$$\frac{dV}{dt} = \text{“Rate in ”} - \text{“Rate out”} = 0 - \lambda V = -\lambda V.$$

To relate all this back to the height of the froth, we use the relationship $V(t) = \pi r^2 h(t)$ between the height and volume of a cylinder where r is the radius of the cylinder's base. From this relationship, we can see that

$$h(t) = cV(t)$$

where $c = 1/(\pi r^2)$ is a constant. This relabeling of constants is a common trick we shall use to keep our computations from getting too messy. If we differentiate both sides of the above relationship with respect to t, we obtain the expression

$$\frac{dh}{dt} = c\frac{dV}{dt} = c(-\lambda V) = -\lambda h \tag{1}$$

where in the second equality, we use our previously derived expression for $V(t)$ and in the last, we use the substitution $h(t) = cV(t)$.

Separating Variables

Comparing what we have done so far with the example from the prologue, we have basically accomplished task (I) : translating dynamic information into a DE for the unknown quantity of interest $h(t)$. Because we are viewing h as an evolving function of t, we will often refer to h as the ***dependent*** variable in the problem and t as the ***independent*** variable although this nomenclature can be ambiguous at times. We would like to be able to determine a

Physics, 23, 21–26.

general solution to Equation (1) which we can use to track froth height. In our Hayabusa example, we simply integrated both sides of the equation and if the DE in (1) read

$$\frac{dh}{dt} = -\lambda t$$

we could do the same thing here and obtain $h(t) = -\lambda t^2/2 + c$ for some constant c. But that is not what (1) says: both the left and right side involve the unknown, evolving function $h(t)$ which could have some complicated relationship with t such as $h(t) = e^{-\ln(\sin(t))}$. Instead we will "separate variables[2]", dividing both sides by the dependent variable h (which is valid as long as $h(t) \neq 0$!) and multiplying by the differential dt to obtain an alternate expression in which all quantities depending on h are on one side of the equation and all quantities involving t are on the other[3]. This leads to the equation

$$\frac{1}{h}\, dh = -\lambda\, dt.$$

Now we can integrate both sides:

$$\begin{aligned} \int \frac{1}{h}\, dh &= -\int \lambda\, dt, \\ \ln h &= -\lambda t + c_1 \end{aligned}$$

where c_1 is an integration constant. We do not need absolute value signs around h here because we know the quantity of interest, the height of the foam, is non-negative. Then we can exponentiate both sides to solve for h:

$$h(t) = e^{-\lambda t + c_1} = e^{c_1} e^{-\lambda t} = ce^{-\lambda t} \tag{2}$$

with $c = e^{c_1}$. This is essentially the completion of stage II in the prologue: we have obtained a general solution to Equation (1) which has two unknown constants c, λ. We will call such constants ***parameters*** because their values depend on experimental data.

Parameter Estimation

The next step is to match our general solution to the data in Table 1 by finding appropriate values for the parameters c, λ. This is like Stage III of the prologue where we plug in static information about the variables of interest. The data point implies that $h(0) = 47$ mm so

$$c = h(0) = 47.$$

The second line then leads to the equation

$$41 = h(5) = 47e^{-5\lambda}$$

which has solution $\lambda \approx 0.0273$. Plugging the estimated parameter values into Equation (2) yields the "naive" model $h(t) = 47e^{-.0273t}$. And using this naive model we can now answer our original question of interest by finding the time t at which the head reaches 10 mm:

$$10 = 47e^{-.0273t} \iff t \approx 57 \text{ seconds}.$$

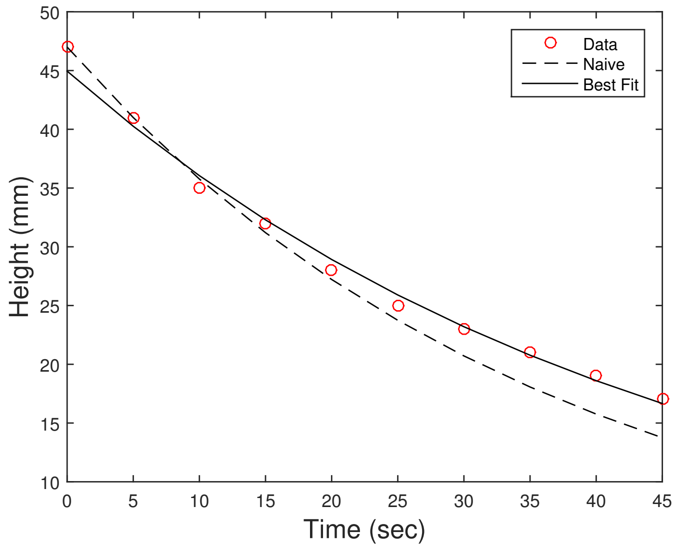


Figure 2: Data from beer pouring experiment (red dots) showing foam height (mm) vs. time (sec) along with the following theoretical fits: (i) using the naive estimate based only on the first two data points in our experiment) and (ii) using Matlab's polyfit function (black solid) to get $\lambda \approx -0.0221, c \approx 44.95$.

Figure 2 shows the predictions of the "naive" model alongside the actual data from the experiment. Note that while the fit is initially right on, the model starts to underestimate the froth height at later times and hence will most likely underestimate the time we will need to wait for the froth to subside to 1 cm. The reason why we call this the naive model is because we only used our first two data points to estimate parameter values and ignored the rest of the data. The problem of parameter estimation from data is a complex mathematical problem and many courses (such as Math 110, Math 145, or Math 132) delve more deeply into the topic of choosing the "best" estimate. Matlab's "polyfit" function is capable of performing some of these more sophisticated algorithms by calculating best fitting polynomials to sets of data. In our case, we do not want to fit a polynomial, but an exponential so we first need to transform our model (2) into a linear model by taking logarithms of both sides:

$$\ln h(t) = \ln \left(h_0 e^{-\lambda t}\right) = \ln h_0 - \lambda t.$$

In other words, since h is exponential, $\ln h$ is linear. Therefore, we can estimate h_0 and λ by finding the best fitting line $a + bt$ for the data $(t_i, \ln h_i)$ and then setting $a = \ln h_0$ and $b = -\lambda$ to solve for h_0 and λ. The solid line in Figure 2 shows the resulting fit which we can see is much better than our previous naive method. In fact, if we use this more clever model to answer our original question (when will the forth subside to 10 mm), we get the estimate $t \approx 68$ seconds which seems more reasonable from the graph[4].

[2]Please see Section 1.4 of Edwards and Penney for a more formal proof of why Separation of Variables works. Spoiler alert: its the chain rule.

[3]The constants can go on either side

[4]Admittedly, it took 45 seconds to get this estimate and many more to compute it so we might as well have just waited 68 seconds in this case.

Even after using this more sophisticated method for parameter estimation though, the model is far from perfect. Do the discrepancies come from measurement error alone? Or was the Dr. Leike's simplifying model bogus? These are complicated questions which someone getting paid to model the decay of beer foam would need to answer (and if such a job ever exists, sign me up!)

Exercises on exponential growth and decay

Differential equations of the form $y' = ay$ show up frequently in applications and can always be solved using the methods described in this section. The following exercises give some additional examples. Please see also the Section 1.1 Matlab Supplement for related problems.

(1) In the early stages of a population's growth, it can often be assumed that the population grows at a rate proportional to its current size. This assumption is reasonable as long we assume that growth due to new births dominates the growth rate due to immigration and that deaths due to resource limitations are negligible, and leads to what is called the Malthusian Law of Population Growth. The table below shows some values for the US population according to census measurements from 1790 to 1890

Year	Population
1890	62,947,714
1880	50,155,783
1870	38,558,371
1860	31,443,321
1850	23,191,876
1840	17,063,353
1830	12,860,702
1820	9,638,453
1810	7,239,881
1800	5,308,483
1790	3,929,214

(a) Translate the Malthusian Law into a differential equation for $p(t)$, the US population in millions t years after 1790.

(b) Determine the general solution to your differential equation from part (a) and use the US population measurements from 1790 and 1800 to estimate any unknown parameter values.

(c) Use your equation from (b) to estimate the *doubling* time of the US population (that is, the time it took the population to double its 1790 value). How does the model's estimate compare with what the data suggests?

(d) What happens if you try to use your equation to estimate the 1870 population? Why do you think the answer is so far off (think about what happened in the 1860s)?

(2) As previously mentioned, the beer froth model was inspired by modeling the decay of radioactive isotopes which is know to satisfy the assumption of "proportional decay". This knowledge is at the root of the process of Radiocarbon dating, a technique commonly

used (and misused!) in movies to date fossils or other samples of once living matter. Carbon-14 is a particular radioactive isotope with a half-life of 5730 years and therefore, the age of a sample containing residual amounts of this isotope can be estimated from the current amount of carbon-14 in the remains if one assumes that (i) the rate at which carbon-14 disintegrates is the same now as it was when the sample first died and (ii) the amount of carbon-14 in the sample at the time of its death is the same as the amount of carbon-14 in a currently living version of the sample[5]. Assumptions (i) and (ii) imply that if $r(t)$ is the rate at which carbon-14 is currently disintegrating (measured in disintegrations per minute per gram or dpm), then $r(t)$ is a solution to the initial value problem

$$\begin{aligned} \frac{dr}{dt} &= -ar \\ r(0) &= r_0 \end{aligned}$$

where t is the current age of the sample, r_0 is the disintegration rate of carbon-14 in a living sample, and a is the decay rate of carbon-14. In 1950, researchers discovered a portion of the Lascaux Cave in France containing charcoal giving an average carbon-14 reading of 0.97 dpm [6]. If living wood (from which charcoal is derived) gives an average reading of 6.68 dpm, estimate the age of the charcoal.

(3) Caffeine in your bloodstream decays at a rate proportional to its current amount with a half-life of about 6 hours. Suppose that Johnny takes a 200 mg double shot of Espresso at noon. At what time will the amount of caffeine in his bloodstream reach 150 mg?

(4) Statements along the lines of "assume the rate of change of a quantity is proportional to its current value" were seen in many of the above exercise including the Beer Froth, Radioactive Decay, Caffeine, and Population Growth models. This assumption is actually based on a more intuitive analysis of birth and death rates. To illustrate, suppose that you have some population of individuals in which each individual gives birth to an average of b new individuals per time period and has an average lifespan of d time-units. Let $Q(t)$ denote the population size after t time-units[7]. Note that b is often called the per capita birth rate for the population.

(a) Explain why there will be an average of bQ new individuals born per time unit.

(b) Explain why there will be an average of Q/d individuals die per time unit.

(c) Combine (a) and (b) with the Balancing Principle to conclude that

$$Q' \approx rQ$$

where $r = b - 1/d$. In other words, the rate of change of Q is proportional to Q.

[5]Carbon dating is often misused in movies. For example, in Prometheus, Elizabeth Shaw uses it to estimate the age of a dead alien specimen on another planet, but to do so, she would need to know the current level of carbon-14 in a living alien of that same species and the rate at which carbon-14 decays on that planet. It isn't clear why she would have either piece of knowledge in her arsenal.

[6]Braun, M. (1983) *Differential Equations and Their Application*, 3rd Edition. Springer, New York

[7]These assumptions are reasonable in a population with unlimited resources where the birth and death rates stay constant over time

(5) A researcher believes that a certain type of tumor cell has an average lifespan of d days and while alive, divides an average of b times per day to create two identical cells. Tumors are typically detected once they reach a size of about 10^6 cells. If a tumor initially contains 100 cells and two days later contains 122 cells, how many days will it take until the tumor is detectable (i.e. contains 10^6 cells)?

(6) A more general version of the "proportional growth" assumption occurs when we believe that one quantity, say q_1, changes at a rate proportional to another, say q_2. Whenever you see this statement, it means that $q_1' = cq_2$ for some constant c. As an example, suppose that a spherical capsule with an initial radius of 3 mm melts at a rate proportional to its current surface area. After 3 seconds, the radius has decreased to 2 mm. Setup and solve a differential equation for the radius of the capsule after t seconds. Use your results to determine the time it takes for the capsule to completely dissolve.

(7) Let $y(t)$ be a solution to the differential equation $y' = ay$ with initial condition $y(0) > 0$. Describe how

$$\lim_{t\to\infty} y(t)$$

depends on the value of a. For example, what is the limit for $a < 0$? How about $a > 0$?

(8) Suppose that instead of a cylindrical glass, one uses a conical or half-spherical goblet in the beer pouring experiment. In both cases, determine a differential equation satisfied by h.

(9) Conduct your own beer pouring experiment (if over 21) or soda experiment (if under) and record the results. Include a comparison with our in-class experiment and a plot of your results with an appropriate exponential fit. For soda, I recommend getting a 2-liter bottle and shaking it up instead of pouring. Alternatively, pouring a soapy-water solution may work better.

Matlab Supplement 1: Importing and Plotting Data

Objectives:

- Manipulate vectors in Matlab
- Create 2D plots in Matlab with axes labels and legends
- Import Data into Matlab
- Access entries, rows, and columns of Matlab variables
- Use exponential, logarithm, and sinusoidal functions in Matlab

MATLAB is short for *matrix laboratory*. This is not just a clever name. MATLAB was designed to manipulate matrices. At this point in your mathematical journey, the word matrix may not mean much, but that will change over the course of this semester. For now, think of a matrix as an array of numbers which has a specified number of rows (horizontal cross sections) and columns (vertical cross sections)

Throughout the MATLAB supplements in this book, we will slowly develop the syntax and use of the software through a series of activities connected to our study of ODEs. We will be entering commands into MATLAB and displaying the **input** in dark font. We recommend that you follow along with an open copy of MATLAB and practice entering these inputs into your command window as well. In this way, you will re-enforce what you are reading with actual muscle memory. You can then check the *output* you get against what we got in the book (displayed in light font). If they don't match, go back and try again.

For example, open up a new MATLAB command window. Then type the following series of commands, hitting "enter" after each step

```
x=[1;2]
```

```
x =

     1
     2
```

```
y=[-1;2]
```

```
y =

    -1
     2
```

```
x+y
```

```
ans =

     0
     4
```

Do you see what happened? You created two vectors x, y and took their vector sum. Vectors are actually the first example of matrices we shall encounter: they are arrays of numbers with 1 column and n rows.

Vectors are foundational for understanding plotting in MATLAB's. Suppose we want to create a plot of the function $y = x^2$ over the interval $x \in [0, 1]$. The first step is to define a vector of x values from 0 to 1 with equal spacing. Let's start with a fairly large spacing of 0.5

```
x=[0;0.5;1]
```

```
x =

         0
    0.5000
    1.0000
```

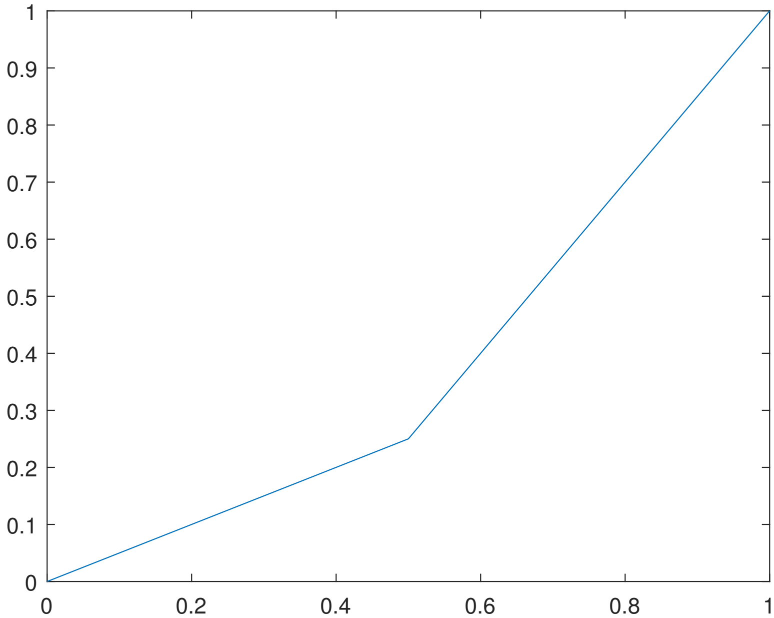
Now we calculate y^2 for each of the x values above.

```
y=x.^2
```

```
y =

         0
    0.2500
    1.0000
```

We will get more into the "." before the exponent later; for now just use it whenever you want to apply an operation (power, multiplication, etc.) to a vector of values. Finally, we get to the plot of y vs x:

```
plot(x,y)
```



Note the "jagged" nature of the plot. That happened because we used too coarse of a spacing for our x values. We could fix this by typing the command "x=[0;0.1;0.2;etc]", but that gets kind of tedious. To this end, we will make use of the colon operator

```
x=0:.1:1
```

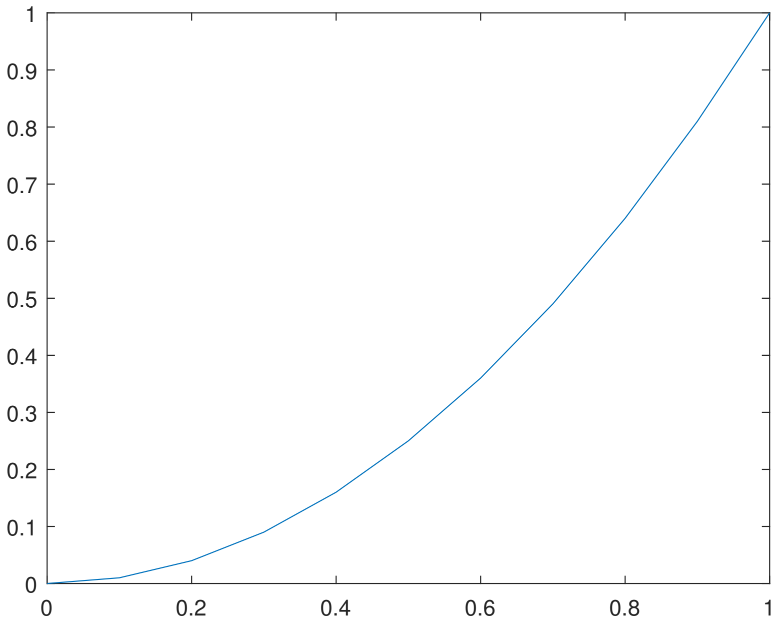
```
x =

  Columns 1 through 7

         0    0.1000    0.2000    0.3000    0.4000    0.5000    0.6000

  Columns 8 through 11

    0.7000    0.8000    0.9000    1.0000
```

The colon operator creates a list of values, starting at the first number and ending at the last with increments given by the middle. You can "suppress" the output (which gets annoying to look at) by placing a semi-colon at the end of a line

```
x=0:.1:1;
```

Now we create a new vector of y values and plot

```
y=x.^2;
plot(x,y)
```



That looks better! We'll explore more plotting options later in these notes.

As a second example of matrices in action, suppose we wish to manipulate a data set (such as our beer pouring data) within Matlab. There are two easy ways to import data from excel files into Matlab. Method 1 is to simply open Matlab by double clicking the icon on your desktop, and drag the excel file of interest into the Matlab command window. The import wizard, a user interface showing the variables which Matlab will import, is displayed. Select "numeric matrix" on the top and then click "Import Selection". A variable with the name of your file should now appear in your workspace. I am going to do the rest of this tutorial with Keller.xlsx, which can be downloaded from Canvas so please try this method with this dataset now if you prefer this drag and drop method

For Method 2, just type

```
Keller=xlsread('keller.xlsx')
```

```
Keller =

     0    47
     5    41
    10    35
    15    32
    20    28
    25    25
    30    23
    35    21
    40    19
    45    17
```

into the Matlab command window and Matlab will do the work for you. Either way, if you followed directions properly, you should end up with an ordered array or matrix of numbers called 'Keller'. The first column contains the time measurements (in sec) from a beer froth experiment while the second contains the corresponding heights (in mm). You can learn the size of the matrix by

```
size(Keller)
```

```
ans =

    10     2
```

The first number in the output is the number of rows and the second is the number of columns. This is also the order you use if you want to access a particular entry in the data frame

```
Keller(2,1)
```

```
ans =

     5
```

vs.

```
Keller(1,2)
```

```
ans =

    47
```

To make a plot of the heights vs times, it is convenient (although not necessary) to create separate vectors for the times and heights.

```
t=Keller(:,1)

h=Keller(:,2)
```

```
t =

     0
     5
    10
    15
    20
    25
    30
    35
    40
    45


h =

    47
    41
    35
    32
    28
    25
    23
    21
    19
    17
```

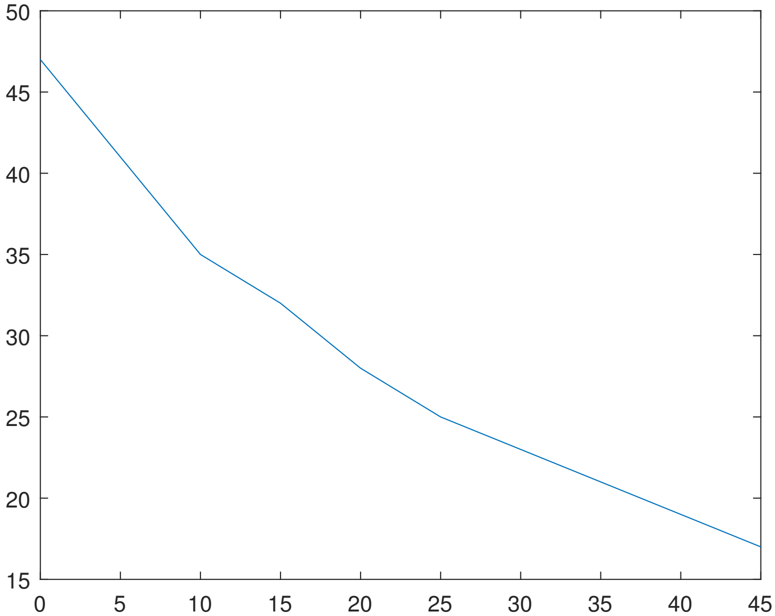
Note the colon comes first since we want all rows to be selected. Before proceeding, take a moment to compare the output t above with the output of the command:

```
Keller(1,:)
```

```
ans =

     0    47
```

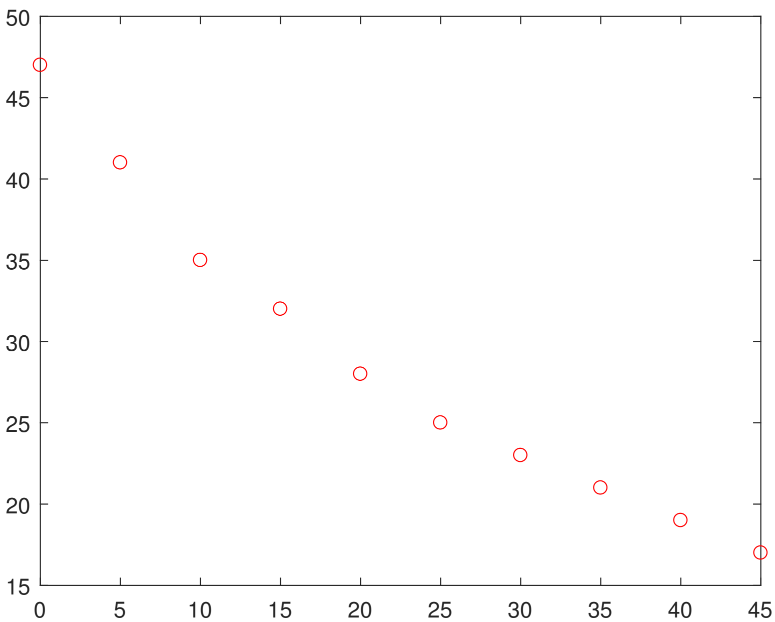
Now that we have defined our x-axis variable t and y-axis variable h, we simply plot h vs t as follows

```
plot(t,h)
```



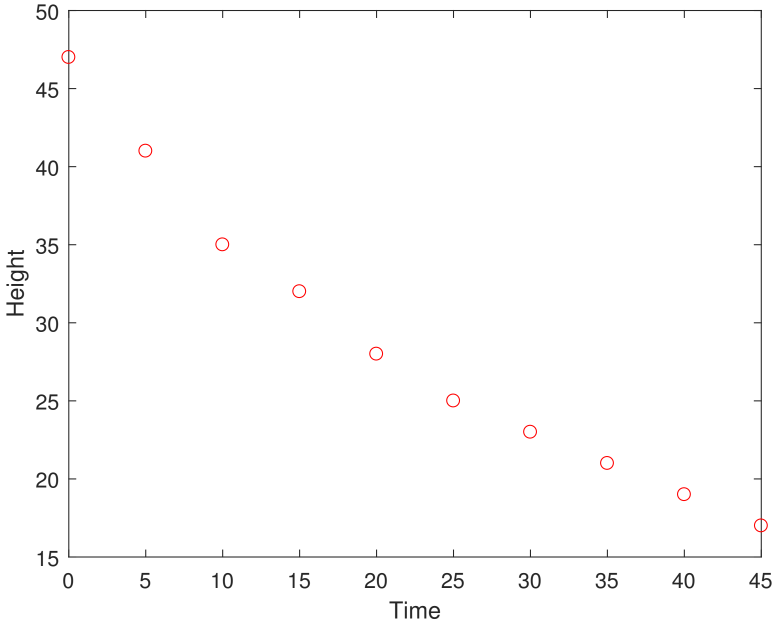
You can check out additional options of the plot function by typing 'help plot'. One particulary interesting modification is to change the straight lines to red circles

```
plot(t,h,'ro')
```



the 'r' changes the color to red; 'ro' makes red o's. Another option is to add axes labels

```
plot(t,h,'ro')
xlabel('Time')
ylabel('Height')
```



You can also compute new vectors from the old. For example, suppose that we want to see how well the model $h(t) = ae^{bt}$ (see Section 1.1) fits the data. First, we set $a = h(0)$ and $\lambda = -(1/5)\ln(h(5)/h(0))$. Matlab does not allow you to call variables by greek letters so we will just use b for λ. Also note that indexing in Matlab starts from 1 though so that 'h(1)' in Matlab is $h(0)$ for our function, 'h(2)' in Matlab is $h(5)$ for our function, and so on. Note also that 'log' means natural log in Matlab

```
a=h(1)
b=-(1/5)*log(h(2)/h(1))
```

```
a =

    47


b =

    0.0273
```

After computing the estimates of a, b, we calculate the values of h estimated by our model. Note that Matlab does this in one line

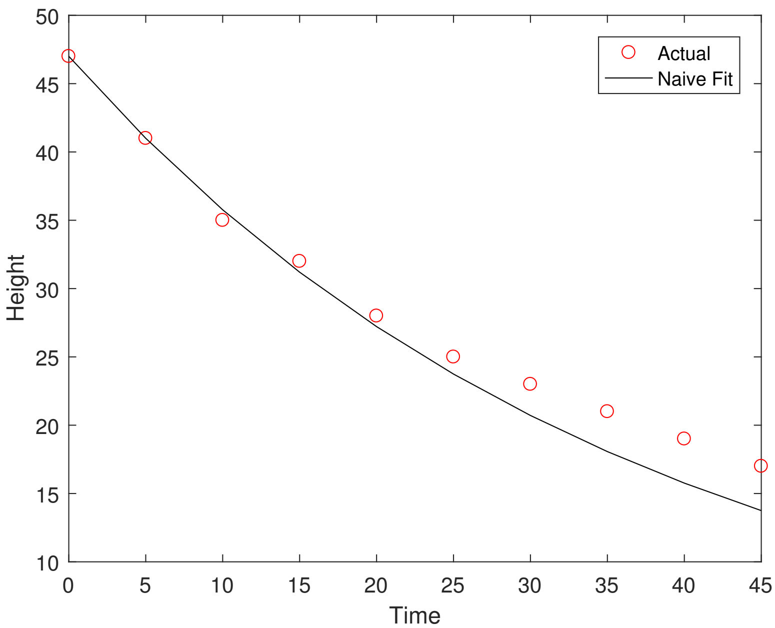
```
hmodel=a*exp(-b*t)
```

```
hmodel =

   47.0000
   41.0000
```

```
35.7660
31.2001
27.2171
23.7426
20.7116
18.0676
15.7611
13.7490
```

Then we add these estimates to our previous plot as a black line. The final line adds a legend - see 'help legend' for more information

```
plot(t,h,'ro',t,hmodel,'k')
xlabel('Time')
ylabel('Height')
legend('Actual','Naive Fit','Location','NorthEast')
```



Once you have created a figure (like the one above), you can save it as an image file by going to File > Save As and then changing the 'Save As Type' to your favorite extension. And now you are ready for some practice!

Exercises

1. Create a Matlab graph which compares the actual census measurements of the US population from 1790 to 1990 with the values predicted by your model from Problem 1 in the Section 1.1 Exercises. You will have to look up the values for 1890-1990 online. Write a few sentences analyzing the "fit".

2. As we have seen, solutions to the differential equation $y' = ay$ always have the form $y(t) = ce^{at}$ for some constant c. Create a single plot which compares the graphs of ce^{at} over the interval $[0, 1]$ for the following combinations of c, a.

   (a) $c = 1, a = 1$

   (b) $c = -1, a = 1$

   (c) $c = 1, a = -1$

   (d) $c = -1, a = -1$.

3. Create Matlab plot to compare the following functions over the interval $0 \leq t \leq 6\pi$ using the following spacings between points on the x-axis: 1 and 0.001. What happens if you choose the spacings too far apart?

   (a) $y_1(t) = (4/\pi)\sin(t)$

   (b) $y_2(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t)$

   (c) $y_3(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t) + (4/(5\pi))\sin(5t)$

Section 1.2: Differential Equations from Physical Principles

Class objectives:

- Setup linear first order differential equations based on modeling assumptions.
- Solve equations of the form $y' = ay + b$ using separation of variables and investigate the solution.
- Describe the difference between a general solution to a differential equation and a particular solution to an initial value problem.

Last lecture, we considered several applications in which the differential equation for an unknown function $y(t)$ had the form:

$$y' = ay \tag{1}$$

where a was a constant parameter. By separating variables, we were able to show that this equation yielded solutions of the form

$$y(t) = y_0 e^{at}$$

where $y_0 > 0$ was the value of y at time 0. Equation (1) is the simplistic example of what we will call a first order, ***linear*** differential equation: first order (as we mentioned in the Prologue) because it involves only a first derivative of y and linear because the dependence on y and y' is linear. Linear equations will make up about 90% of the course content, but before getting into the more general theory of how to solve linear first order equations (and give a more detailed discussion of what they are and are not!), we will spend the next couple of classes building up slightly more complicated examples. This section will focus on differential equations that arise in the translation of two important physical principles.

Model for a Cooling Light Bulb

Figure 1 shows the results of an experiment in which a light bulb was heated to a temperature of approximately 120°C and then turned off. After 30 seconds, measurements of the light bulb's temperature were recorded every five seconds for approximately five minutes. Notice that the rate of decay does appear to be exponential, but approaches a constant.

We will apply the following principle, known as Newton's Law of Cooling (NLC), to set up a model for heat loss in this scenario:

(NLC) The rate of change of an object's temperature is proportional to the temperature difference between the object and its surrounding medium.

To translate this into a differential equation, let T_m denote the temperature of the surrounding medium (i.e. the room) and $T(t)$ be the temperature of the bulb in degrees Celsius after t seconds. Then the statement above translates to the equation:

$$\frac{dT}{dt} = -k(T(t) - T_m) \tag{2}$$

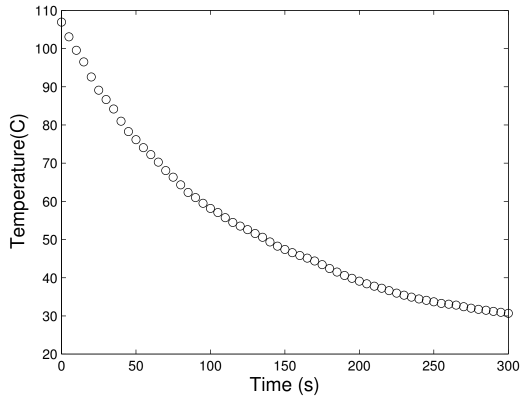


Figure 1: Plot of Temperature (in degrees Celsius) vs. Time (in seconds).

where $k > 0$ is an unknown constant of proportionality (something we should estimate from the data). The negative sign is there because while the temperature of the bulb is greater than the temperature of the surrounding medium, the temperature should be decreasing. Separating variables, we rewrite this as

$$\frac{1}{T(t) - T_m}\,dT = -k\,dt$$

which we can now integrate to obtain

$$\ln|T(t) - T_m| = -kt + C.$$

Assuming that $T > T_m$ (the bulb is hotter than its surroundings, we can solve for T by exponentiating both sides of the equation:

$$T(t) = T_m + e^{-kt+C} = T_m + ce^{-kt}.$$

where $c = e^C$ is another arbitrary constant. If we are given an initial temperature $T(0) = T_0$, then $c = T_0 - T_m$ and we obtain the solution

$$T(t) = T_m + (T_0 - T_m)e^{-kt}.$$

Since the power of the exponential term is negative, we have $T(t) \to T_m$, the temperature of the medium as $t \to \infty$ as our intuition might suggest. Exercise 3 will ask you to estimate the unknown parameters in this equation from the data used to make Figure 1

Example

As a numeric example of Newton's Law of Cooling, suppose that a 92° C cup of coffee is placed in a 25° room. After 20 seconds, it has cooled to 91°. How long will it take for the coffee to reach a drinkable 80°? To answer this question, let $y(t)$ be the coffee cup's temperature (in C) after t seconds. Then Newton's Law of Cooling implies that

$$y' = -k(y - 25)$$

where k is a positive constant. Separating variables and integrating as in the light bulb example, we obtain

$$y(t) = 25 + ce^{-kt}$$

where c is another constant. At this point, we plug in our two static conditions $y(0) = 92, y(20) = 91$ to obtain

$$92 = y(0) = 25 + C \iff C = 67$$

and

$$91 = y(20) = 25 + 67e^{-20k} \iff k = -\frac{1}{20}\ln(66/67).$$

Since the coffee is drinkable when $y(t) = 80$, solving for t gives

$$80 = y(t) = 25 + 67e^{-kt} \iff t = -\ln(55/67)/k \approx 262 \text{ seconds}.$$

Looks like you have time to use the restroom first!

Newton's Second Law

Newton's Second Law (NSL[1]) states that the sum of the forces acting upon an object is equal to the product of the object's mass and acceleration. This is often simply stated as $F = ma$ where F is the sum of all forces acting upon the object (usually measured in Newtons, N), m is the object's mass (usually in kg), and a is the objects acceleration (usually in m/s^2). Since NSL involves acceleration, it can be translated into the first order differential equation

$$mv' = F$$

for an object's velocity v or the second order equation

$$my'' = F$$

for an object's displacement y. We shall encounter numerous applications of both throughout the semester, many of which you may also see (or have seen) in Phys 53.

One simple application of NSL is in the study of linear motion. For example, consider the falling body from Exercise 1 in the Prologue. Let $y(t)$ and $v(t)$ to be the distance fallen and velocity of the body after t seconds (in m and m/s, respectively). To set up a differential equation for v, we need to quantify the forces acting upon the body. Two things come to mind:

Gravitational Force (F_g) : According to classical mechanics, the gravitational force acting upon a falling object near the Earth's surface is constant and proportional to the object's mass. This leads to the equation $F_g = mg$ where $g \approx 9.81\, m/s^2$ is a known constant.

Drag Force (F_d) : Drag opposes changes in speed (In other words, the faster an object is moving, the more drag will act to slow it down). Drag force is proportional to the square of an object's velocity, but acts in the opposite direction of motion so $F_d = -kv^2$ where k is a positive constant. The value of k depends on the shape/size of the object being dropped and the density of the surrounding medium (air, water, etc).

[1] Like NLC, NLS is a basic principle or assumption that may be valid in some situations and not in others, however we'll let you take Modern Physics to find out more about that.

If we ignore the second force (as you usually do in Phys 53), then NSL says that

$$mv' = mg \tag{3}$$

or $v' = g$. This leads to the standard "kinematic equations" from Phys 53 (see Section 0). However, if we add drag forces, we obtain the more complicated differential equation

$$mv' = mg - kv^2 \tag{4}$$

or $v' = g - \frac{k}{m}v^2$. Despite being nonlinear, this equation is actually separable and can be solved using partial fractions or hyperbolic functions, but we won't go there right now. Instead, we would like to note that in some cases[2], we can approximate drag by a force proportional to velocity. This yields the alternative, now linear DE

$$mv' = mg - kv \tag{5}$$

which is actually the same mathematical equation as NLC with different labels for the constants. Therefore, solving a generic version of this equation (see Exercise 5 below) actually allows one to tackle either application.

The challenge with real world examples of NSL is estimating k. This is an interesting, but somewhat tangential diversion from our course and we encourage you to take fluid dynamics if you wish to learn more about such things. For now, just know how to translate NSL from words into a differential equation and solve the resulting equation when possible (see Exercise 4 below).

General vs. Particular Solutions

When I was in high school, I saw math as an unconnected set of random problems that were sometimes interesting, but mostly irrelevant to my life. That's the danger of focusing too much on individual problems and not enough on how they tie together through deeper principles. You miss the forest for the trees so to speak. The trees are beautiful in their own way, but you'll get lost if you don't have a map to follow. So even though this is a course in applied differential equations, we are going to pause every once and a while to talk about bigger picture stuff.

The Prologue introduced a four-step algorithm for modeling with differential equations. Now that we have seen some more examples, let's go back and re-examine that process. First, we have the modeling phase where we translate words into mathematical equations. This will often be the most challenging and diverse phase of the process, but so far it has involved translating simple assumptions like NLC or the proportional growth assumption into differential equations relating dynamic quantities.

Second, we have the solving phase. This is where we develop techniques (like SoV) for finding both *general* and *particular* solutions to differential equations. If we take the DE

$$y' = -2y$$

for example, SoV yields the general solution $y(t) = ce^{-2t}$. This formula actually represents a whole family of different solutions obtained by plugging in different values for c, some

[2] in short, small objects or viscous mediums

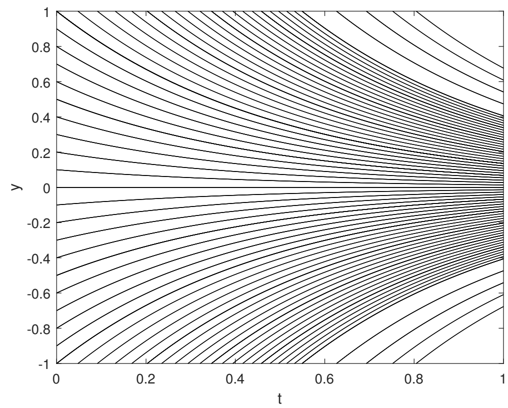


Figure 2: Some solutions to the differential equations $y' = -2y$ found by substituting in different values for c in the general solution $y(t) = ce^{-2t}$.

examples of which are shown in Figure 2 below. Sometimes, we will only be interested in finding this general solution and analyzing its behavior. In such examples, we can skip the third PSI phase and proceed to the final analysis phase of modeling. For example, we can see that all solutions in the figure below converge to 0 as $t \to \infty$. Therefore, if our question of interest is just to analyze the asymptotic behavior of solutions to $y' = -2y$ as $t \to \infty$, then we are done. It is only situations that supply additional requirements (eg. $y(0) = 1$) which require us to complete the PSI phase and obtain particular solutions to the equation (eg. $y(t) = e^{-2t}$).

The most common situation in which we will require a particular solution is when we are given both a differential equation AND a specific initial condition $y(0) = y_0$. We shall refer to such examples as ***initial value problems*** or IVPs for short. Many exercises will ask you to find a solution to a given initial value problem and validate the solution.

To demonstrate these instructions, suppose we wish to find and validate a solution to the IVP

$$y' = -2y, \ y(0) = 2.$$

Then we first find the general solution $y(t) = ce^{-2t}$ and proceed to plug in our initial condition $y(0) = 2$ to find that $c = 2$. Thus, the solution to this initial value problem is $y(t) = 2e^{-2t}$.

"Validate your solution" is a fancy way of saying "check your answer". Soon we will see how to do this using Matlab, but for now you can validate solutions by plugging them back into your IVP. In our example, this means we differentiate our solution and show that $y' = -2y$:

$$y' = -4e^{-2t} = -2(2e^{-2t}) = -2y.$$

Then we plug in $t = 0$ and show that we get $y(0) = 2$. Since both equations how, our solution is validated.

In a more theoretical course on differential equations, we would spend more time at this point looking at the existence and uniqueness of solutions to differential equations. In short, this boils down to two questions: when does a given IVP have a solution and when is that solution unique? Section 1.3 of Edwards and Penney has a good discussion on this topic for the interested reader with some examples of situations where solutions don't exist or exist, but are not unique. In fact, one of the Clay Institute's "Millenium Problems" in mathematics is about the existence and uniqueness of solutions to a particular differential equation from fluid dynamics (the Navier-Stokes equation). So it is, quite literally, the "million dollar question" in certain scenarios. But it won't be that relevant to the rest of this course because most of the equations we consider rather trivially have unique solutions. Therefore, we'll move onto some further applications at this point.

Exercises

(1) At 2 pm on a cool 34° F day in March, Detective Bunk measures the temperature of a dead body at 38° F. One hour later. the temperature has dropped to 36°F. Using 98°F as the approximate temperature of a living body, estimate the time of death. Then discuss some potentially unreasonable assumptions in your calculation.

(2) A 200 degree cup of coffee is placed in a 70 degree car. 10 seconds later, it has cooled to a temperature of 195. How many seconds will it take to cool to a drinkable 140?

(3) Download the data "Raw Temp Data.xlsx" from the experiment shown in Figure 1 and estimate T_m, k from the data. Create a plot which shows how the fitted model matches the data.

(4) Determine the solutions $y(t)$ to each of the following initial value problems.

(a) $y' = \frac{y}{5},\ y(0) = 2$.

(b) $y' = 2 - 4y,\ y(0) = 1$.

(c) $y' = y(1-y), y(0) = 1/2$.

For each case, verify your solution by plugging it back into both sides of the given DE.

(5) Solve the differential equation in (5) from this Section's notes assuming 0 initial velocity. Your answer will depend on m, g, k. Determine what happens to the body's velocity as $t \to \infty$ (the so-called terminal velocity of the object).

(6) Suppose that $v(t) =$ the velocity of an object after t seconds (in m/s) satisfies the initial value problem

$$v' = -9 - 3v,\ v(0) = 3$$

Find a formula for $v(t)$.

(7) One of the common themes of this course will be investigating the dependence of solutions on unknown parameters in differential equations. We conclude the exercises with a simple example describing "one ring to rule" all the models in this section. Consider the initial value problem

$$\begin{aligned} y' &= ay + b \\ y(0) &= y_0 \end{aligned}$$

where $b, y_0 > 0$ and $a \neq 0$.

(a) Determine a formula for the solution y in terms of the parameters a, b, y_0.

(b) What is $\lim_{t\to\infty} y(t)$? How does the answer depend on a, b or y_0.

.

Section 1.3: Mixing Tanks and Euler's Method

Class objectives:

- Setup equations to describe the mixing of chemicals in tanks.
- Describe Euler's Method for numerically approximating solutions to first order equations
- Implement Euler's Method in Matlab (see supplement)
- Describe properties of solutions to differential equations from numerical approximations

To summarize the models we have looked at so far:

(i) Equations of the form $y' = ay$ yield exponential solutions.

(ii) Equations of the form $y' = ay + b$ yield solutions which are constants + exponentials.

Today, we will look at an application which gives rise to a model of the form $y' = ay + f(t)$ for a more complicated function f. Such equations cannot (in general) be solved using separation of variables and hence, motivate our next important technique for investigating solutions to differential equations: Euler's Method.

A mixing problem

A pipeline carrying an undesirable chemical flows into a 100 gallon well mixed reservoir at a rate of 200 gal/day and the mixture flows out through an overflow channel at the same rate. The concentration of chemical in the incoming water is $\gamma = \frac{1}{100}$ g/gal. Determine the maximum amount of chemical which will be present in the reservoir. Assume that the reservoir initially contains pure water.

To answer this question, we start off in the modeling phase and let $y(t)$ denote the amount of chemical (measured in g) in the tank after t days. From the balancing principle in Section 1.1,

$$\frac{dy}{dt} = \text{Rate at which chemical flows in} \; - \; \text{Rate at which it flows out.}$$

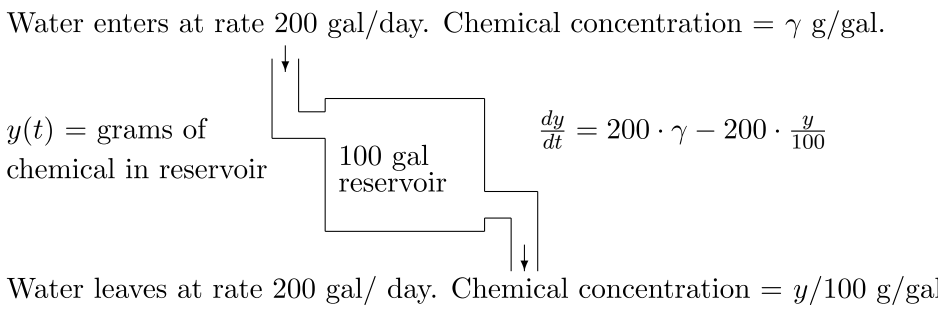
We are also told that the reservoir initially contains pure water so we have the initial condition $y(0) = 0$. Now the units for dy/dt are g/day so the right hand side should have the same units. A common mistake on mixing problems is to assume that the "rate in" part of this equation is 200 gal/day, but this is the rate at which water flows in, not chemical. To find the rate at which chemical flows into the reservoir, we multiply the rate at which water flows in, 200 gal/day, by the concentration of chemical, $\gamma = 1/100$ g/gal. Note that the resulting quantity has the right units of g/day and keeping track of units will help you avoid many pitfalls in mixing problems. To find the rate out, we multiply the rate at which water flows out, again 200 gal/day, by the concentration of chemical in the out flowing water. Assuming that the reservoir stays well mixed, this concentration can be approximated by the number of grams of chemical currently in the reservoir, which is the unknown $y(t)$, divided by the size of the reservoir, 100 gal. Piecing this information together yields the differential equation

$$\begin{aligned} \frac{dy}{dt} &= (200)\left(\frac{1}{100}\right) - (200)\left(\frac{y}{100}\right) \\ &= 2 - 2y \end{aligned} \tag{1}$$

This leads to an initial value problem

$$\begin{aligned} \frac{dy}{dt} &= -2y + 2 \\ y(0) &= 0 \end{aligned} \tag{2}$$

for the grams $y(t)$ of pollutant in the reservoir after t seconds. For more pictoral learners, all this information is condensed in the diagram below.



The IVP in equation (2) is similar to equations from the last section and can be solved using Separation of Variables to obtain

$$y(t) = 1 - e^{-2t}.$$

This equation tells us the amount of chemical in the reservoir at time t and we can make a couple of simple observations once we have the solution

- The limiting or asymptotic amount of chemical in the reservoir will be

  $$\lim_{t \to \infty} y(t) = 1 \ g$$

  while the limiting concentration (amount/volume) will be 1/100 g/gal, the same as the incoming concentration.

- The concentration of chemical in the tank will never be larger than 1 since $y'(t) > 0$ implies that y is always increasing towards its limit.

For constant input, the above two statements will always hold (see Exercises 1, 2, and 3 below).

Mixing problems get considerably more interesting when we add a time-varying input although the modeling phase is just as easy. For example, let's suppose that the concentration of chemical in the incoming water now varies periodically with time according to the expression $\gamma(t) = \frac{1}{100}(1 + \sin 2t)$ g/gal. Returning to the first line of Equation (1), we replace $\gamma = 1/100$ with $\gamma(t)$ to obtain

$$\begin{aligned} \frac{dy}{dt} &= (200)\left(\frac{1}{100}(1 + \sin 2t)\right) - (200)\left(\frac{y}{100}\right) \\ &= 2 + 2\sin 2t - 2y \end{aligned}$$

which leads to the new IVP

$$\begin{aligned} \frac{dy}{dt} &= -2y + 2(1 + \sin 2t) \\ y(0) &= 0 \end{aligned} \tag{3}$$

for the grams $y(t)$ of pollutant in the reservoir after t seconds. This new equation cannot be solved by SoV (but give it a try to make sure you are convinced!) Therefore, we will need to develop some other techniques for solving first order equations.

Euler's Method

Separation of variables is one example of an *analytic method* for solving differential equations. It is a procedure that when properly applied yields an explicit formula for the solution $y(t)$ to a given IVP. However, as mentioned above, there are some ODEs that cannot be solved with SoV. As we progress through the course, this shall be a common theme in the analytic techniques we develop: they all have strict limitations on when they can be applied.

In contrast, *numeric methods* are techniques which only yield approximate solutions of IVPs, but typically apply to much broader classes of ODEs. In other words, they are more flexible, but less accurate than analytic methods. Here we will give a brief introduction to one of the oldest. simplest, and most famous[1] algorithms of this type: Euler's Method.

For motivation, let's first consider an example of an ODE where we already know the solution, say $y' = -y$ with initial condition $y(0) = 1$. If we want to estimate the value of $y(1)$, we could use linear approximation:

$$y(1) \approx y(0) + y'(0)(1 - 0).$$

Since $y' = -y$, we know that $y'(0) = -y(0) = -1$ so plugging this in yields the approximation

$$y(1) \approx 1 - 1 = 0,$$

not a very good approximation considering the actual value is $e^{-1} \approx 0.3679$. However, we could do better by using linear approximation to first approximate $y(1/2)$ and then a second step of linear approximation to get $y(1)$. In other words,

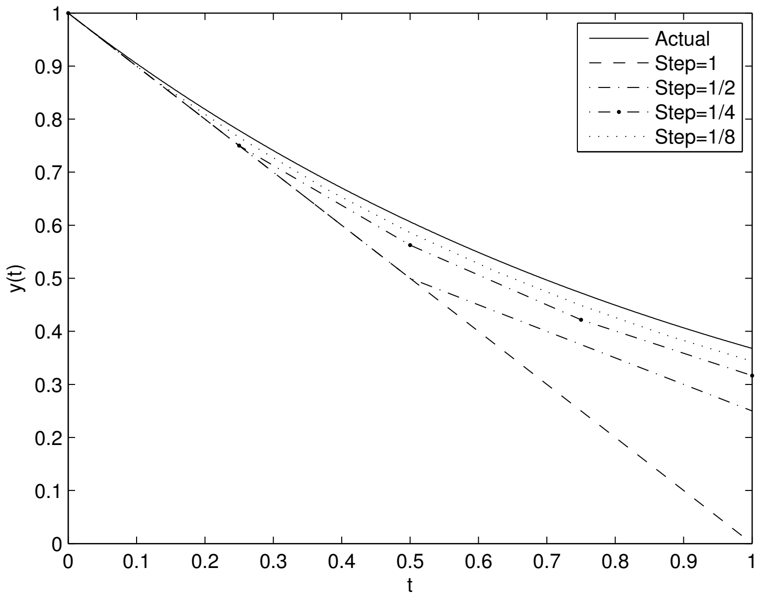
$$y(1/2) \approx y(0) + y'(0)(1/2 - 0) = 1/2$$

and then

$$y(1) \approx y(1/2) + y'(1/2)(1 - 1/2) = 1/2 - 1/4 = 1/4.$$

Much better! But why stop at intervals of $1/2$? We could use four successive approximations of size $1/4$ or eight of size $1/8$ or more of even smaller size. An example comparing the approximations obtained by this method are shown in the figure below and this process illustrates the basic idea behind *Euler's Method*.

[1] See Hidden Figures



If you must for some reason perform Euler's by hand (and yes, in this class you must do it by hand a few times to understand the process!), then it is helpful to organize your work in a table. As an example, suppose we estimate the solution $y(t)$ to the IVP

$$y' = -y + t,\ y(0) = 1.$$

at time $t = 0.2$ using a step size of 0.1. This will require two steps:

$$\begin{aligned} y(0.1) &\approx y(0) + y'(0)(0.1) \\ y(0.2) &\approx y(0.1) + y'(0.1)(0.1) \end{aligned}$$

If we label the sequence of times $t_0 = 0, t_1 = 0.1, t_2 = 0.2$ and call the time step of 0.1 Δt, both these equations can be written in the form

$$y(t_j) \approx y(t_{j-1}) + y'(t_{j-1})\Delta t$$

or perhaps more tellingly

$$y(next) \approx y(previous) + y'(previous)\Delta t. \tag{4}$$

The work required in this process is shown in the table below.

Step	t_j	$y(t_j) \approx y(t_{j-1}) + y'(t_{j-1})\Delta t$	$y'(t_j) = -y(t_j) + t_j$
0	0	$y(0) = 1$	$y'(0) = -1 + 0 = -1$
1	0.1	$y(0.1) \approx 1 + (-1)(0.1) = 0.9$	$y'(0.1) \approx -0.9 + 0.1 = -0.8$
2	0.2	$y(0.2) \approx 0.9 + (-0.8)(0.1) = 0.82$	Not needed

As you can see, implementing numerical methods by hand can get tedious because they typically rely on iteration. However, if you look carefully at the algorithm above, you can see that there is a pattern to the process: during each step, you estimate y at the current time by Equation (4) and then use the given ODE to estimate y'. Such iterative procedures are easy for computers to carry out. So in the next Section, we will see how to carry out this process in Matlab. We will also see how to implement a more advanced numerical procedure using Matlab's ode45 function.

Exercises

1. Victor Frankenstein's reanimation recipe critically depends on obtaining a 20 liter mixture consisting of water with a concentration of 1 g/liter bone dust, but because of rough economic times, his equipment has been reduced to a 20 liter vat that leaks at a rate of 1 liter/min. To obtain the desired mixture, Dr. Frankenstein initially fills the vat with pure water and immediately begins pouring in a bone dust mixture with a concentration of 2 g/liter at a rate of 1 liter/min.

   (a) Write an IVP for the amount of bone dust present after t minutes

   (b) Solve the IVP above to find a formula for the amount of bone dust in the vat after t minutes

   (c) Find the time t at which the ***concentration*** of bone dust will reach the desired level of 1 g/l.

   (d) Determine the limiting concentration of bone dust in the vat.

2. Eben's parents place an unknown amount of soap into a 4 gallon tub of water. After bathing Eben in the soapy water, they begin to pump fresh water into the tub at a rate of a quarter gallon per second and the well stirred mixture flows out of the tub through a hole in the bottom at the same rate. How long should Eben's parents wait until the amount of soap in the tub reaches 10% of its original value?

3. A 40 L tank is initially filled with pure water. A solution contained 10 g/L of salt is pumped into the tank at a rate of 4 L/min and the well stirred mixture flows out of the tank at the same rate. How much salt will be in the tank after 10 minutes?

4. (The following is a modification of Exercise 2 with the added twist that the volume of water is changing) Eben's parents place 2 grams of soap into 10 liters of water. After bathing Eben in the soapy water, they begin to pump fresh water into the tub at a rate of a 1 liter per second and the well stirred mixture flows out of the tub through a hole in the bottom at a rate of 1/2 liter per second.

   (a) Determine an appropriate IVP for the number of grams of soap in the tub after t seconds.

   (b) Solve the IVP in (a) and use your answer to calculate how long it will take for the *concentration* of soap in the tub to reach 1/10 grams/liter.

5. Bunce's farm has a 10 L water tank fed by a stream running from Boggis' Chicken Ranch. Water from the stream flows into and out of Bunce's tank at a rate of 10 L/minute which keeps it well-mixed. One day Bunce finds out that Boggis has been dumping chicken manure into the river which adds Salmonella to the stream at an unknown concentration of a mcfu/L (Note: mcfu is a unit of measurement for bacteria standing for "millions of colony forming units"). At the time of his discovery, Bunce measures the concentration of Salmonella in his tank to be 1 mcfu/L and one minute later measures it again to be 2 mcfu/L. Use these measurements to determine the concentration of Salmonella in the river **and** the limiting concentration of resulting Salmonella in Bunce's tank.

6. Estimate $y(1)$ for the solution to $y' = 2y$ with initial conditions $y(0) = 1$ using Euler's Method with two steps and then again with four steps. How does his estimate compare with the actual value of e^2?

7. Repeat the previous exercise for the differential equation $y' = -2y$ and the same initial conditions. In this case, do you get an over or under estimate?

8. In general, what do you think we can say about using Euler's Method to approximate solutions to differential equations of the form $y' = ay$?

9. Use Euler's Method with step size $\Delta t = 1/2$ to estimate $y(1)$ for the solution y to

$$\begin{aligned} \frac{dy}{dt} &= -4y\cos(2\pi t) \\ y(0) &= 1. \end{aligned}$$

   Then separate variables to solve this IVP and compare Euler's approximation with the exact solution $y(1)$.

10. Toricelli's Law says that the height of water in a cylindrical tank draining through a hole at the bottom will satisfy the first order equation

$$\frac{dh}{dt} = -k\sqrt{h}$$

   where t is time in seconds, h is the height of water in meters, and k is a constant. Suppose that $k = 2$ and that the initial height of the water is four meters. Use Euler's method with step size 0.5 to estimate the height of the water after one second. Then solve the equation using separation of variables and compare the results.

Matlab Supplement 2 - Numerical Methods in Matlab

Euler's In Matlab

In Section 1.3, we introduced Euler's Method, a low level numerical scheme for finding approximate solutions to first order equations. By hand, Euler's is slow and tedious, but it lends itself well to implementation via "for" loops. To demonstrate, suppose we want to solve the initial value problem $y' = -y + t$, $y(0) = 1$ for $0 \leq t \leq 1$. Euler's method says start by choosing a time step Δt and then computing $y(\Delta t)$ via the linear approximation

$$y'(\Delta t) \approx y(0) + y'(0)\Delta t$$

Note that in our example, $y'(0) = -y(0) + 0$ so we can substitute that above. Then we estimate $y(2\Delta t) \approx y(\Delta t) + y'(\Delta t)\Delta t$ and so on until we get to $y(1)$. If we take $\Delta t = 1/2$, for example, then we would need to store the values of y at the three times $0, 1/2, 1$. So we start by defining a vector of times t and empty vectors of values for y, y' of the same length

```
t=0:1/2:1
y=zeros(1,length(t))
dy=zeros(1,length(t))


t =

         0    0.5000    1.0000


y =

     0     0     0


dy =

     0     0     0
```

We then fill in the initial values of y, y' in the first elements

```
y(1) = 1
dy(1)=-y(1)+t(1)


y =

     1     0     0
```

```
dy =

    -1     0     0
```

We proceed to Step 1 of Euler's Method: computing $y(1/2)$. This will be stored in the second element of y. We then compute $y'(1/2)$ for use in the next step and store this in the second element of y'

```
y(2)=y(1)+dy(1)*(1/2)
dy(2)=-y(2)+t(2)
```

```
y =

    1.0000    0.5000         0


dy =

    -1     0     0
```

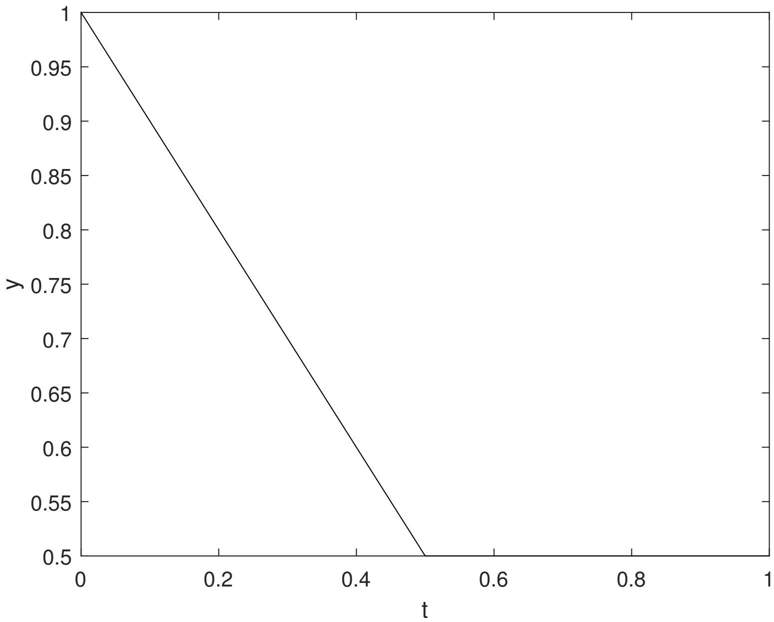
Finally, we do the second step of Euler's to estimate $y(1)$ and $y'(1)$:

```
y(3)=y(2)+dy(2)*(1/2)
dy(3)=-y(3)+t(3)
```

```
y =

    1.0000    0.5000    0.5000


dy =

   -1.0000         0    0.5000
```

Plotting the three values in y against their times shows us the approximate solution over the interval $[0, 1]$

```
plot(t,y,'k')
xlabel('t')
ylabel('y')
```



If we want to use a smaller step size, say $\Delta t = 1/10$, the above gets tedious. To avoid this, we need a general formula which relates the jth entry in y with the previous (j-1)st entry. Looking at the examples above with $j = 2$ and $j = 3$ show that $y(j) = y(j-1) + y'(j-1)\Delta t$ provides the desired relationship. So suppose we want to use $\Delta t = 1/10$, requiring us to store 11 entries in y. Then the initialization (with MATLAB comments) is:

```
dt=0.1;                 % Define Step Size
t=0:dt:10;              % Define interval of times
y=zeros(size(t));       % Initialize variable for storing y
dy=zeros(size(t));      % Initialize variable for storing dy

y(1)=1;                 % Enter initial y
dy(1)=-y(1)+t(1);       % Compute initial y'
```

Notice that the semi-colons suppress the output from showing up, but still execute the commands. Next, we would compute y(2), the value of y at time $t(2) = 1/10$:

```
j=2;
y(j)=y(j-1)+dy(j-1)*dt;
dy(j)=-y(j)+t(j);
```

and then change to $j = 3$ to compute the value of y at time $t_(3) = 2/10$ and so on. A "for" loop streamlines this process by cycling through each value of j

```
for j=2:n
    y(j)=y(j-1)+dy(j-1)*dt;
    dy(j)=-y(j)+t(j);
end

y
```

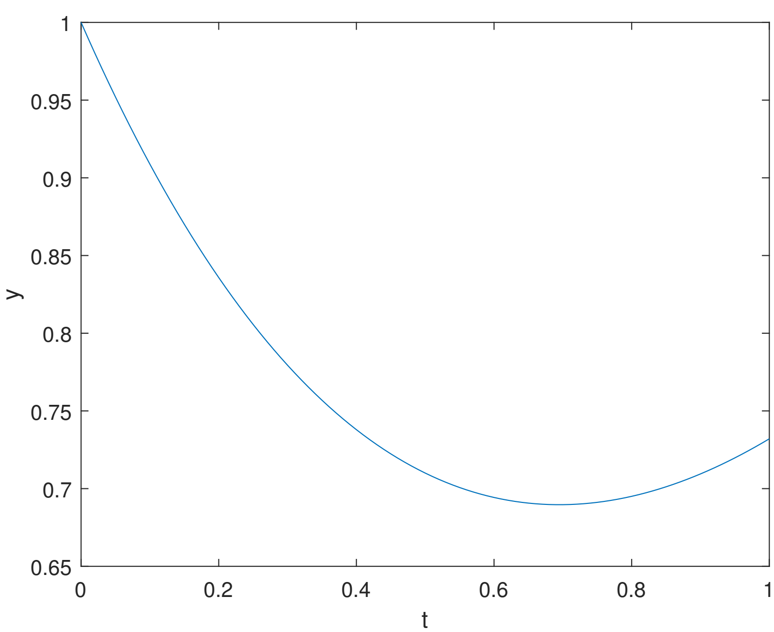
```
y =

  Columns 1 through 7

    1.0000    0.9000    0.8200    0.7580    0.7122    0.6810    0.6629

  Columns 8 through 11

    0.6566    0.6609    0.6748    0.6974
```

Voila, y is filled in. If we want to do this for 100 steps, no problem; just change the first line of the previous code.

```
dt=1/100;
t=0:dt:1;
n=length(t);
y=zeros(1,length(t));
dy=zeros(1,length(t));
y(1)=1;
dy(1)=-y(1)+t(1);
for j=2:n
    y(j)=y(j-1)+dy(j-1)*dt;
    dy(j)=-y(j)+t(j);
end
```

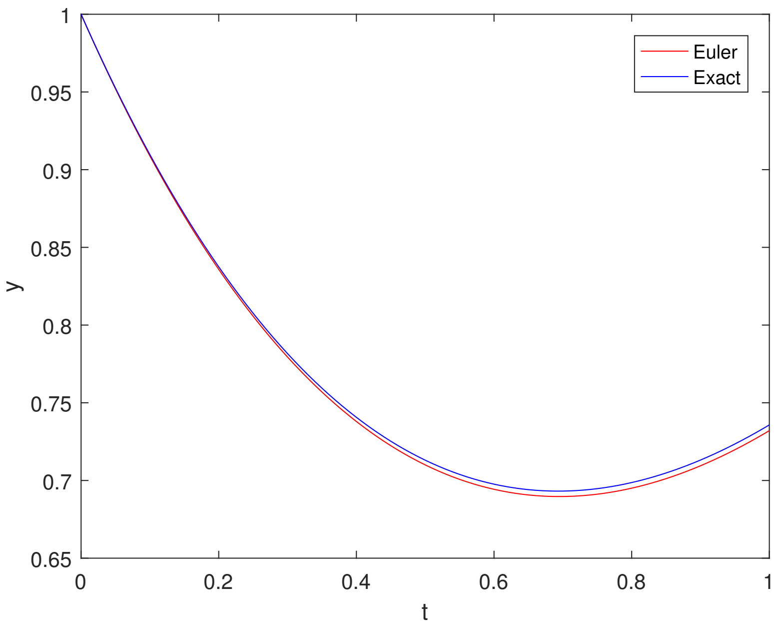
Now we plot the solution

```
plot(t,y)
xlabel('t')
ylabel('y')
```



In the next section, we will see that the exact solution to the given IVP is $y(t) = 2e^{-t} + t - 1$. We can compare this exact solution with the one obtained using Euler's and a step size of 1/100

```
plot(t,y,'r',t,2*exp(-t)+t-1,'b')
xlabel('t')
ylabel('y')
legend('Euler','Exact')
```



Pretty good match! At the end of the interval, the exact solution is $y(1) = 2e^{-1} + 1 - 1$ so the approximation error is

```
abs(y(end)-2*exp(-1))
```

```
ans =

    0.0037
```

Of course, if we want to obtain a more accurate estimate of $y(1)$, we could use an even smaller step size. $\Delta t = 1/1000$ yields

```
dt=1/1000;
t=0:dt:1;
n=length(t);
y=zeros(1,length(t));
dy=zeros(1,length(t));
y(1)=1;
dy(1)=-y(1)+t(1);
for j=2:n
```

```
    y(j)=y(j-1)+dy(j-1)*dt;
    dy(j)=-y(j)+t(j);
end

error=abs(y(end)-2*exp(-1))


error =

   3.6803e-04
```

In a course on numerical methods (Math 110), you would study how the error depends on the step size in more detail, but in brief, to get an additional decimal of accuracy, you typically need to decrease the step size by one order of magnitude in Euler's. For example, decreasing the step size from 1/100 to 1/1000 improved the absolute error from 0.0037 to 0.00037. Here, we will instead just end with a complete commented code from the example above. The lines with an asterisk are the ones you may need to change to tackle problems in the exercises below

```
% Code for using Euler's Method to approximate the solution to y'=t-y, y(0)=1
% using a step size of Delta t = 0.001

dt=0.001;              % *Define Step Size
t=0:dt:10;             % *Define interval of times
y=zeros(size(t));      %  Initialize variables for storing y
dy=zeros(size(t));

y(1)=1;                % *Enter initial condition for y
dy(1)=t(1)-y(1);       % *Compute initial y'=f(t_0,y_0);
                       % * in the example here f(t,y) = t-y

for j=2:length(t)
    y(j)=y(j-1)+dy(j-1)*dt;     %  Euler's Method Line to compute y(next)
    dy(j)=t(j)-y(j);            % *Compute y'(next) from the rate function
end

plot(t,y,'k')          % *Plot Solution
xlabel('t')
ylabel('y')
```

MATLAB's ode45

A course like Math 110 would also cover how to get "more bang for your buck" (that is, larger error reduction with fewer steps) compared to Euler's by more cleverly approximating y at each step. Lucky for us, Matlab has a whole suite of built-in functions that implement some of these techniques for the casual user. In this section, we shall illustrate the use of one such solver ode45 (pronounced "O-D-E-4-5" because it is based on a Runge-Kutta (4,5) formula). The "O" stands for "ordinary". All differential equations talked about in this course fall under the category of "ordinary differential equations," or ODEs for short, because they involve a single independent variable (usually t). In contrast, partial differential equations (PDEs) involve multiple independent variables and consequently, are much more complicated. But that is the topic of Math 157!

Before explaining how to use this function in Matlab, we first need a bit of terminology and a brief introduction to function handles. If you can write a first order differential equation in the form

$$y' = f(t, y)$$

for some function f, then f is called the *rate function* for the differential equation. For example, $y' = t - y$ has rate function $f(t, y) = t - y$. One can enter this (or any) rate function into Matlab using function handles as follows:

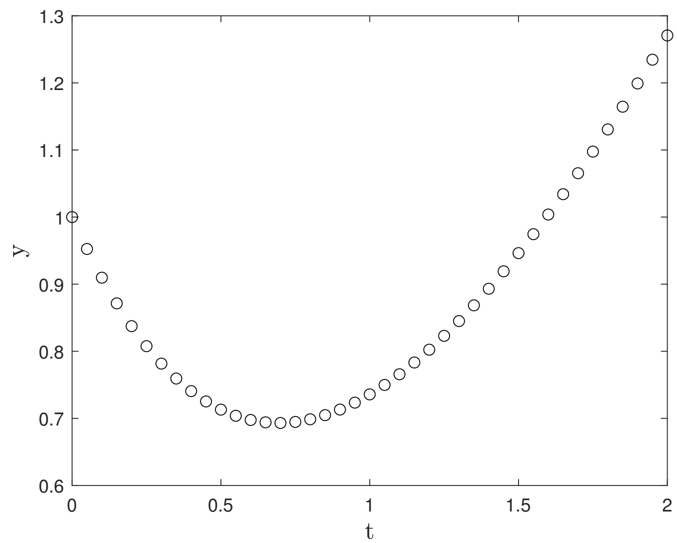
```
f=@(t,y)  t-y


f =

  function_handle with value:

    @(t,y)t-y
```

The notation used to define function handles is a bit funky at first glance, but it is easy to explain. The command "@(t,y)" tells Matlab to define a function with t and y as inputs. Whatever comes after the right parenthesis above then tells Matlab what to do with the inputs (in this case subtract them). We can then ask Matlab to evaluate our new function at various inputs, for example

```
f(3,1)


ans =

     2
```

The solver ode45 is a built-in Matlab function with three primary inputs. The first is a function handle where we input the rate function for our differential equation. The second is the span of t values over which we want to solve the differential equation and the third is the initial condition. For example, suppose we want to solve the IVP $y' = t - y$, $y(0) = 1$ over the interval $[0, 2]$ and then plot our solution using black circles. This can be accomplished with the following commands:

```
f=@(t,y) t-y;                  % Rate function
y0=1;                          % Initial Condition
[t,y]=ode45(f, [0,2],y0);      % Second input is time interval
plot(t,y,'ko')
xlabel('t')
ylabel('y')
```



Note that the recreated plot is similar to what we got using Euler's Method (as it should be), but for some examples, ode45 is much more accurate.

Exercises

(1) Katniss is watching a group of trackerjackers swarm the body of a fellow tributary. She estimates that $y(t)$ = number of trackerjackers on the body (***in hundreds***) after t seconds is changing according to the equation

$$\frac{dy}{dt} = \frac{y}{2}\left(1 - \frac{y}{10}\right).$$

Suppose that there are initially two hundred trackerjackers (so $y(0) = 2$).

(a) Create a single plot which compares the Euler's method approximations of $y(t)$ over the interval $0 \leq t \leq 5$ for three different step sizes: $\Delta t = 2.5, 0.25, 0.025$.

(b) What are the three different estimates of $y(5)$ from your plots in (a)? Why do you think the estimates are so different?

(c) Re-plot your approximate solution for step size $\Delta t = 0.025$ using the longer time interval $0 \leq t \leq 20$. What appears to be happening to the number of trackerjackers as $t \to \infty$?

(2) Consider the initial value problem

$$y' = 2 - 4y, \; y(0) = 1$$

Solve this IVP by hand (see Section 1.2, exercise 2b) and with ode45. Create a plot which compares the two solutions (they should match) using solid lines for your hand solution and circles for the ode45 solution to distinguish between them. This type of plot is what we will call “validating a solution” throughout the rest of this chapter.

(3) Recall the following IVP from Section 1.3 which arose in the context of a mixing problem:

$$\begin{aligned} y' &= -2y + 2(1 + \sin 2t) \\ y(0) &= 0 \end{aligned}$$

(i) Approximate the solution $y(t)$ to this IVP over the interval $0 \leq t \leq 20$ using either ode45 or Euler's with a step size of 0.001. Include a nicely labeled plot showing the result.

(ii) Discuss the behavior of the solution including an estimation of its midline, amplitude and period (review these terms if you forgot what they mean).

Section 1.4 - Linear Equations and the Principle of Superposition

Class objectives:

- Define, recognize, and classify linear first order equations as homogeneous or non-homogeneous equations.
- State and apply the principle of superposition to combine solutions to linear differential equations
- Determine solutions to linear first order equations with constant coefficients by the method of undetermined coefficients (MUC)

In the final exercise of Matlab Supplement 2, you were asked to solve the non-separable differential equation

$$y' = 2(1 + \sin(2t)) - 2y \tag{1}$$

using numerical methods. This equation was motivated by an application in which $y(t)$ was the amount of undesirable chemical in a reservoir being fed by a polluted pipeline after t days. Figure 1 presents the results of this analysis and compares $y(t)$ with the incoming chemical amount $g(t) = 2(1 + \sin 2t)$, suggesting several predictions:

- After a short initial period, the amount of chemical in the reservoir will oscillate between ≈ 0.3 and 1.7 g.
- Oscillations occur with period $\approx \pi$.
- Oscillations occur with a slight phase shift from the "input" function $g(t)$. In other words, there is a slight delay between when the amount of chemical in the inflowing water is at its peak and when the amount of chemical in the reservoir reaches its maximum value.

While numerical schemes like Euler's are powerful, it would be nice to confirm our predictions by computing an exact solution. There are, in fact, numerous ways to "skin the cat" so to speak and obtain the desired formula, but we will pursue here a technique called the Method of Undetermined Coefficients (or MUC for short as a mispelling of one of my favorite Pokémon characters). In addition to being a relatively straightforward method, it also has a natural analog to second order equations as we shall see in Chapter 2. First, we need to place Equation (1) in its proper context within the subject of ODEs.

Linear Equations

Recall from the last section that when a first order ODE is written in the form $y' = f(t, y)$, the function f is called the rate function for the ODE. A ***linear*** first order differential equation is one in which the rate function depends linearly on y and hence can be written in the form

$$y' + p(t)y = g(t) \tag{2}$$

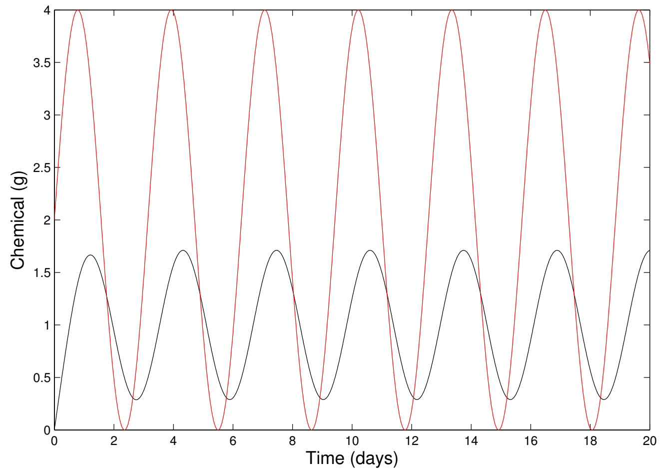


Figure 1: Plot of the solution $y(t)$ to equation (1) (in black) along with the input $g(t) = 2(1 + \sin 2t)$ (red).

where p, g are given functions of the independent variable t. If $g(t) = 0$, we further say that the first order equation is ***homogeneous***[1]. If $g(t) \neq 0$, then the system is ***non-homogeneous*** and we call g the ***forcing function***.

The reason for this latter name comes from thinking about first order equations as input-output problems: the forcing function $g(t)$ is received as an input and the solution $y(t)$ is sent as an output. For example, in the tank mixing problem from Equation (1), the forcing function $g(t) = 2(1 + \sin(2t))$ represents the rate at which chemical enters the tank. This "input" mixes with pure water already in the reservoir and then the tank outputs a different amount of chemical $y(t)$ on the other side. Most first order, non-homogeneous ODEs arise in similar scenarios.

The exponential growth and decay equation, $y' = ay$, which can be rewritten as $y' + p(t)y = 0$ with $p(t) = -a$, is the simplest example of a linear equation and it is homogenous (the righthand side $= 0$). In general, however, linear equations can be quite complicated, since the name "linear" comes from the linear dependence on y and NOT on t. For example, the differential equation $y' = 2(1 + \sin 2t) - 2y$ from the reservoir problem is linear (even though it has a $\sin 2t$ term) because it can be written in the form (2) with $p(t) = 2$ and $g(t) = 2 + 2\sin 2t$. Here are some other examples:

- The equation $yy' = t + 1$ is a nonlinear first order equation because it contains the product yy'.
- The equation $(t+1)^2y' + e^{-t}y = \sin t$ is a linear first order equation and nonhomogeneous because we can divide both sides by $(t+1)^2$ to obtain (2) with $p(t) = e^{-t}/(t+1)^2$ and $g(t) = \sin t/(t+1)^2$ (provided $t \neq -1$).
- The equation $y' + 7\sin(t + y) = 1$ is nonlinear because the solution y appears inside the sin term.

[1] As a side note, this is pronounced "ho-mo-ge-ne-ous" and is actually a different word than "homogenous" with no "e"

- The equation $y' = e^t y$ is linear and homogeneous because it can be written in the same form as Equation (2) with $p(t) = -e^t$ and $g(t) = 0$.

For most of the course, we shall only look at scenarios where $p(t) = a$ is a constant. In this case, we call Equation 2 a first order constant coefficient homogeneous or non-homogeneous equation depending on whether $g(t)$ is or is not 0. We will use the abbreviations FOCCHE and FOCCNE, respectively, because we are lazy.

FOCCHE are easy to solve since they can always be rewritten as $y' = -ay$ and then we can separate variables to obtain $y(t) = ce^{-at}$ as a general solution. But our tank equation (1) is a FOCCNE and we have already noted that SoV fails in this scenario. So it is FOCCNE towards which we shall dedicate the remainder of this section.

The Method of Undetermined Coefficients

A solution y to a first order IVP $y' = f(t, y), y(0) = y_0$ must satisfy two requirements: its derivative must be $f(t, y)$ and its value at 0 must be y_0. Sometimes it is much easier to first find a function y whose derivative is f and then adjust our solution so that $y(0) = y_0$. This is similar to the idea behind antiderivatives. If you want to find a function y such that $y' = t$, we know that $y(t) = t^2/2$ does the job, but if we want a function for which $y' = t$ and $y(0) = 1$, $t^2/2$ is close, but no cigar. Instead, we have to consider all function of the form $t^2/2 + C$ and then PSI $y(0) = 1$ to find the appropriate value of C (in this case $C = 1$).

We shall apply a similar technique for solving non-homogeneous linear equations: we will first use good-old fashioned "guess-and-check" to find a function y for which $y' + ay = g(t)$ and then adjust it to satisfy our initial conditions. To illustrate, let's start by looking at the FOCCNE from our Matlab supplement $y' = -y + t$ which in standard from reads

$$y' + y = t. \tag{3}$$

We shall define a *particular solution* to a first order linear ODE as any function y which when plugged into the lefthand side of Equation (2) yields the right. We shall often use the subscript p to denote particular solutions. So in our present example (3), a particular solution is a function y_p such that $y_p' + y_p = t$.

To "guess" a particular solution, consider what the equation is saying about y_p. It must be a function that when differentiated (wrt t) and added to itself yields back t. Since derivatives of linear functions are constants, we might therefore guess that $y_p(t) = At + B$ for some constants A, B. To find appropriate values for A, B, we plug our "trial solution" into (3):

$$(At + B)' + (At + B) = t.$$

This simplifies to

$$(A + B) + At = 0 + 1t.$$

From here we can "match coefficients" of like terms to obtain a system of two equations

$$\begin{aligned} A + B &= 0 \\ A &= 1 \end{aligned}$$

which has solution $A = 1, B = -1$. Therefore, the function $y_p(t) = t - 1$ gives a particular solution to our non-homogeneous DE (3).

But in the problem from Matlab Supplement 2, we were also given the initial condition $y(0) = 1$ which our particular solution does not satisfy ($y_p(0) = -1$). So we need to add a " plus C" to our $y_p(t)$ which adjusts for this fact. However, if we simply add 2 to our $y_p(t)$ the resulting function $y_p(t) = t + 1$ is no longer a solution to Equation (3) and we are back where we started. So we need to be a little bit more clever in our choice.

The key to solving this mystery is to look at the complementary homogeneous problem $y' + y = 0$. We know that this equation has general solution $y_c(t) = Ce^{-t}$ (the subscript c is for "complementary"). What happens if we add THIS to our particular solution? Plugging $y(t) = y_p(t) + y_c(t) = t - 1 + Ce^{-t}$ into $y' + y$ yields

$$1 - Ce^{-t} + t - 1 + Ce^{-t} = t.$$

Sweet! It solve the DE. Now we PSI with our initial conditions $y(0) = 1$ to solve for C

$$C - 1 = 0 \Longrightarrow C = 1$$

and we have a solution $y(t) = t - 1 + e^{-t}$ to our given IVP.

This process of finding solutions to FOCCNE IVPs is called the Method of Undetermined Coefficients (MUC). To summarize, it involves three steps:

(1) "Guess-and-check" a particular solution $y_p(t)$ to the non-homogeneous equation of interest.

(2) Determine the general solution $y_c(t) = Ce^{-at}$ to the complementary equation $y' + ay = 0$. The general solution is then $y(t) = y_c(t) + y_p(t)$

(3) PSI with the initial conditions to find C.

Of course the hardest part is making an educated guess at an appropriate trial solution to plug in for Part 1 which leads us to the next section.

Trial Solutions

MUC only works if the forcing function $g(t)$ is simple enough for us to guess an appropriate trial solution. For example, if $g(t)$ is a linear function, we saw in the last section that guessing $y_p(t)$ to be a generic linear function $at + b$ worked. In general, there are really three situations in which such educated guessing is possible

- when $g(t)$ is a polynomial.
- when $g(t)$ is a exponential function
- when $g(t)$ is a combination of sin or cos functions.

In all three cases, the idea is to choose $y_p(t)$ as a function which includes *all terms in g, plus their derivatives, with undetermined coefficients.* The exercises below give some additional practice and exploration involving the first two cases, but we would like to demonstrate an example of the third by finally returning to our tank example from Equation (1).

The forcing term in this equation is $g(t) = 2 + 2\sin 2t$. The first term is a constant so our trial solution $y_p(t)$ should have an A in it (the derivative of a constant is 0 so we don't need to worry about that). The second term is $\sin 2t$ whose derivative is $2\cos 2t$ so we need both a $B\sin 2t$ and $C\cos 2t$ term[2]. Putting this all together yields the trial solution

$$y_p(t) = A + B\sin 2t + C\cos 2t.$$

Now we checking our candidate solution by plugging it into $y' + 2y$ (the part of Equation (1) without g) and equating the result with g as required by the equation:

$$y_p' + 2y_p = 2A + (2B - 2C)\sin 2t + (2B + 2C)\cos 2t = 2 + 2\sin 2t$$

Equating coefficients of like terms, we can see that we need

$$\begin{aligned} 2A &= 2 \\ 2B - 2C &= 2 \\ 2B + 2C &= 0. \end{aligned}$$

We can solve these equations for the unknowns to obtain $A = 1, B = 1/2$, and $C = -1/2$. Therefore, a particular solution is

$$y_p(t) = 1 + (1/2)\sin 2t - (1/2)\cos 2t.$$

To complete MUC, we note that the general solution to the complementary equation $y' = -2y$ is $y_c(t) = Ce^{-2t}$ so the general solution is

$$y(t) = 1 + (1/2)\sin 2t - (1/2)\cos 2t + Ce^{-2t}$$

Starting off with pure water tells us that $y(0) = 0$ so we substitute in $t = 0$ to obtain

$$0 = C + (1/2)(0) - (1/2)(1) + 1 \Longrightarrow C = -1/2.$$

Let's compare this exact solution with the approximate solution shown in Figure 1. First, notice that $(1/2)e^{-2t} \to 0$ as $t \to \infty$ and so after a short initial period, $y(t) \approx 1 + (1/2)\sin 2t - (1/2)\cos 2t$. For this reason, we call $(-1/2)e^{-2t}$ the ***transient*** part of the solution. It is the part of the solution coming from our initial conditions and it literally gets washed out over time in this example. We call $(1/2)\sin 2t - (1/2)\cos 2t + 1$ the ***steady-state*** part of the solution because this is the part that dominates in the long run. After the transient part gets washed away, the steady state is indeed a function of period π (as we predicted). This will be further investigated in the next section where we discuss the phase-amplitude form of solutions to RC circuits with sinusoidal input.

[2] We don't need a "2" in front of $\cos 2t$ because the C is already an arbitrary constant.

The Principle of Superposition

To conclude this section, we return to the rather mysterious step of adding y_c to y_p. It worked in one example, but that is hardly a mathematical proof. True, this is a course in applied differential equations, but it is still a math course and as such, we want to give some attention to deeper principles at work beneath the surface. Besides, we shall see the same "trick" come into play, albeit in different forms, throughout the text and as such, it seem worth attaching a name to this phenomenon. And so we shall.

The reason why linear equations are nice to work with is because we can easily form new solutions from old ones via the process of *superposition*. To describe this principle, suppose that $y_1(t)$ is a solution to a linear differential equation of the form

$$\frac{dy_1}{dt} = -p(t)\,y_1 + g_1(t) \tag{4}$$

and y_2 is a solution to a different linear differential equation

$$\frac{dy_2}{dt} = -p(t)\,y_2 + g_2(t). \tag{5}$$

Notice both equations have the same function p, but potentially different forcing functions g. Define a new function $y(t) = y_1(t) + y_2(t)$. If we differentiate y, we get

$$\begin{aligned} \frac{d}{dt}(y_1 + y_2) &= \frac{dy_1}{dt} + \frac{dy_2}{dt} \\ &= (-p(t)\,y_1 + g_1(t)) + (-p(t)\,y_2 + g_2(t)) \\ &= -p(t)[y_1(t) + y_2(t)] + [g_1(t) + g_2(t)] \end{aligned}$$

where the first equality follows from the linearity of differentiation, the second from the fact that y_1 and y_2 are solutions to Equations (4) and (5), respectively, and the final from rearranging. If we compare the first and last expressions above, we can see that $y = y_1 + y_2$ is a solution to the equation

$$\frac{dy}{dt} = -p\,y + g$$

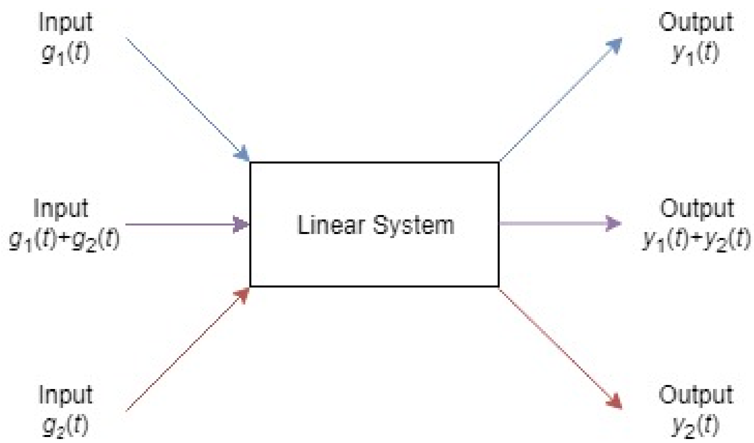
where $g(t) = g_1(t) + g_2(t)$. Put into words, the solution to a linear equation whose forcing term is the superposition (or sum) of two separate functions is the sum of the solutions that you would obtain by solving the two linear equations obtained from separating out the forcing functions. In fact, we could easily extend the argument above to any finite sum of forcing functions $g_1, g_2, ...g_N$.

This Principle of Superposition is typically stated in the more general sense of inputs and outputs to linear systems. To reconcile this with our calculation above, again think of a forcing function as a stimuli or input and the solution to a linear differential equations as the corresponding output or response[3]. We can then state the principle of superposition as follows

> The net output caused by two or more inputs in a linear system is the sum of the outputs that would have been caused by each input individually.

[3]As in the mixing tank problem from the last section.

This process is also illustrated in the figure below.



We will use the Principle of Superposition (PoS) in numerous ways throughout the semester, but all of them will take one of the following forms:

PoS1 Suppose that we find two different solutions $y_1 + y_2$ to a homogeneous linear equation $y' + p(t)y = 0$. If we apply the PoS with $g_1 = g_2 = 0$ then since $0 = 0 + 0$, we can see that $y = y_1 + y_2$ solves the same homogeneous equation. In other words, **the sum of two or more solutions to a homogeneous linear equation is also a solution**.

PoS2 Suppose we are given a nonhomogeneous equation $y' + p(t)y = g(t)$ and find a particular solution $y_p(t)$ along with a solution $y_c(t)$ to the complementary homogeneous equation $y_c' + p(t)y_c = 0$ (i.e. same left hand side, but no forcing). Then applying the PoS with $g_1 = g$, $g_2 = 0$, we can see that $y(t) = y_p(t) + y_c(t)$ is also a solution to the nonhomogeneous equation $y' + p(t)y = g(t)$.

PoS3 Suppose we are trying to solve $y' + p(t)y = g(t)$ where g is a sum of two simpler forcing function g_1, g_2. Then we can separately solve the simpler equations $y_1' + p(t)y_1 = g_1(t), y_2' + p(t)y_2 = g_2(t)$ and superimpose the solutions to solve the original equation of interest.

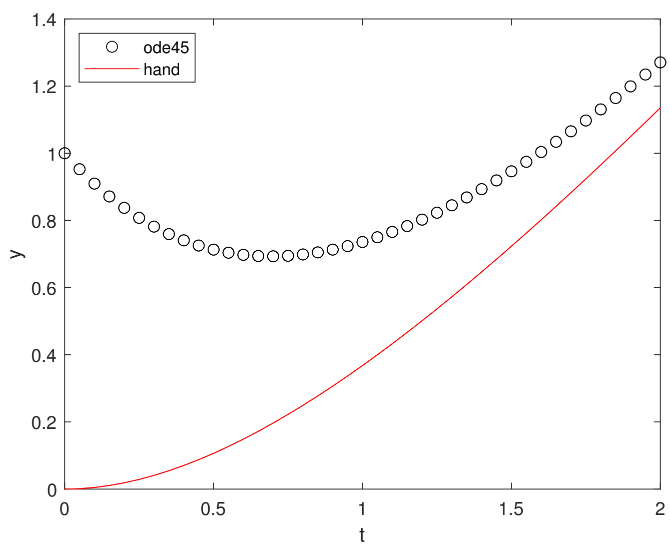
The first will come into play in Chapters 2 and 3 (see also Exercise 1 below) while the usefulness of the third is demonstrated in the exercises below. PoS2 justifies us adding $y_p + y_c$ in MUC[4] and concludes our discussion on the topic.

A Quick Note on Validating Solutions

MUC, like SoV, is an analytic method for solving ODEs: you manipulate symbols and obtain an exact formula for the solution. It is easy to make and miss algebraic mistakes in such problems and it is important to have a computer-based method for validating your solutions once you have them. For example, suppose in solving the IVP $y' = t - y$, $y(0) = 1$ over the interval $[0, 2]$ a series of unfortunate events leads us to obtain $y(t) = e^{-t} + t - 1$ as our final answer (which is wrong). We will catch our mistake if we plot our hand solution against the output of ode45

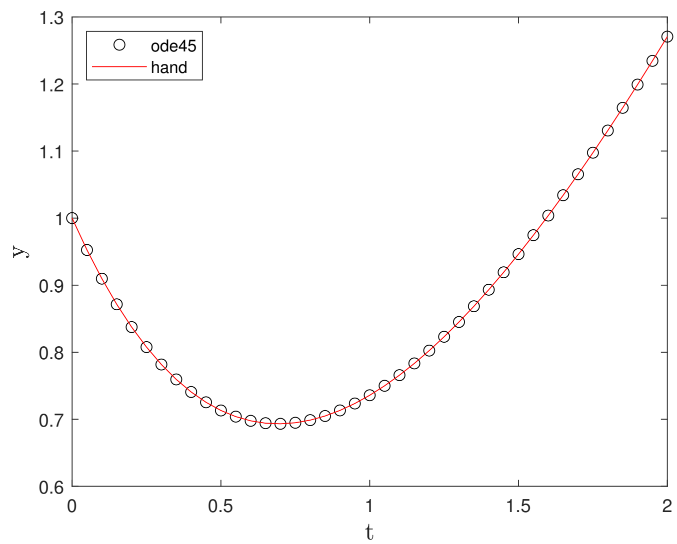
[4]A colleague likes to refer to this one as "The Fundamental Theorem of Math 57". They have a good point: it is pretty fundamental to the course.

```
% Numeric solution
f=@(t,y) t-y;               % Rate function
y0=1;                       % Initial Condition
[t,y]=ode45(f, [0,2],y0);   % Second input is time interval
plot(t,y,'ko')
hold on

% Analytic Solution
yhandWrong=t-1+2*exp(-t);
plot(t,yhandWrong,'r')
xlabel('t')
ylabel('y')
legend('ode45','hand','Location','northwest')
```



The two do not match up so we go back and recheck our algebra arriving at the correct solution $y(t) = 2e^t + t - 1$. We then change the second part of our code to the following and re-plot to confirm:

```
yRight=2*exp(-t)+t-1;
plot(t,yRight,'r')
xlabel('t')
ylabel('y')
legend('ode45','Hand','Location','northwest')
```



Yay, its a match! Now you can submit your assignment with confidence.

Exercises

In Exercise 2 below (as in other MUC problems), you may have to experiment around with a couple of different trial solutions $y_p(t)$ before you get one that works. But as Steven Tyler used to say in one of my favorite childhood arcade games Revolution X to get you to pump another quarter in the machine, "Don't Give up". Go back, modify, and try again!

(1) A ***linear combination*** of two functions y_1, y_2 is any function of the form $y(t) = c_1 y_1(t) + c_2 y_2(t)$ where c_1, c_2 are constants. Show that any linear combination of solutions to a first order linear homogeneous equation is also a solution to the same equation. **Hint**: Follow the same argument laid out when deriving the Principle of Superposition.

(2) Solve the following IVPs and validate your solutions using ode45 (this means include a plot which shows that the solution you got by hand matches the one spit out by ode45; see examples at the end of the section). Briefly describe the asymptotic behavior of the solution as $t \to \infty$.

(a) $y' + 2y = t + t^2,\ y(0) = 1$

(b) $2y' + 4y = 6e^{-3t}, y(0) = 0.$

(c) $y' = 5y + 5 - 4e^t,\ y(0) = -1.$

(d) $y' - 2y = 5\sin t,\ y(0) = 1.$

(e) $y' = -y + 5\cos(2t),\ y(0) = 1.$

(3) For each of the following forcing functions, identify an appropriate trial solution to use in solving the DE $y' + ay = g(t)$. You do not need to actually solve for the undetermined. coefficients.

(a) $g(t) = 2\sin 4t + 4\cos 2t$

(b) $g(t) = 1 + t^2 + 3\sin 5t$

(c) $g(t) = e^{-10t} + e^{2t} + t + 2\cos 3t$

(4) Find a particular solution to the differential equation $y' + 3y = 3 + 2e^{-t}$ and determine what will happen to this solution as $t \to \infty$.

(5) Given that $y_1(t) = te^{-t}$ is a solution to $y_1' + y_1 = e^{-t}$, determine the general solution to the differential equation $y' + y = e^{-t} + t$. Hint: We already solved $y' + y = t$ so use the general Principle of Superposition.

(6) Given that $y_p(t) = (t+1)^2$ is a particular solution to the differential equation

$$(t+1)y' - y = (t+1)^2,$$

determine the solution with initial condition $y(0) = 0$.

(7) Find a general formula for the particular solution to the equation

$$y' + y = e^{rt}$$

in terms of r in the case when $r \neq -1$. Discuss how the solution depends on the parameter r.

(8) If $r = -1$ in the last exercise, find values of A, B for which $y_p(t) = (At + B)e^{-t}$ is a particular solution to the given DE. **Remark:** This demonstrates a special case when the trial solution you want to try doesn't work because it is already a solution to the complementary equation, but one can still find a particular solution after modifying the form of the trial solution.

(9) Find values of a, r such that the function

$$y_p(t) = at^r$$

is a particular solution to the differential equation

$$ty' + y = t^2.$$

Remark: This exercise shows an example of how "guess and check" can sometimes work to find particular solutions of non-constant coefficient, linear equations.

Section 1.5 : Applications to Circuits

Class objectives:

- Set up and solve differential equations for the response of RC circuits to both constant and sinusoidal inputs.
- Write steady state responses in phase-amplitude form to identify the phase, amplitude, and period of the response to sinusoidal inputs.

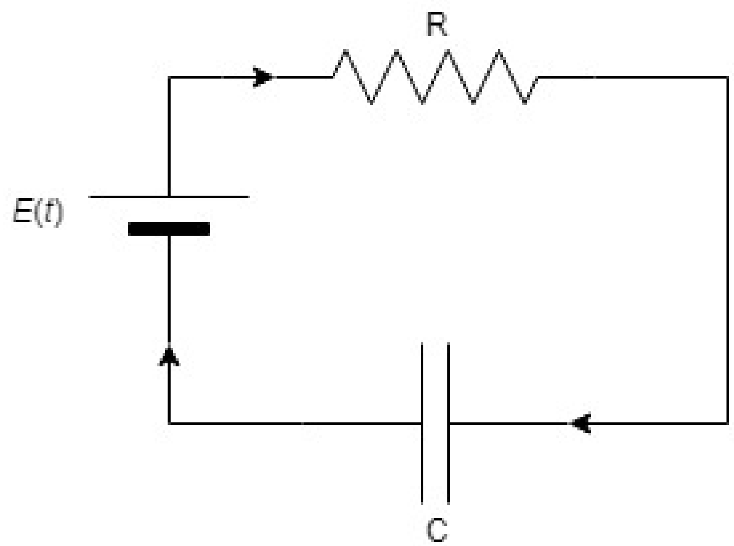
RC Circuit Description

Another application in which students frequently encounter differential equations is in the design and analysis of electrical circuits. In this section, we will consider a simple electrical circuit with the following components (see the diagram below for a schematic representation):

(i) An electromotive force ($E(t)$).

(ii) A resistor (R).

(iii) A capacitor (C).



We call this an *RC circuit*. Although our goal will mainly be the mathematical analysis of such systems, we'll give a one paragraph introduction to how this circuit works, recognizing that it is probably too few words to do any justice to the topic and too many words for an uninterested reader! Electrons naturally want to move from areas of high concentration (negative charge) to areas of low concentration (positive charge). An electromotive force (emf) works against this by pushing electrons through the circuit in the opposite direction and the force at which it does so is measured in Volts (V). Both the resistor and the capacitor oppose this force, resulting in a *voltage drop* across each component. However, the mechanisms of opposition are quite different. The resistor opposes the flow of electrical charge created by the emf at a rate proportional to its current (how fast the charge is moving) while the capacitor stores up charge. This is often explained by an analogy with water flowing through a system of pipes: voltage is like water pressure, the emf is like a pump forcing water into the system, the current (obviously) is the rate at which the water is flowing, the resistor is like a valve or narrowing in the pipe which decreases pressure on the opposite side, and the capacitor is like a big reservoir which stores up water. Now onto the math!

The Response of an RC Circuit

Our goal will be to set up and study a differential equation for $q(t)$, the difference in electrical charge (measured in Coloumbs) on opposing ends of the capacitor. Let $i(t)$ denote the corresponding current (measured in Amperes) feeding the capacitor. Kirchhoff's voltage law says that the voltage drop around a closed circuit is zero. With an emf providing a voltage of $E(t)$ V at time t, this yields the equation:

$$V_R + V_C = E(t)$$

where V_R and V_C are the voltage drops across the resistor, and capacitor, respectively. To express V_R and V_C in terms of the unknown quantity of interest $q(t)$, we use the following additional electrical laws:

(Ohm's Law) The voltage across the resistor is proportional to the current: $V_R = iR$ for some positive constant R called the resistance of the resistor (units of Ohms)

(Capacitor-Voltage Law) The voltage across the capacitor is proportional to its charge: $V_C = q/C$ for some positive constant C called the capacitance (units of Farads)

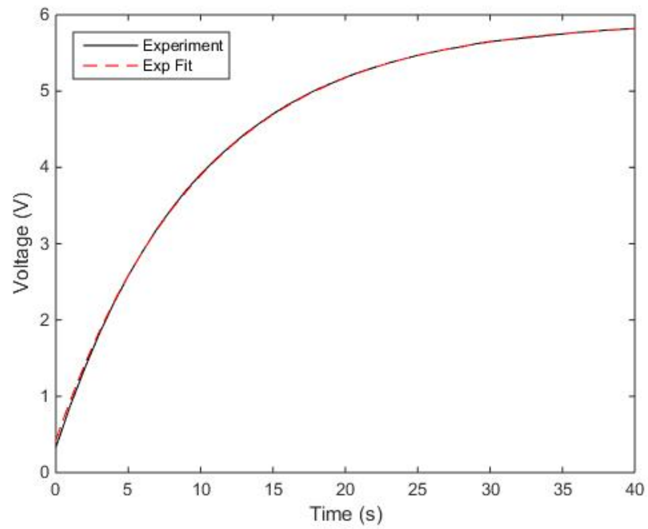
Since the current is the rate of change of charge ($i = q'$), we obtain the differential equation.

$$Rq' + \frac{1}{C}q = E(t)$$

This is similar to the non-homogeneous linear equations encountered last class in our study of mixing tanks. Note also that if we instead want an equation for the *voltage* $v(t)$ across the capacitor after t seconds, we can plug $v = q/C$ into the above DE for q to obtain

$$RCv' + v = E(t) \tag{1}$$

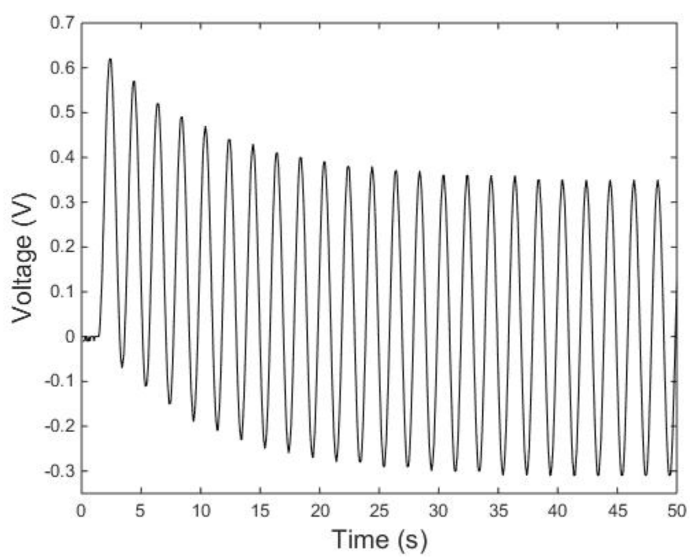
We leave it as an exercise to show that if we supply a constant or Direct Current (DC) input $E(t) = \varepsilon > 0$, then $v(t) = \varepsilon(1 + e^{-t/(RC)})$. The figure below compares the results of an experiment in which an RC circuit was driven with a DC input of $6V$. The exponential fit was obtained using Matlab's polyfit function and we see a good agreement between the two curves.



When $E(t) = \varepsilon \sin(\omega t)$, the emf is producing an alternating current (AC). From our work in the last section, we know that the general solution to Equation (1) will be given by

$$v(t) = A \sin \omega t + B \cos \omega t + c_1 e^{-t/(RC)} \tag{2}$$

for some constants c_1, A, B. The second part of the solution is transient, quickly decaying towards 0 as $t \to \infty$. The first part of the solution represents the *steady state* response of the circuit. A good schematic for thinking about this problem is as an input-output scenario: the RC circuit receives an input $E(t) = \varepsilon \sin(\omega t)$ and produces a steady state output $v_s(t) = A \sin \omega t + B \cos \omega t$. An example, in which an RC circuit was driven by an alternating current with an amplitude of 10 V and a frequency of 0.5 Hz (forcing function $= 10 \sin(\pi t)$) is shown below.



Typical applications involve determining how RC circuits will respond to different inputs. Take a moment to jot down a few observations based on the figure above. Ask yourself:

- How does the period/frequency of the output compare with the period/frequency of the input?
- How do the amplitudes compare?
- What is the phase shift of the output?

Once you have ruminated on these a while, proceed to the next section for answers.

Phase-Amplitude Form of Solution

Before we begin, it is worth recalling once again a few facts about the function $f(t) = \alpha \sin(\omega t + \phi)$:

- α is called the *amplitude* and represents the maximum displacement of f from 0.
- ω is called the *angular frequency* and is measured in radians/sec. The *period* (length of a single oscillation) is $2\pi/\omega$ while the *frequency* (how many oscillations occur per second) is $\omega/2\pi$. Frequency is commonly measured in hertz (1 hz = 1 cycle per second).
- ϕ is called the *phase* and indicates a time lag in the amount of time it takes to reach the first peak. In particular, the time-lag in the graph of f relative to a normal sine graph is $-\phi/\omega$ which can be found by setting $\omega t + \phi = 0$ and solving for t.

Analyzing the steady-state solution in equation 2 (the "output") involves determining how its amplitude, frequency/period, and phase shift relate to the corresponding properties of the forcing function (the input). The problem is that the steady-state is currently written as a sum of a sin and cos function with the same frequency and hence, directly reading off its oscillatory properties is more difficult. Fortunately, in trig, you probably learned (and forgot) that a combination of a sin and cos function with the same frequency can be rewritten as a single sin function of the form

$$\alpha \sin(\omega t + \phi)$$

using the sum of angles formula:

$$\sin(a + b) = \cos b \sin a + \sin b \cos a.$$

How? We work backwards and apply the sum of angles formula with $a = \omega t$ and $b = \phi$ to yield

$$\alpha \sin(\omega t + \phi) = \alpha \cos\phi \sin(\omega t) + \alpha \sin\phi \cos(\omega t). \tag{3}$$

Low and behold, the right side above is of the form $A\sin(\omega t) + B\cos(\omega t)$ with $A = \alpha\cos\phi, B = \alpha\sin\phi$. Since

$$A^2 + B^2 = \alpha^2$$

we can find the amplitude α by the formula

$$\alpha = \sqrt{A^2 + B^2}. \tag{4}$$

The phase can be found by the solution ϕ to

$$\tan\phi = B/A \tag{5}$$

and the angular frequency is of course just ω.

To see how this works in an example, suppose that we obtain a steady state solution

$$v_s(t) = (1/2)\sin(2\pi t) + \cos(2\pi t).$$

We want to write this in the alternative form

$$\alpha \sin(\omega t + \phi).$$

The angular frequency $\omega = 2\pi$, the amplitude is given by equation (4) as

$$\alpha = \sqrt{(1/2)^2 + 1} = \sqrt{5}/2,$$

and the phase shift is given by equation (5) as

$$\phi = \tan^{-1}\left(\frac{1}{1/2}\right) = \tan^{-1}(2).$$

Now to tie all this back in with AC circuits, suppose that we have an RC circuit with $R = 17k\Omega = 170,000\Omega$ and $C = 2.2\mu F = 2.2 \times 10^{-6}F$, driven by an AC force of $2\sin(2\pi t)$ (2 V with a frequency of 1 hz). MUC yields the steady state solution

$$v_s(t) = \frac{2}{1 + (2\pi RC)^2}\sin(2\pi t) - \frac{4\pi RC}{1 + (2\pi RC)^2}\cos(2\pi t).$$

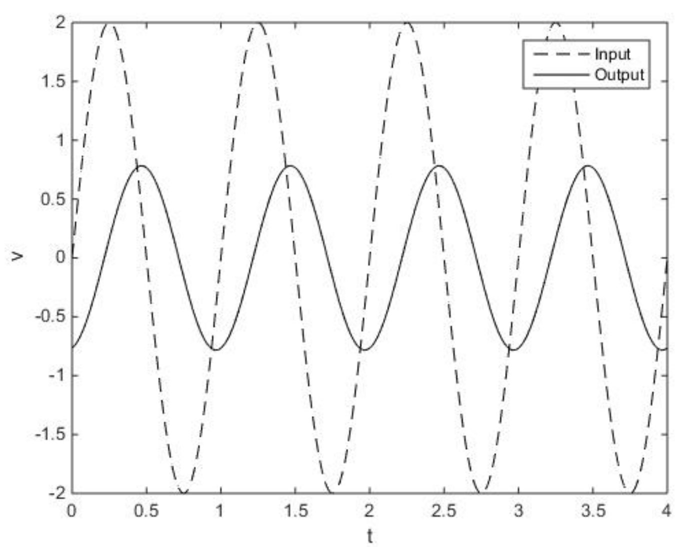
Therefore, the frequency of the steady state output is also 1 hz, the amplitude is

$$\sqrt{\left(\frac{2}{1 + (2\pi RC)^2}\right)^2 + \left(-\frac{4\pi RC}{1 + (2\pi RC)^2}\right)^2} \approx 0.3916$$

and the phase-shift is

$$\phi = \tan^{-1}(-4\pi RC/2) = -1.3611.$$

The steady state solution and input are shown in the figure below.



Frequency Response Curves

It is not hard to show by repeating the procedures outlined in the previous section that the steady state solution for a generic RC circuit driven by AC input $E(t) = \varepsilon \sin(\omega t)$ is

$$v_s(t) = \frac{\varepsilon}{\sqrt{1 + (\omega RC)^2}} \sin(\omega t + \phi) \tag{6}$$

where $\phi = tan^{-1}(-\omega RC)$. A tool commonly employed by engineers to visualize this information is a *frequency response* curve. A frequency response curve shows how a system responds to inputs of different frequencies by plotting frequency (in either radians or hertz) on the x-axis and amplitude gain or phase shift on the y-axis. In practical applications, amplitude gain is typically measured in decibel) as a ratio of output to input power on a logarithmic scale, however, to avoid a digression into things like Joule's Laws, we will simply measure amplitude gain in this text as the ratio of output to input amplitude.

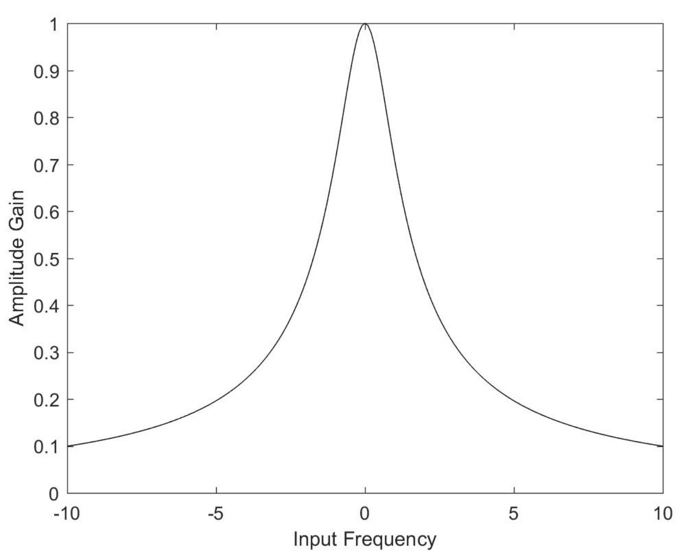
In the context of an RC circuit, the input refers to the electromotive force $E(t)$ and the output refer to the steady state response $v_s(t)$. If the input is sinusoidal $E(t) = \varepsilon \sin(\omega t)$ then Equation (6) implies that the steady state response gain at angular frequency ω is

$$f(\omega) = \frac{\text{Amplitude of Output}}{\text{Amplitude of Input}} = \frac{\frac{\varepsilon}{\sqrt{1+(\omega RC)^2}}}{\varepsilon} = \frac{1}{\sqrt{1 + (\omega RC)^2}}.$$

We write $f(\omega)$ because we want to think of this as a function of ω. In fact, we call f the frequency-response function for the system (or FRF for short). As an example, suppose that we have an RC circuit with $RC = 1$. Then the FRF is

$$f(\omega) = \frac{1}{\sqrt{1 + \omega^2}}.$$

A segment of this function is plotted in the figure below and is frankly quite uninteresting; the gain is a strictly decreasing function of input frequency with the largest gain achieved at 0 (constant input). For this reason, RC circuits are often used as "low-pass" filters (low frequency signals pass, but high frequency signals get filtered out). We will see more exciting scenarios when we return to this concept in Module 2.



Exercises

(1) Determine the solution to Equation (1) in the case where the emf produces a Direct Current (DC) of $E(t) = \varepsilon$ Volts and $v(0) = 0$. What happens to the voltage as $t \to \infty$?

(2) Write the following functions in phase-amplitude form and as a result, compute the phase, amplitude, period, and frequency of the solution.

(a) $y(t) = 3\cos 2t + 4\sin 2t$

(b) $y(t) = -\cos t + \sqrt{3}\sin t$

(c) $y(t) = 4\cos 3t - 2\sin 3t$

(d) $y(t) = -2\cos \pi t - 3\sin \pi t$.

(3) A closed circuit consists of a resistor with a resistance of $R = 10^4$ Ohms, a capacitor with a capacitance of $C = 10^{-3}$ Farads, and a driving force $E(t) = 10\sin(\pi t)$ volts. Suppose that the initial voltage drop across the capacitor is 0.

(a) Determine a formula for the voltage across the capacitor after t seconds.

(b) Use Matlab to make a plot of your solution from part (a) over 20 seconds. Compare the period, amplitude, and phase of the solution with the period, amplitude, and phase of the driving force $E(t)$.

(4) A driving force of $E(t) = \cos(4t)$ volts is applied to an RC circuit with resistance $R = 10^6$ Ohms, capacitance $C = (1/2) * 10^{-6}$ Farads, and zero initial voltage across the capacitor. Determine a formula for the steady state response and calculate the amplitude reduction from the input.

(5) Consider an RC circuit with $R = 10^4$ Ohms, $C = 10^{-4}$ Farads, and $E(t) = \varepsilon \cos t$ Volts for some positive constant ε.

(a) Determine a formula for $v(t)$ assuming that $v(0) = 0$ (your answer will depend on ε). Include a Matlab validation of your solution for the case where $\varepsilon = 1$ and identify the transient/steady state parts of your solution.

(b) Calculate the period, amplitude, and phase shift of the steady state part of your solution in terms of ε. Compare with the amplitude and period of the driving force $g(t)$ and create a Matlab plot which illustrates the comparison for several different values of ε.

(6) Suppose that the voltage $v(t)$ across the capacitor, after t seconds, in a series RC circuit connected to an AC power source is

$$(RC)v' + v = E(t)$$

with $R = 10^4$ Ohms, $C = 2 \times 10^{-4}$ Farads, and $E(t) = 4\cos(2t)$ Volts.

(a) Determine a formula for $v(t)$ assuming that $v(0) = v_0 > 0$. Your answer will depend on v_0.

(b) Identify the steady state part of your solution. Does it depend on v_0?

(c) Calculate the period and amplitude of the steady state part of your solution.

Section 1.6 : Variation of Parameters

Class objectives:

- Determine general and particular solutions to first order linear equations using Variation of Parameters
- Compare the efficiency of variation of parameters with previous methods (method of undetermined coefficients, separation of variables, Euler's)

The Method of Undetermined Coefficients is sufficient for solving all first order linear equations of the form

$$y' + ay = g(t)$$

for some constant a as long as g is a combination of polynomials, sines, cosines, or exponentials. While this covers many common applications, our discussion of first order linear systems would not be complete without coverage of *variation of parameters*[1], a technique that is capable of solving a larger class of first order linear equation (including those with non-constant coefficients) and also has extensions to higher order equations.

The Basic Idea

Like undetermined coefficients, variation of parameters (VoP) is a method for solving non-homogeneous linear equations using the solution to a corresponding homogeneous equation as a starting point. As an example of how the method works, suppose we want to find the general solution to the equation $y' + 2y = e^{-t}$. The complementary equation is $y_c' + 2y_c = 0$ which has the general solution $y_c(t) = c_1 e^{-2t}$. Instead of "guessing" a trial solution, VoP says to propose the particular solution

$$y_p(t) = u(t)e^{-2t}. \tag{1}$$

In other words, we replace the constant c_1 in the complementary solution with a function of t. (Since c_1 is a parameter, we say that we are "letting the parameter vary", hence the name for the technique.) We then plug this back into the equation $y' + 2y = e^{-t}$, apply the product rule, simplify, and integrate to obtain a formula for u:

$$\begin{aligned} \left(ue^{-2t}\right)' + 2\left(ue^{-2t}\right) &= e^{-t} \\ u'e^{-2t} - 2e^{-2t}u + 2ue^{-2t} &= e^{-t} \\ u'e^{-2t} &= e^{-t} \\ u' &= e^{t} \\ u(t) &= e^{t}. \end{aligned}$$

Plugging this back into proposed solution in (1) yields the particular solution

$$y_p(t) = e^t e^{-2t} = e^{-t}$$

and the general solution is then

$$y(t) = y_c(t) + y_p(t) = c_1 e^{-2t} + e^{-t}$$

[1] Many books instead cover the closely related Method of Integrating Factors which is essentially equivalent as will be shown later in this section.

Why VoP Works

Any homogeneous linear equation $y' + p(t)y = 0$ can be solved using separation of variables:

$$\frac{dy}{dt} = -p(t)y \Longrightarrow \frac{1}{y}\,dy = -p(t)\,dt$$

resulting in the general solution

$$y_c(t) = c_1 e^{-\int p(t)\,dt}. \tag{2}$$

Suppose we are looking for a particular solution to $y' + p(t)y = g(t)$ of the form

$$y_p(t) = u(t)I(t)$$

where

$$I(t) = e^{-\int p(t)\,dt}$$

is a solution to the complementary equation $y_c' + p(t)y_c = 0$. Plugging our proposed solution y_p into the DE $y' + p(t)y = g(t)$ yields

$$\begin{aligned} (uI)' + p(uI) &= g \\ u'I + uI' + puI &= g \\ u'I + u(I' + pI) &= g. \end{aligned}$$

Looking at the last line, we note that since I is a solution to the complementary equation $y_c' + py_c = 0$, we have $I' + pI = 0$ so the second term on the left drops out and we obtain a new DE

$$u'I = g$$

Assuming that $I \neq 0$ (which only happens for the trivial IVP $y' = 0, y(0) = 0$), we can divide by I and integrate to obtain a formula for u

$$u(t) = \int_t I^{-1}(s)g(s)\,ds. \tag{3}$$

Therefore, a particular solution to $y' + p(t)y = g(t)$ is

$$y_p(t) = I(t)\left(\int_t I^{-1}(s)g(s)\,ds\right)$$

and the general solution is

$$y(t) = y_c(t) + y_p(t) = c_1 I(t) + I(t)\left(\int_t I^{-1}(s)g(s)\,ds\right) \tag{4}$$

This is a really nice, compact way of writing the general solution to a linear first order, nonhomogeneous DE, but trying to memorize and apply directly is not necessarily the best way to solve DEs using VoP. It is much safer to just go through the process which led us to this formula. However, if you insist on taking this path under the mountain, then note that you cannot just move the $I(t)$ inside the integral to cancel out the $I^{-1}(s)$. Instead, equation (4) says that you first multiply I^{-1} and g, then integrate the product, and then multiply by I.

Connections with Integrating Factors

Many books implicitly cover VoP for first order equations under the guise of the *Method of Integrating Factors*. The latter actually starts from a different perspective, but the end result is the same. The function $I(t)$ which appears in the previous section is called an *integrating factor*[2]. For more on this alternative, but ultimately equivalent approach, just google "integrating factors".

Chapter Summary

In this Chapter, we covered three techniques for obtaining exact solutions to first order equations

- Separation of Variables (SoV)
- Method of Undetermined Coefficients (MUC)
- Variation of Parameters (VoP)/Method of Integrating Factors

as well as one method for obtaining numerical approximations (Euler's Method). Deciding on which to use in a given application is not always easy, but my advice is as follows:

- If the equation is separable (i.e. can be written as $y' = p(t)m(y)$), use SoV. This includes all linear homogeneous equations.
- If the equation is a FOCCNE with sinusoidal or polynomial forcing, use MUC.
- If the equation is a FOCCNE with any other type of forcing or is a non-constant coefficient linear equation, use VoP. If VoP fails, use Euler's to get an approximate solution.
- If the equation is nonlinear and not separable, use Euler's to approximate the solution.

The last exercise below will give you some practice.

Exercises

(1) Determine general solutions to the following first order DEs. For each example, state whether or not it would be possible to use MUC as well.

   (a) $y' + y = 1 + 2t$

   (b) $ty' + y = 1/t$

   (c) $y' + ty = te^{t^2/2}$

(2) Determine the solution $y(t)$ to the Initial Value Problem

$$\begin{aligned} ty' &= y + t^2 e^{-t} \\ y(1) &= 0 \end{aligned}$$

for $t \geq 1$.

[2] Or rather, it is the inverse of the integrating factor

(3) Determine the solution to the equation

$$t\frac{dy}{dt} = y + t^2 e^{-t}$$

under the constraint $y(1) = y_1$ and find the value of y_1 which will force $y(t) \to 0$ as $t \to \infty$.

(4) Solve the initial value problem

$$\begin{aligned} \frac{dy}{dt} &= 2y + 4 + e^{-2t} \\ y(0) &= y_0. \end{aligned}$$

and determine the value of y_0 for which your solution converges to a constant value as $t \to \infty$?

(5) State the preferred solution method (SoV, MUC, VoP, or Euler's) for each of the following equations

(a) $y' + ty = 0$

(b) $y' + ty = e^t$

(c) $y' + y = \sin t$

(d) $y' + 2y = 1$

(e) $y' + 2y = 1 + t$

(f) $y' = 2y(1 - y)$

(g) $y' = \sin(y + t)$

Part II
Second Order Differential Equations

Section 2.1: Spring-Mass Systems and Euler's Revisited An Introduction to Second Order Equations

Section objectives:

- Set up second order equations for spring-mass systems using Newton's Second Law.
- Apply Euler's Method to numerically solve second order equations.
- Compare the behavior of spring-mass systems with different parameter choices (damping, spring constants, masses, initial conditions)

So far in this course, we have focused on studying first order equations; that is, equations which only involve a first derivative of an unknown function y. In this next segment of the course, we will focus on the analysis of second order equations which frequently arise in both mechanical and electrical engineering. We start with an example from dynamics.

Spring-mass systems

Newton's second law states that the sum of the forces acting upon an object is equal to the product of the object's mass time its acceleration. In physics courses, this is often shorthanded as

$$F = ma.$$

Because acceleration is the second derivative of position, Newton's second law commonly gives rise to second order differential equations. As an example consider the spring-mass system shown in the video `http://www.youtube.com/watch?v=q-NZWDOrn4U`. This system consists of a car attached by two springs to opposite ends of the track as diagrammed in Figure 1 below.

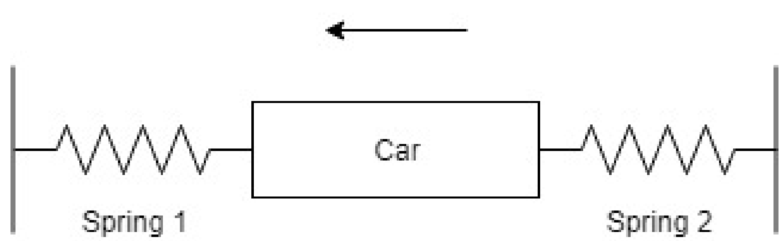


Figure 1: Diagram of 1 car, 2 spring mechanical oscillator.

At time 0, the car is displaced from equilibrium and released. Its motion is tracked and plotted as a function of time. Figure 2 shows one sample run from this experiment and we observe that the car oscillates about its equilibrium with decaying amplitude.

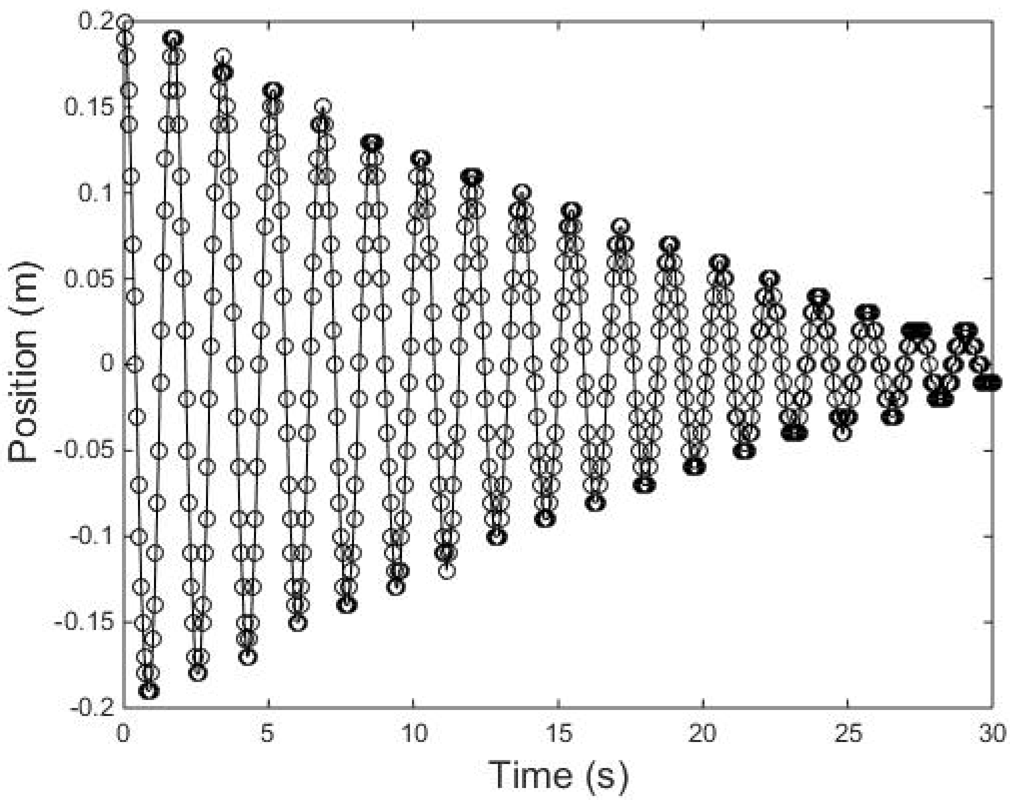


Figure 2: Sample plot of car's position vs time in the spring-mass oscillator.

Our first goal will be to develop a model for the motion of the car which describes this behavior. We will then use our model to address the following questions about how the behavior of the car in the system above depends on various parameters

- How do the weight of the car and the stiffness of the springs used impact the period and amplitude of the oscillations?
- How do the initial magnitude of displacement and velocity of release impact the oscillations?
- What effect does the damping coefficient γ have on the decay rate of the oscillations? Do we always get oscillations?
- What happens if we supply a periodic forcing to the systems?
- Is there a forcing frequency which will give us an optimal response amplitude?

In our model, we will let $y(t)$ denote the displacement of the car from equilibrium (in m) after t seconds with left (if you're looking at the cars) being the positive y direction[1]. Let m denote the mass of the car (in kg). Then Newton's Second Law implies that

$$my'' = F$$

[1]In the video, the car is initially displaced by a positive amount under this orientation.

where F is the sum of the forces acting upon the car. There are several forces acting upon the car, but let's start off by talking about the springs. A spring pulls when stretched from its resting position and pushes when contracted. The simplest way of modeling this relationship is to assume that the force exerted by each spring is directly proportional to its current displacement x from rest, but works in the opposite direction. Mathematically, we can write this as $F_S = -kx$ for some positive constant k called the *spring constant*[2]. This relationship is known as Hooke's Law and is usually a good approximation as long as the spring is not compressed too much or stretched too far. To work this into our model, we shall refer to the spring on the left as Spring 1 and the spring on the right as Spring 2, labelling all associated constants with the matching subscript. Let L_1 and L_2 denote the amount by which each spring is stretched when the car is in equilibrium and let k_1 and k_2 denote the corresponding spring constants. Then spring 1 will be stretched by $L_1 - y(t)$ at time t, but will exert a positive force while spring 2 will be stretched by $L_2 + y(t)$, but will exert a force in the negative direction. Mathematically,

$$F_1 = k_1(L_1 - y(t))$$
$$F_2 = -k_2(L_2 + y(t)).$$

Now when the car is in equilibrium ($y(t) = 0$), the forces acting upon the car are balanced yielding the equation

$$k_1 L1 - k_2 L_2 = 0.$$

Therefore, the combined spring force of the two springs is simply

$$F_S = -ky$$

where $k = k_1 + k_2$.

The others forces acting upon the car are damping forces which in this example, come in three forms: friction between the car wheels and the track, friction between the wheel axels and the car, and air resistance. (You could also add additional damping forces like a magnet to the front of the car). Damping forces directly oppose the motion of an object i.e. they depend on the object's velocity. The faster the car is moving, the greater the effect of the frictional forces. Again, the simplest way to model this force is by assuming that

$$F_{\text{damping}} = -\gamma y^{'}$$

for some positive constant γ which depends on gravity, the frictional coefficient of the track, the mass of the car, the axles, and the surrounding air conditions.

Putting all this information about forces together yields the equation

$$my^{''} = -ky - \gamma y^{'}.$$

Moving all terms involving y to the same side of the equation yields the second order differential equation

$$my^{''} + \gamma y' + ky = 0 \tag{1}$$

[2]Since force is typically measured in Newtons, the spring constant has the units of N/m and can be interpreted as the amount of restoring force added per additional meter of displacement. Basically, the spring constant is a measure of stiffness: larger spring constants correspond to stiffer springs.

for $y(t)$. If we initially displace the car by y_0 m and use an initial release velocity of v_0 m/s, then we have an initial value problem with the initial conditions $y(0) = y_0$ and $y'(0) = v_0$. If we add a forcing term which applies a force of $f(t)$ N, then we obtain a nonhomogeneous equation

$$my'' + \gamma y' + ky = f(t). \tag{2}$$

We shall refer to Equation 1 as a SOCCHE (Second Order Constant Coefficient Homogeneous Equation) and Equation (2) as a SOCCNE (replace Homogeneous with Nonhomogeneous).

Before we move on to the next section, it is worth highlighting the fact that second order initial value problems involve two initial conditions: one for the initial position and one for the initial velocity. This makes sense since both are quantities which can be tweaked at the start of the experiment to yield different trajectories: we can pull the car by a certain amount in either direction to give it an initial displacement and then give it a push in either direction to give it an initial velocity. As a general rule of thumb, an nth order linear equation requires n initial conditions to uniquely determine future behavior[3].

Euler's Method for Second Order equations

We will start our investigation of second order equations by adapting Euler's Method to this new scenario. To this end, suppose we wish to approximate the solution $y(t)$ to the differential equation

$$y'' = -y \tag{3}$$

with initial conditions $y(0) = 1$ and $y'(0) = 0$. Starting with a step size of 1/2, we start Euler's in motion by estimating $y(1/2)$ in the same manner as before:

$$y(1/2) \approx y(0) + y'(0) * (1/2) = 1.$$

However, if we wish to estimate $y(1)$, we need to do

$$y(1) \approx y(1/2) + y'(1/2) * (1/2)$$

and we run into a problem because equation (3) tells us the second derivative of y, not the first derivative so we do not know $y'(1/2)$. But the second derivative is the derivative of the first derivative so we can apply Euler's and equation (3) to estimate $y'(1/2)$ by

$$y'(1/2) = y'(0) + y''(0) * (1/2) = 0 + (-y(0))/2 = -1/2$$

and then plug this into our previous expression for $y(1)$ to get

$$y(1) \approx 1 + (-1/2)(1/2) = 3/4.$$

As before, we could summarize the calculations in a table with one new column

Step	t	$y(t)$	$y'(t)$	$y''(t) \approx -y(t)$
0	0	1	0	-1
1	0.5	1	$-1/2$	-1
2	1	3/4	Not needed	Not needed

[3]See Section 3.1 in Edwards and Penney for restrictions and conditions.

So the only new trick is that we estimate both y and y' with Euler's. If we let t_j denote the jth time, then we can summarize the general procedure as follows:

(i) Estimate $y(t_j) \approx y(t_{j-1}) + y'(t_{j-1})(t_j - t_{j-1})$.

(ii) Estimate $y'(t_j) \approx y'(t_{j-1}) + y''(t_{j-1})(t_j - t_{j-1})$.

(iii) Calculate $y''(t_j)$ using the second order equation.

In the Matlab supplement, we will discuss how to adapt both ode45 and our previous Euler's program to this new scenario.

Exercises

(1) Newton's Second Law is often used in physics to model projectile motion (those of you who took Physics 53 or well acquainted with these problems). As a simple example, suppose that a ball with mass m kg is dropped from a tall building. Then there are two main forces acting upon the ball: gravity and air resistance. Gravity accelerates the ball's descent at a roughly constant rate of $g = 9.81\, m/s^2$. Air resistance is more complicated, but assume for the sake of this problem that it exerts a force proportional to the ball's velocity, but in the opposite direction of motion.

   (a) Use Newton's Second Law to set up a second order differential equation for the height $y(t)$ of the ball in meters after t seconds.

   (b) Use the substitution $v = y'$ to convert the second order DE in (a) to a first order equation, and then solve for the ball's velocity $v(t)$ using separation of variables. Note that the initial condition is $v(0) = 0$.

   (c) Calculate the ball's terminal velocity $\lim_{t\to\infty} v(t)$.

(2) In the notes from this section, we modeled the motion of a car attached with horizontal springs to the end of a track. Suppose instead that we have an object of m kg hanging vertically from the ceiling by a single spring with spring constant k. The car is pulled down y_0 meters and released with an initial velocity of v_0 meters per second. Assuming that air resistance is proportional to the object's velocity, show that we can still use Equation (1) to model the object's motion.

(3) Suppose that we apply Euler's Method with a step size of 1 to the second order equation

$$y'' + y' + 2y = 0$$

with initial conditions y(0)=1, y'(0)=-1. What would be our estimate of y(2)?

Matlab Supplement 3: Numerical Methods for Second Order Equations

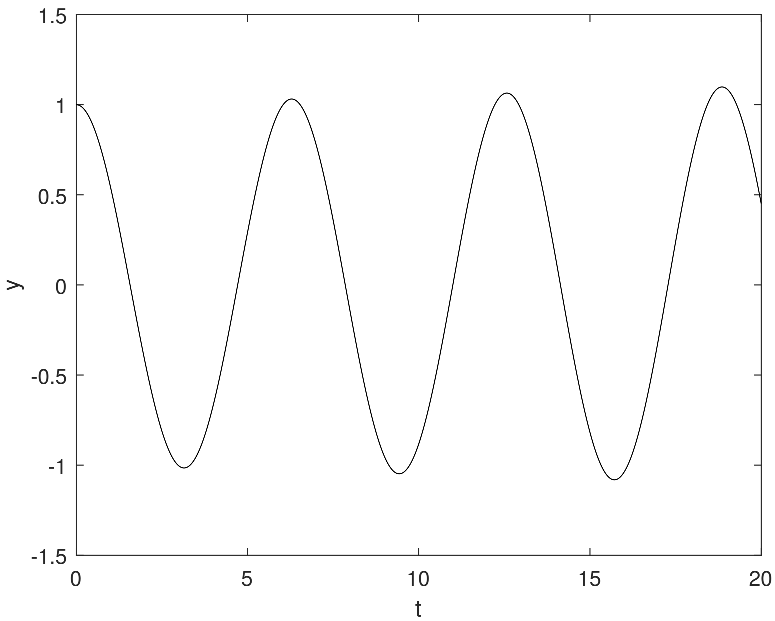
Euler's

In Section 1.3, we introduced Euler's Method, a low level numerical scheme for finding approximate solutions to first order equations. We then implemented Euler's in Matlab to solve the differential equation $y' = -y + t$ with initial condition $y(0) = 1$ up to time 1 using the following code

```
dt=1/100;                       % Define Step Size - here 1/100
t=0:dt:1;                       % Interval of t values
n=length(t);                    % Number of values
y=zeros(1,n);                   % Initialize y
dy=zeros(1,n);                  % Initialize dy
y(1)=1;                         % Initial Condition
dy(1)=-y(1)+t(1);               % Compute y'(0) using DE

for j=2:n
    y(j)=y(j-1)+dy(j-1)*dt;    % Euler's to estimate y(t_j)
    dy(j)=-y(j)+t(j);          % Compute y'(t_j) using DE
end
```

To apply Euler's to second order equations, we will use the example $y'' = -y$ with initial conditions $y(0) = 1, y'(0) = 0$. The trick to modifying our program will be to add a new line which applies Euler's to estimate y' and then use our DE to compute y'':

```
dt=1/100;                       % Define Step Size - here 1/100
t=0:dt:20;                       % Interval of t values
n=length(t);                    % Number of values
y=zeros(1,n);                   % Initialize y
dy=zeros(1,n);                  % Initialize dy
ddy=zeros(1,n);                 % Initialize ddy
y(1)=1;                         % Initial value y(0)=1
dy(1)=0;                        % Initial value y'(0)=0
ddy(1)=-y(1);                   % Compute y''(0) from DE

for j=2:n
    y(j)=y(j-1)+dy(j-1)*dt;     % Euler's to estimate y(t_j)
    dy(j)=dy(j-1)+ddy(j-1)*dt;  % Euler's to estimate y'(t_j)
    ddy(j)=-y(j);               % Compute y''(t_j) using DE
end

plot(t,y,'k')
xlabel('t')
ylabel('y')
```



Creating a function out of your code

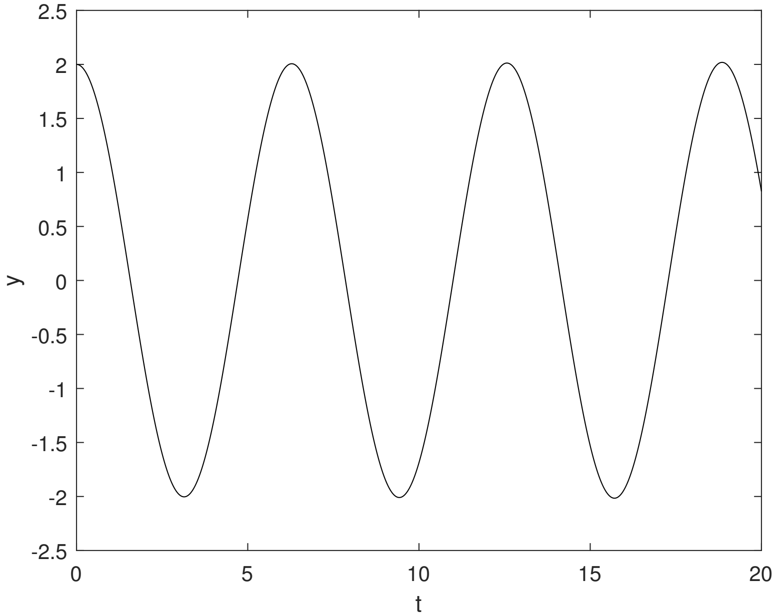
To adapt this code for different DE's only requires the modification of a few lines: the two lines giving initial conditions for y, y' and the two lines used to compute y" So if we wish to investigate how solutions to the spring-mass equation

$$my'' + \gamma y' + ky = 0$$

depend on the parameters m, γ, k or the initial conditions $y(0), y'(0)$, we could just cut and paste the above code and alter these lines accordingly. It is convenient in such situations to instead define a *function* which takes the parameters as inputs and puts them in the appropriate places within the above code. Then we need only *call* the function with a single line of code. A sample of one such function is given below. Take a moment to compare each line of this new code with the first example and note changes. Also note that in addition to the DE parameters, we have created inputs for the step size dt and the final time tfin so that we can change these as needed between problems.

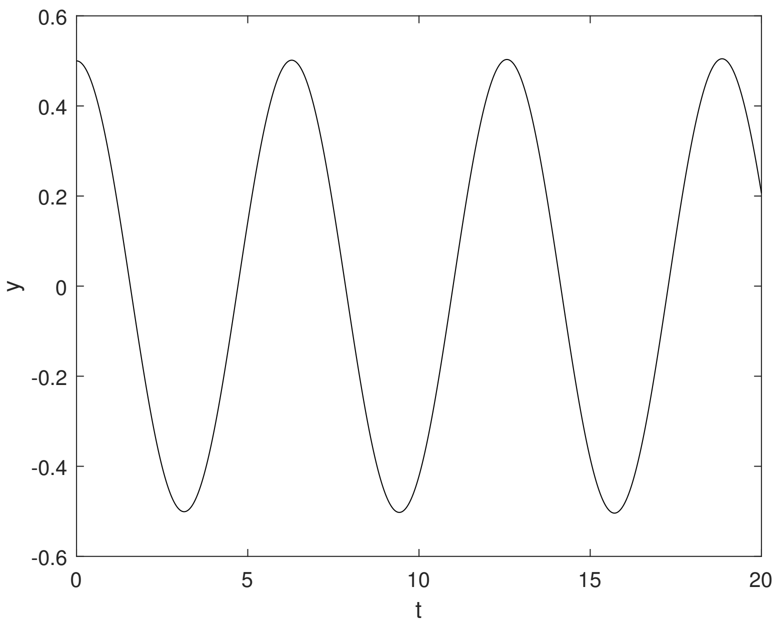
Once we have defined our function, we can use it to investigate how the solution y changes when we change one or more parameters in the DE. For example, suppose that we want to look at how the solution $y(t)$ to the DE $y^{''} = -y$ from earlier depends on the initial displacement $y(0)$. Then we can call our function with $m = k = 1$, $g = v0 = 0$ and different values of $y0$. We tried $y0 = 1$ in the last section so let's see what happens if we increase it to $y0 = 2$

```
[t,y]=socche(1,0,1,2,0,1/1000,20);
```



It looks like this just increased the amplitude of our solution to 2. Let's try $y0 = 1/2$

```
[t,y]=socche(1,0,1,1/2,0,1/1000,20);
```



At this point you can probably conjecture that $y0$ defines the amplitude of the oscillations and that the solution is just $y(t) = y_0 \cos t$. This is easily confirmed and in a couple classes, we will explain how we could come up with this on our own without Matlab. The exercises below will ask you to investigate some more interesting scenarios

Validating Solutions

In Section 1.5, we talked about the use of ode45 to validate solutions to first order DEs. You can (and should) use the SOCCHE function from this section for a similar purpose in the exercises throughout the remainder of the chapter. But there is one caveat: you MUST use a small step size.

Code for SOCCHE function

```
function [t,y]=socche(m,g,k,y0,v0,dt,tfin)
% Inputs: step size dt, final time tfin.
%         coefficients m,g,k in my''+gy'+ky=0
%         initial displacement and velocity y0,v0


t=0:dt:tfin;                        % Interval of t values
n=length(t);                        % Number of values
y=zeros(1,n);                       % Initialize y
dy=zeros(1,n);                      % Initialize dy
ddy=zeros(1,n);                     % Initialize ddy
y(1)=y0;                             % Initial value y(0)=y0
dy(1)=v0;                            % Initial value y'(0)=v0
ddy(1)=-(k/m)*y(1)-(g/m)*dy(1);    % Compute y''(0) from DE

for j=2:n
    y(j)=y(j-1)+dy(j-1)*dt;        % Euler's to estimate y(t_j)
    dy(j)=dy(j-1)+ddy(j-1)*dt;     % Euler's to estimate y'(t_j)
    ddy(j)=-(k/m)*y(j)-(g/m)*dy(j);% Compute y''(t_j) using DE
end

plot(t,y,'k')
xlabel('t')
ylabel('y')

end
```

ode45

We can also solve second order equations using ode45 as long as we first convert them into two first order equations. This is similar to the idea used in adapting Euler's Method where we estimated y'' as the derivative of y'. To demonstrate, let's again use the example $y^{''} = -y$ with initial conditions $y(0) = 1, y^{'}(0) = 0$.

First, define two new variables $y_1(t) = y(t)$, $y_2(t) = y'(t)$. In our car application, y_1 now represents position and y_2 now represents velocity. Furthermore, it is easy to see based off these new definitions, that $y_1' = y_2$ and

$$y_2' = y^{''} = -y = -y_1.$$

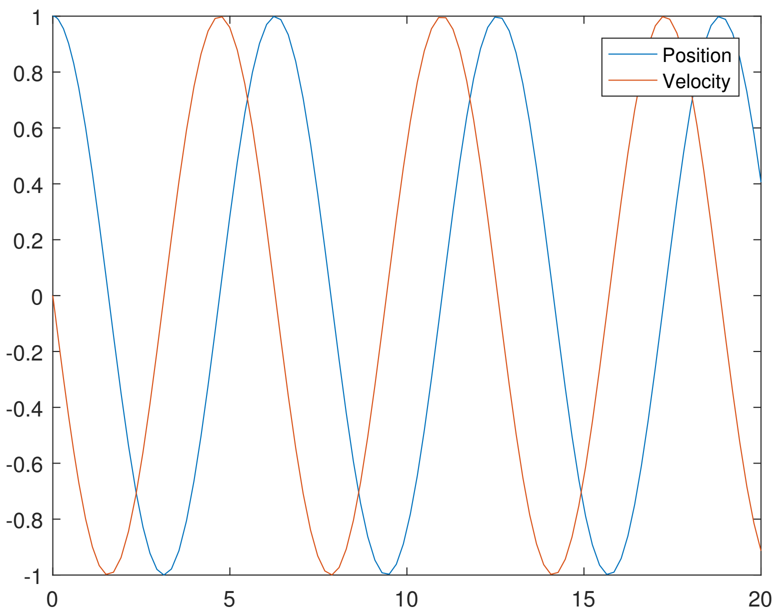
We therefore have converted our second order equation $y^{''} = -y$ into two first order equations

$$y_1' = y_2,\ y_2' = -y_1.$$

Our initial conditions have become $y_1(0) = 1, y_2(0) = 0$.

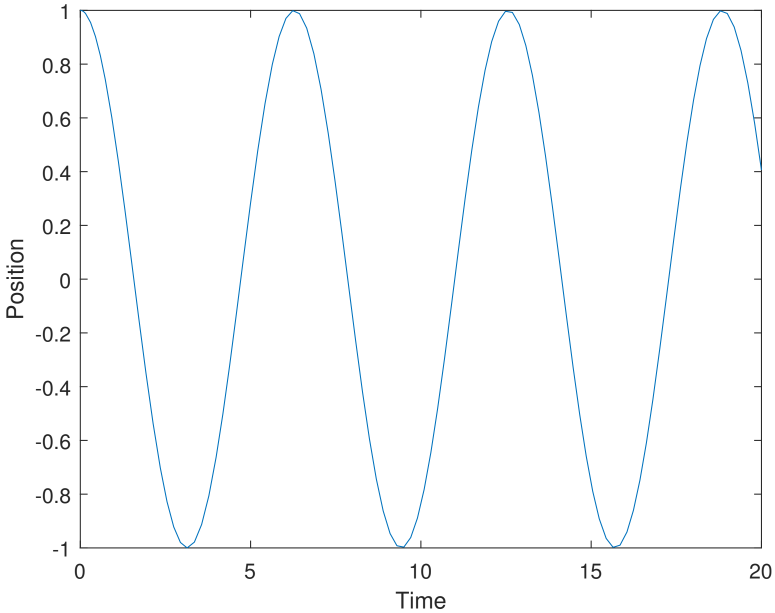
The trick to using ode45 for this new scenario is to define our rate function f as a vector-valued function of t, y_1, and y_2 whose first component is the function $f_1(t, y_1, y_2) = y_2$ and whose second compontent is the function $f_2(t, y_1, y_2) = -y_1$. Recall the Matlab notation "[1;2]" defines a vector whose first value is 1 and second is 2. We can do the same thing with function handles, defining a function of t, y whose first value is $y(2)$ and second is $-y(1)$. We can also input our initial conditions a vector [1;0]. Putting this all together suggests the following code.

```
f=@(t,y) [y(2);-y(1)];
[t,y]=ode45(f,[0,20],[1;0]);
plot(t,y)
legend('Position','Velocity')
```



Note that this code actually plots both y_1 and y_2 which is why we slapped on a legend. If you want to just plot position, you have to isolate the first column of the Matlab variable y as this contains all the position measurements. This can be done with the colon operator, whose use we will delve into deeper during Chapter 3.

```
plot(t,y(:,1))
xlabel('Time')
ylabel('Position')
```



Exercises

(1) Euler's Method can lead to inaccurate "blow-ups" of solutions if you don't use a small enough step size. As an example, use Euler's Method to solve the differential $y^{''} = -y$ with initial conditions $y(0) = 1, y^{'}(0) = 0$ over the interval $[0, 20]$ using two different step sizes: $\Delta t = 0.1$ and $\Delta t = 0.001$. Plot the solutions on the same axes and explain the issue with the larger step size.

(2) Use either Euler's Method or ode45 to investigate how the solution of the second order equation

$$y^{''} + \gamma y' + 2y = 0$$

depends on the damping coefficient γ. Include nicely labeled plots $y(t)$ vs. t for each choice of γ in (a)-(d) below (either on the same or separate axes). Use the initial conditions $y(0) = 0.1, y'(0) = 0$ and step size $\Delta t = 0.001$. Write a few sentences describing how the behavior of $y(t)$ depends on γ and comment on which choice of γ best matches the behavior we observed in the real live spring-mass system.

(a) $\gamma = 0$ (the case of an ***undamped*** or ***harmonic*** oscillator)

(b) $\gamma = 0.1$.

(c) $\gamma = 1$.

(d) $\gamma = 3$.

(3) In our class demo, we also made some observations about how the displacement function $y(t)$ depends on the mass of the car. Investigate this dependence by plotting solutions to the DE

$$my^{''} + y = 0$$

with initial conditions $y(0) = 0.1, y'(0) = 0$ for at least four different values of m. Discuss.

Section 2.2 : Second Order Constant Coefficient Homogeneous Equations 1

Characteristic Equations with Real Roots

Class objectives:

- Find exponential solutions to second order constant coefficient homogeneous equations (SOCCHE) via characteristic equations
- Apply the Principle of Superposition to form general solutions to SOCCHE.
- Solve initial value problems involving SOCCHE with real roots.

In Section 2.1, we modelled the motion of a car attached with springs to two ends of a track by setting up the second order differential equation

$$my'' + \gamma y' + ky = 0 \tag{1}$$

for the displacement $y(t)$ of the car after t seconds. We then investigated solutions to such equations using Euler's method. For example, Figure 1 compares how the solutions depends on the damping coefficient γ.

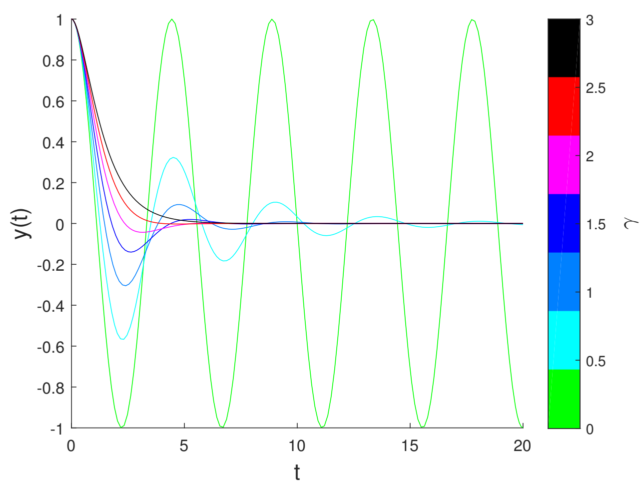


Figure 1: Euler's Method approximations of the displacement $y(t)$ when $m = 1$, $k = 2$ for different values of γ ranging from 0 to 3.

Notice that there are essentially four regimes of behavior

Undamped: When $\gamma = 0$, the output oscillates with constant amplitude

Underdamped: When γ is small, but positive, the output oscillates with decaying amplitude

Critically Damped: When γ is at an intermediate level, the output passes through equilibrium once, but then converges towards equilibrium

Overdamped: When γ is large, the output decays exponentially fast towards equilibrium.

Today, we will develop an analytic technique for deriving general solutions to equations of the form (1) which will serve as a general example of a Second Order Constant Coefficient Homogeneous Equation (SOCCHE). Our goal will be to explain the four regimes listed above and provide simple criteria for classifying systems into these different cases. The idea behind our method will be to find two separate solutions and then use the Principle of Superposition to combine them into a general solution. This process bears many similarities to the developments of Section 1.4 which we encourage you to review before reading this section.

Characteristic Equation

By analogy with first order constant coefficient homogeneous equations $y' = ay$ which have exponential solutions, we might guess that Equation (1) also has a solution of the form $y(t) = e^{\lambda t}$ for some constant λ. The idea, similar to what we did in MUC, is to plug this candidate solution into Equation (1) for y and find all values of λ for which:

$$m(e^{\lambda t})'' + \gamma(e^{\lambda t})' + k(e^{\lambda t}) = 0$$

or after taking derivatives,

$$m\lambda^2 e^{\lambda t} + \gamma\lambda e^{\lambda t} + ke^{\lambda t} = 0.$$

Dividing both sides by $e^{\lambda t}$ yields the equation

$$m\lambda^2 + \gamma\lambda + k = 0. \tag{2}$$

Any λ which solves Equation (2) will yield a solution $e^{\lambda t}$ to the general SOCCHE (1) and conversely, any solution of the form $e^{\lambda t}$ must have an exponent λ satisfying Equation (2). For this reason, we call (2) the *characteristic equation* associated with (1). It is a quadratic equation in λ and can easily be solved in any given situation using high school algebra - see Exercise 1 at the end of the section for some examples

The characteristic equation allows us to generate solutions, but suppose we need to calibrate our solution to meet some specific initial position and velocity. For example, let's consider the DE $y'' + 3y' + 2 = 0$ which appears in part (a) of Exercise 1 below and has exponential solutions e^{-2t}, e^{-t}. Suppose that we want a particular solution to this equation which satisfies the conditions $y(0) = 1$ and $y'(0) = 0$. Both e^{-2t} and e^{-t} would work for $y(0) = 1$, but neither has $y'(0) = 0$. We need some more free parameters (or *degrees of freedom*)! For first order of equations, an initial condition $y(0) = y_0$ required one free constant in our general solution. Now we will need two free constants to meet the constraints $y(0) = y_0$ and $y'(0) = v_0$ so the general solution to (1) should look something like

$$y(t) = c_1 \times (stuff) + c2 \times (stuff).$$

To find a solution of this form, we appeal to an old friend.

The Principle of Superposition Revisited

Suppose that $y_1(t)$ and $y_2(t)$ are two different solutions to equation (1). The first consequence of the principle of superposition (POS1) in Section 1.4 stated that the sum of any two solutions to a linear, homogeneous first order equation is also a solution. Now we will provide a generalization of this fact to second order equations (which is similar to Exercise 1 from that section). A *linear combination* of y_1 and y_2 is any function of the form

$$y(t) = c_1 y_1(t) + c_2 y_2(t)$$

where c_1, c_2 are arbitrary constant. If we have a linear combination of two functions y_1, y_2, then we know from elementary calculus that

$$y' = c_1 y_1' + c_2 y_2'$$

and

$$y'' = c_1 y_1'' + c_2 y_2''.$$

Therefore,

$$\begin{aligned} my'' + \gamma y' + ky &= c_1 m y_1'' + c_2 m y_2'' + c_1 \gamma y_1' + c_2 \gamma y_2' + c_1 k y_1 + c_2 k y_2 \\ &= c_1(m y_1'' + \gamma y_1' + k y_1) + c_2(m y_2'' + \gamma y_2' + k y_2) \end{aligned}$$

where in the second line we have simply grouped together the terms involving y_1 and y_2. Now since y_1 and y_2 are solutions to (1), both terms in parentheses are 0 and hence

$$my'' + \gamma y' + ky = 0.$$

Thus we have extended POS1 to show that any linear combination of solutions to (1) is itself a solution.

To see this in application, consider our quandary from the end of the last section: we wanted to find a solution y to the equation $y'' + 3y' + 2y = 0$ with initial conditions $y(0) = 1$ and $y'(0) = 0$. Consider the function

$$y(t) = c_1 e^{-2t} + c_2 e^{-t}$$

formed by taking linear combinations of the two solutions found in part (a) of Exercise 1. The principle of superposition tells us that this is also a solution to the differential equation of interest. Since

$$y'(t) = -2c_1 e^{-2t} - c_2 e^{-t}$$

we can plug in our two initial conditions to obtain a system of two equations

$$\begin{aligned} 1 &= c_1 + c_2 \\ 0 &= -2c_1 - c_2 \end{aligned}$$

for the two unknowns c_1, c_2. This system has solution $c_1 = -1$, $c_2 = 2$ so

$$y(t) = -e^{-2t} + 2e^{-t}$$

solves the desired initial value problem.

Does this process always work for SOCCHE?

Yes and no depending on what you mean by this question. To elucidate, let's go back and re-question the procedure we used to solve the DE $y'' + 3y' + 2y = 0$. The characteristic equation gives us two solutions e^{-t}, e^{-2t} to the given DE, but the principle of superposition immediately implies that $2e^{-t}$ and $2e^{-2t}$ are also solutions. So why not form the general solution

$$y(t) = c_1 e^{-t} + c_2(2e^{-t})$$

for example? Well, if we try this, then plugging in the initial conditions leads to the system of equations

$$\begin{aligned} 1 &= c_1 + 2c_2 \\ 0 &= -c_1 - 2c_2. \end{aligned}$$

Adding these two equations together, we get the ridiculous statement $1 = 0$. In other words, we obtain an *inconsistent system of equations* for c_1, c_2; a system with no solutions. On the other hand, in the previous section when we applied the principle of superposition with two solutions which were not related by a scalar multiple, e^{-t} and e^{-2t} in our example, we obtained a system of equations which had a unique solution for c_1, c_2. So the key to solving SOCCHE is to start with two solutions which are not multiples of each other and we call such solutions *linearly independent*. The key to solving second order linear equations is to find two linearly independent solutions.

How can you tell if two solutions $y_1(t), y_2(t)$ are linearly independent? Well, it seems pretty obvious: you just check that there is no constant c such that $y_1(t) = cy_2(t)$ or vica versa. Some textbooks also discuss the computation of a determinant called the Wronskian. Although I personally love the name, it is actually not so important for the continuity of our story to indulge in this topic here, but you can check out Section 3.1 of Edwards and Penney for further details.

For our present pursuits, it is more relevant to discuss *when* the characteristic equation will yield two linearly independent solutions and what to do if it doesn't. The characteristic equation is a quadratic function of λ so from high school algebra, we know that there are three possible scenarios for its roots:

(I) Two distinct real roots

(II) One (repeated) real root

(III) No real roots.

Examples of all three scenarios are illustrated in Exercise 1 below and every characteristic equation will result in one of these three situations (see also Exercise 4 below). In Scenario (I), we are guaranteed two linearly independent solutions $e^{\lambda_1 t}, e^{\lambda_2 t}$ corresponding to the two distinct real roots $\lambda_1 \neq \lambda_2$. However, in Scenario (II) we obtain only one linearly independent solution $e^{\lambda_1 t}$ corresponding to the one real root of the characteristic equation. The example below shows how to generate a second linearly independent solution in this scenario.

The case of one repeated root

Consider the SOCCHE

$$y'' + 6y' + 9y = 0$$

The characteristic equation is

$$\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2$$

which has just one solution $\lambda = -3$. Hence, one solution to the SOCCHE is e^{-3t}. Interestingly, the product rule, shows that te^{-3t} also works:

$$(te^{-3t})'' + 6(te^{-3t})' + 9(te^{-3t}) = 0.$$

So the general solution is

$$y(t) = c_1 e^{-3t} + c_2 te^{-3t}.$$

The procedure applied above will work in any similar single root scenario[1] - once a single solution of the form $e^{\lambda t}$ is found, a second solution is always $te^{\lambda t}$. See the exercises for further examples.

The case of no real roots

Now we move onto the interesting case when the characteristic equation was no real valued solutions. Consider the DE

$$y'' + 4y = 0$$

from part (d) of Exercise 1 below. The characteristic equation has no real roots, but does have two imaginary roots of $\pm 2i$. Therefore, if we can make sense out of the imaginary exponentials $e^{\pm 2it}$, then the principle of superposition will yield the general solution

$$y(t) = c_1 e^{2it} + c_2 e^{-2it}.$$

That will be our task for the next section. If you are suspicious of the usefulness of imaginary numbers (as I once was), then that section will hopefully change your opinion as we shall see that imaginary numbers are a key component of modeling spring-mass and other mechanical/electrical oscillators.

A brief comment on solutions to systems of equations

In solving second order IVPs, we will always encounter linear systems of algebraic equations for the unknown coefficients c_1, c_2. For example, earlier in the section we saw the system

$$\begin{aligned} c_1 + c_2 &= 1 \\ -2c_1 - c_2 &= 0 \end{aligned}$$

[1] To justify this statement, suppose that we are given a generic second order equation $ay'' + by' + cy = 0$ whose characteristic equation has a single root. Then $b^2 - 4ac = 0$ and this fact can be used to show that $e^{(-b/2a)t}$ and $te^{(-b/2a)t}$ are both solutions in a manner similar to the example.

which has solution $c_1 = -1$, $c_2 = 2$. Elimination of variables is sufficient for solving any systems encountered in this chapter, but it is worth mentioning that there is a whole mathematical field motivated by the search for solutions to such linear systems, namely *linear algebra*. We shall delve a bit deeper into this field in Chapter 3, but if you have previously taken a course in this subject, we invite you to apply what you learned there at will in this course.

Exercises

(1) Find all solutions $e^{\lambda t}$ to equation (1) for the following choices of m, γ, k.

  (a) $m = 1, \gamma = 3, k = 2$
  (b) $m = 1, \gamma = -6, k = 9$
  (c) $m = 1, \gamma = -1, k = -2$
  (d) $m = 1, \gamma = 0, k = 4$.

(2) Find a solution to the DE in part (c) above which has initial conditions $y(0) = 3$ and $y'(0) = 3$.

(3) Find a solution to the DE in part (b) above which has initial conditions $y(0) = 0, y'(0) = -1$.

(4) Determine a relationship between m, γ, k in Equation (1) which will tell us whether the characteristic equation has two, one, or no real roots. *Hint*: Use the quadratic formula and classify roots using the discriminant

(5) Apply your rule from the last exercise to classify each of the following SOCCHE as having characteristic equations with two, one, or no real roots.

  (a) $y'' + y = 0$
  (b) $2y'' + 8y' + 4y = 0$
  (c) $4y'' + 8y' + 4y = 0$

(6) Write out the characteristic equations and general solutions to the following SOCCHE

  (a) $y'' + y' - 6y = 0$
  (b) $2y'' + 8y' + 8y = 0$

(7) Solve the following IVPs and validate your solutions in Matlab. Describe what happens to the solutions as $t \to \infty$.

  (a) $y'' + 4y' + 3y = 0$, $y(0) = 2$, $y'(0) = 0$.
  (b) $y'' + 2y' + y = 0$, $y(0) = 0$, $y'(0) = 1$.
  (c) $y'' + 7y' + 6y = 0, y(0) = 1, y'(0) = 0$

(8) A 1 kg mass is suspended from a spring with constant $k = 6$ in a medium with damping coefficient $\gamma = 5$. The mass is pulled 1 meter from equilibrium and pushed with an initial velocity of -1 m/s. Determine a formula for the displacement of the mass after t seconds. Validate your solution using Matlab.

Section 2.3: SOCCHE Part 2

Euler's Formula and Characteristic Equations with Complex Roots

Section objectives:

- Perform basic algebraic operations on complex numbers
- Define e^{ix} via Euler's Formula
- Solve SOCCHE IVPs in the case where the characteristic polynomials has no real roots.

Last class, we laid the framework for solving equations of the form

$$my'' + \gamma y' + ky = 0 \tag{1}$$

by finding the roots to the characteristic equation

$$m\lambda^2 + \gamma\lambda + k = 0 \tag{2}$$

and then using the principle of superposition to form the general solution. We distinguished between three different scenarios:

(i) The characteristic equation has two real roots.

(ii) The characteristic equation has only one real root.

(iii) The characteristic equation has no real roots.

The purpose of today's class is to explain why the third scenario leads to oscillations (the un- or under-damped scenarios) by making sense out of e^{it}. We'll start with a brief review of complex arithmetic.

Multiplication gets Complex

When I was first introduced to imaginary numbers in high school, I was outraged that they would waste our time with shit that wasn't even real. I vividly remember the moment in my own undergraduate differential equations course when my perspective changed and I began to appreciate the beauty of this intriguing extension to the real number system. Imaginary numbers came about in the search for solutions to quadratic expressions. For example, the very simple quadratic expression

$$x^2 + 1 = 0$$

has no real roots (the graph of $y = x^2 + 1$ is a parabola that does not touch the x axis), but if we define $i = \sqrt{-1}$ and allow multiples of i to be solutions, then it actually has two solutions: $x_{1,2} = \pm i$. The fundamental theorem of algebra states that if we extend our search for roots to quadratic functions to include expressions of the form $a + ib$, then we will always be able to find two solutions or have a repeated real root. In this sense, the real numbers can say to the complex numbers "you complete me."

A ***complex number*** is any number of the form $z = x + iy$ where x, y are real numbers and $i = \sqrt{-1}$ is called the imaginary unit. x is referred to as the **real part** and y as the **imaginary part**. We add two complex numbers by adding the corresponding real and imaginary parts so that if $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2).$$

If c is a real number, then we take the product of c and z in the obvious way, multiplying the real and imaginary parts of z by c:

$$cz = cx + icy.$$

A slightly more complicated operation is taking the product of two complex numbers. This is done in a manner analogous to the "FOIL" method of multiplying linear terms together and then using the rule that $i^2 = -1$ to simplify the result:

$$\begin{aligned} z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1 x_2 + ix_1 y_2 + iy_1 x_2 + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \end{aligned}$$

To define division, we first need to define the conjugate of a complex number. If $z = x + iy$, then the ***conjugate*** of z, denoted $\bar{z}$, is $x - iy$ (change the sign of the imaginary part). Every real number is its own conjugate (a real number is just a complex number with no imaginary part). Furthermore, if we multiply any complex number by its conjugate, we get a real number whose value is just the sum of the real and imaginary parts squared:

$$z\bar{z} = (x + iy)(x - iy) = x^2 - ixy + ixy - i^2 y^2 = x^2 + y^2.$$

We then define division through multiplication and conjugation:

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{z_2 \bar{z}_2}.$$

One interesting consequence of this fact is that

$$\frac{1}{i} = \frac{i}{i^2} = -i.$$

Geometrically, it is useful to think of complex numbers as points in the xy-plane: a complex number $z = x + iy$ can be associated with the point (x, y). Under this visualization, the x-axis is called the real axis and y-axis is called the imaginary axis. This association also motivates a definition of the absolute value or *modulus* of a complex number $z = x + iy$ as the Euclidean distance from the point (x, y) to the origin:

$$|z| = \sqrt{x^2 + y^2}.$$

For example, if $z = 1 - 2i$, then $|z| = \sqrt{5}$. Also note that $|z| = \sqrt{z\bar{z}}$.

Getting used to complex arithmetic, like most things in mathematics, requires a bit of practice so we recommend completing Exercise 1 at the end of the section before you move on.

Euler's formula and applications.

The key to making sense out of "imaginary exponentials" is Taylor series. Recall that the Taylor series expansions of $e^x, \cos x$, and $\sin x$ about 0 are:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$
$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.$$

Setting $x = iy$ in the Taylor series for e, we define e to an imaginary number iy by

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} \tag{3}$$

and then e to a complex number $z = x + iy$ by

$$e^z = e^x e^{iy}.$$

To obtain a better idea of what this definition means, we write out the first few terms of the series:

$$e^{iy} = 1 + iy + i^2y^2/2 + i^3y^3/3! + i^4y^4/4! + i^5y^5/5! + i^6y^6/6! + i^7y^7/7! + \cdots.$$

Since $i^2 = -1$, $i^3 = i^2 i = -i$, $i^4 = i^2 i^2 = 1$, $i^5 = ii^4 = i$, etc (see the pattern?), we can break the above expression into real and imaginary parts to obtain

$$e^{iy} = (1 - \frac{y^2}{2} + \frac{y^4}{4!} - \cdots) + i(y - \frac{y^3}{3!} + \frac{y^5}{5!} - \cdots).$$

Observing that the two terms in parentheses are the Taylor series expansions of $\cos x$ and $\sin x$, respectively, we obtain a class mathematical equality known as *Euler's Formula*:

$$e^{iy} = \cos y + i \sin y. \tag{4}$$

There are several interesting consequences of Euler's Formula:

- Applying Euler's formula with $y = \pi$ yields

  $$e^{i\pi} = \cos \pi + i \sin \pi = -1$$

  which can be rewritten as

  $$e^{i\pi} + 1 = 0,$$

  a result often referred to an Euler's Identity. You may have seen bumper stickers with this equation (if you run in such circles). It's a favorite of mathematicians because it is a single equation which involves five of our favorite numbers: 1, 0, π, i, and e.

- Since cos is an even function ($\cos(-y) = \cos(y)$)and sin is an odd function ($\sin(-y) = -\sin y$), Euler's formula implies that
$$e^{-iy} = \cos(-y) + i\sin(-y) = \cos y - i\sin y. \tag{5}$$
In other words, e^{iy} and e^{-iy} are complex conjugates.

- Using the last result, we obtain the identities
$$\begin{aligned} e^{iy} + e^{-iy} &= 2\cos y \\ e^{iy} - e^{-iy} &= 2i\sin y. \end{aligned}$$
We shall use these frequently in what follows and refer to them as **Euler's Trig Identities**. They basically state that adding (subtracting) two complex conjugate exponentials results in a cos (sin) function.

- Euler's formula also allows to derive formulas for the derivative and integral of e^{it}. For example,
$$\begin{aligned} \frac{d}{dt}\left(e^{it}\right) &= \frac{d}{dt}(\cos t + i\sin t) \\ &= -\sin t + i\cos t \\ &= i(\cos t + i\sin t) = ie^{it}. \end{aligned}$$
A similar computation yields the nice formula
$$\frac{d}{dt}\left(e^{zt}\right) = ze^{zt}. \tag{6}$$
In other words, derivatives of e to a complex exponent behave just like derivatives of e to a real exponent. Similar analogies hold for integration (see exercises for examples).

- Euler's Formula can also be used to obtain a "polar decomposition" for complex numbers. Recall that the polar coordinates of a point (x, y) in the plane is given by the pair (r, θ) where $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$. For a complex number $z = x+iy$, we can write it as
$$z = re^{i\theta}$$
where $r = \sqrt{x^2 + y^2}$ is the modulus of z and $\theta = \arctan(y/x)$. We don't really need this fact for these notes, but mention it because it is cool and important to the subject of complex analysis!

In this course, our main use of Euler's Formula will be to take expressions such as
$$e^{(2-3i)t} + e^{(2+3i)t}$$
and rewrite them as trig functions. To do so, first break up the exponents and factor out like terms to obtain
$$e^{(2-3i)t} + e^{(2+3i)t} = e^{2t}(e^{-3it} + e^{3it}).$$
Then apply the appropriate Euler's trig identity, in this case
$$e^{3it} + e^{-3it} = 2\cos 3t,$$
to further simplify your result as
$$e^{2t}(e^{-3it} + e^{3it}) = 2e^{2t}\cos(3t).$$
Exercise 3 below provides some additional practice examples.

Return of SOCCHE

Now that we know how to make sense out of e^{it}, let's return to our investigation of solutions to SOCCHE equations. Consider the equation $y'' = -y$ with initial conditions $y(0) = 1$ and $y'(0) = 0$ which we used to demonstrate Euler's Method in lab last week. This DE yields the characteristic equation

$$\lambda^2 + 1 = 0$$

which has complex solutions $\lambda = \pm i$. The principle of superposition then implies that

$$y(t) = c_1 e^{it} + c_2 e^{-it}$$

is a solution to $y'' = -y$ for any constants c_1, c_2, Plugging in our initial conditions yields a system of two equations

$$\begin{aligned} 1 &= c_1 + c_2 \\ 0 &= ic_1 - ic_2 \end{aligned}$$

which has solution $c_1 = c_2 = 1/2$. Now we can use Euler's trig identity $e^{it} + e^{-it} = 2\cos t$ to simplify our solution

$$\begin{aligned} y(t) &= \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} \\ &= \frac{1}{2}(2\cos t) = \cos t. \end{aligned}$$

And voila, from something imaginary something real comes and we get the nice solution $y(t) = \cos t$.

Let's look at one slightly more complicated example. Suppose that we want to solve the IVP $y'' + 2y' + 2y = 0,\ y(0) = 2,\ y'(0) = 0$. The characteristic equation is

$$\lambda^2 + 2\lambda + 2 = 0$$

which has roots $\lambda = -1 \pm i$. Therefore, the general solution is

$$y(t) = c_1 e^{(-1+i)t} + c_2 e^{(-1-i)t}$$

which has derivative

$$y'(t) = (-1+i)c_1 e^{(-1+i)t} - (1+i)c_2 e^{(-1-i)t}.$$

Plugging in the two initial conditions yields the system

$$\begin{aligned} 2 &= c_1 + c_2 \\ 0 &= (-1+i)c_1 - (1+i)c_2. \end{aligned}$$

To solve this, multiply the first equation by $1 + i$ and add to get $2(1 + i) = 2ic_1$ or $c_1 = 1 + 1/i, c_2 = 1 - 1/i$ (note that the two solutions are complex conjugates). This is looking bad, but now we organize together sums and differences of conjugate exponents, applying Euler's trig identities where appropriate

$$\begin{aligned} c_1e^{(-1+i)t} + c_2e^{(-1-i)t} &= e^{-t}\left[(e^{it}+e^{-it}) + (e^{-it}-e^{-it})/i\right] \\ &= e^{-t}(2\cos t + (2i\sin t)/i) \\ &= 2e^{-t}(\cos t + \sin t). \end{aligned}$$

We could furthermore use the techniques of Section 1.5 to rewrite this as

$$y(t) = 2ae^{-t}\sin(t+\phi)$$

where $a = \sqrt{2}$ and $\phi = tan^{-1}(1) = \pi/4$.

Long Term Behavior based on roots

Additional practice problems are provided in the exercises, but let's stop to enjoy the fruit of our labor thus far. We have essentially observed two types of behavior resulting from roots with imaginary components

- Pure imaginary roots ($\pm i$) led to constant amplitude oscillations
- Complex roots with negative real part ($-1 \pm i$) led to exponentially decaying oscillations.

It is no coincidence that these two situations correspond to the previously mentioned un- and under-damped scenarios. In general, complex roots yield oscillations as a consequence of Euler's formula and whether or not these oscillations decay, remain constant, or increase in amplitude over time depends on the real part of the complex valued roots to the characteristic equation. In fact, this relationship can often be ascertained without diving into the messy algebra of solving for c_1, c_2 as we did in the previous section. If our only goal is to deduce long term or asymptotic behavior as $t \to \infty$, then the general solution is usually sufficient and this involves no more algebra than solving a quadratic equation.

To demonstrate, consider the DE

$$y^{''} + 2y' + 2y = 0$$

from the previous section which had general solution

$$y(t) = c_1e^{(-1+i)t} + c_2e^{(-1-i)t} = e^{-t}\left(c_1e^{it} + c_2e^{-it}\right).$$

We know that the term in parentheses will be some linear combinations of $\sin t$ and $\cos t$ as result of Euler's Formula and hence, will oscillate with period[1] 2π. The term out front, however, decays to 0 as $t \to \infty$ so without any further work, we know that this solution will exhibit decaying oscillations.

[1]Technically, this is only the quasi-period of y since the function doesn't strictly repeat itself, but we will mostly ignore this distinction for now.

Similarly, suppose we obtain a DE whose characteristic equation has roots $3 \pm 2i$ so that the general solution is

$$y(t) = c_1 e^{(3+2i)t} + c_2 e^{(3-2i)t} = e^{3t}\left(c_1 e^{2it} + c_2 e^{-2it}\right).$$

The term in parentheses is now a combination of $\sin 2t$ and $\cos 2t$ and hence, has period π while the term out front blows up exponentially fast as $t \to \infty$. We conclude that the solution will exhibit growing amplitude oscillations with period π.

In general, the above examples hopefully show how the numbers a, b in a complex root $a \pm bi$ relate to the behavior of the solution $y(t)$: b adjusts the period of the oscillations (Period $= 2\pi/b$) and a determines if the oscillations grow ($a > 0$), stay constant ($a = 0$), or decay ($a < 0$) as $t \to \infty$. The real power of the Method of Characteristic Equations for solving SOCCHE lies in the simplicity of this approach. In a couple classes, we will learn a more abstract, but often quicker method for solving SOCCHE IVPs if particular solutions are required. But first, we turn briefly to nonhomogeneous equations.

Exercises

(1) Let $z_1 = 1 + 2i$, $z_2 = 2 - i$, and $c = -1$. Calculate the following expressions

- (a) $z_1 + z_2$
- (b) cz_1
- (c) $z_1 z_2$
- (d) $\bar{z}_1$
- (e) $|z_1|$
- (f) $\frac{z_1}{z_2}$

(2) What is (a) i^{100}? (b) i^{2018}?

(3) Simplifying the following expressions using Euler's trig identities

- (a) $-\frac{1}{2i}(e^{4it} - e^{-4it})$
- (b) $e^{(-2+3i)t} + e^{(-2-3i)t}$
- (c) $(1+i)e^{(-2+3i)t} + (1-i)e^{(-2-3i)t}$

(4) Calculate the following complex integrals

- (a) $\int_0^{2\pi} e^{it}\, dt$
- (b) $\int_0^{\pi} e^{it}\, dt$
- (c) $\int_0^{\pi} e^{imt}\, dt$ where m is any even integer.

(5) Determine the solutions $y(t)$ to the differential equation

$$y'' + y = 0$$

which satisfy the initial conditions in (a) and (b) below. Compare each with the solution $y(t) = \cos t$ from the notes which had $y(0) = 1, y'(0) = 0$.

(a) $y(0) = 0, y'(0) = 1$

(b) $y(0) = 1, y'(0) = 1$.

(6) Determine solutions $y(t)$ to

$$y'' + 4y = 0$$

with the following initial conditions. How do the answers compare with solutions to $y'' + y = 0$ with similar initial conditions?

(a) $y(0) = 1, y'(0) = 0$.

(b) $y(0) = 1, y'(0) = 2$

(7) Solve the IVP

$$y'' + 2y' + 5y = 0,\ y(0) = 2,\ y'(0) = 0.$$

Validate your solution using Matlab.

(8) Consider the second order equation $ay'' + by' + cy = 0$ where a, b, c are all constants. For each of the following choices of a, b, c, write down the general solution for y and determine $\lim_{t\to\infty} y(t)$ (or explain why the limit does not exist).

(a) $a = 4, b = 4, c = 1$

(b) $a = 1, b = -2, c = 13/4$.

(9) The charge $q(t)$ across a capacitor in an RLC circuit with no external forcing satisfies the differential equation

$$Lq'' + Rq' + \frac{1}{C}q = 0$$

for some constants $L, R, C > 0$. For each of the following choices of L, R, C, write down the general solution for the charge and determine $\lim_{t\to\infty} q(t)$ (or explain why the limit does not exist).

(a) $L = 1, R = 2, C = 1$.

(b) $L = 2, R = 4, C = 1/10$.

(c) $L = 1, R = 0, C = 9$.

(10) In Matlab Supplement 3, the exercises investigated how the solution to

$$my'' + y = 0$$

depends on the mass m of the car. Write out the general solution to this DE and find a formula for the period of y in terms of m. Compare the results of Exercise 3 from the supplement.

(11) The exercises from Matlab Supplement 3 also investigated how the solution to

$$y'' + \gamma y' + 2y = 0$$

depends on the damping coefficient γ. Write out the general solution to this system and describe how the solution depends on γ. Compare with the results of the Exercise 2 from the supplement

(12) Use the Taylor series expansions for $\cos, e$ to show that

$$\cos(it) = \frac{e^t + e^{-t}}{2}$$

and find a similar identity for $\sin(it)$.

Section 2.4: Resonance and Frequency-Response Functions

Section objectives:

- Find solutions to SOCCNE with constant and periodic forcing functions.
- Compute resonant frequencies and frequency-response gains for second order equations.

Now that we have a complete set of methods for solving SOCCHE equations, we move on to discuss the nonhomogeneous system

$$my'' + \gamma y' + ky = g(t) \tag{1}$$

where $g(t)$ is a suitable forcing function. The procedure will be similar to what we did in Section 1.4 for homogeneous first order equations:

(i) We choose an appropriate trial solution $y_p(t)$ which has the same form as $g(t)$ and solve for the unknown coefficients.

(ii) We find the general solution $y_c(t)$ to the complementary equation

(iii) We form the general solution $y(t) = y_p(t) + y_c(t)$ and plug in the appropriate initial conditions.

Let's demonstrate with two important classes of examples.

Constant forcing

For our first example, let's set $m = 1$, $\gamma = 4$, $k = 3$ and $g(t) = 6$ in equation (1). In the spring-mass example, think of applying a constant force of 6 Newtons to the car. The complementary equation in this case is the SOCCHE

$$y_c'' + 4y_c' + 3y_c = 0$$

which has characteristic equation

$$\lambda^2 + 4\lambda + 3 = (\lambda + 3)(\lambda + 1) = 0 \Longrightarrow \lambda = -3, -1.$$

The general solution to the complementary equation is therefore,

$$y_c(t) = c_1 e^{-3t} + c_2 e^{-t}.$$

To find a particular solution to equation (1), we note that since the form of the forcing term is $g(t) =$ a constant, we might guess that the equation $y'' + 4y' + 3y = 6$ has a solution of the form $y_p(t) = A$ for some value of A. Substituting $y_p(t) = A$ in for $y(t)$ in the equation $y'' + 4y' + 3y = 6$ gives

$$\begin{aligned} y_p'' + 4y_p' + 3y_p &= 6 \\ 0 + 4(0) + 3A &= 6 \end{aligned}$$

implying that $A = 2$ works. Therefore, the general solution to the nonhomogeneous equation is given by

$$y(t) = y_c(t) + y_p(t) = c_1 e^{-3t} + c_2 e^{-t} + 2.$$

Note that $y(t) \to 2$ as $t \to \infty$ and that 2 is the strength of the force (6 N) divided by the spring constant ($k = 3$). Exercise 1 below will show this is no coincidence.

Periodic forcing

Life gets more interesting with periodic forcing terms. As our warm-up example, consider the IVP

$$y'' + y = \sin 2t,\ y(0) = 0,\ y'(0) = 0.$$

We established last class that the complementary equation has general solution

$$y_c(t) = c_1 e^{it} + c_2 e^{-it}.$$

Based on previous examples with first order equations, we set

$$y_p(t) = A\sin 2t + B\cos 2t$$

as our trial solution to match the general family of functions in $g(t) = \sin 2t$. Plugging this in for y in $y'' + y = \sin 2t$ yields the equality

$$-4A\sin 2t - 4B\cos 2t + A\sin 2t + B\cos 2t = \sin 2t$$

so after equating coefficients of like terms, we obtain the system of equations

$$-3A = 1,\ -3B = 0.$$

Our particular solution is therefore

$$y_p(t) = -\frac{1}{3}\sin 2t$$

and our general solution is

$$y(t) = c_1 e^{it} + c_2 e^{-it} - \frac{1}{3}\sin 2t$$

At this point, we employ our initial conditions to obtain the system of equations

$$\begin{aligned} 0 &= c_1 + c_2 \\ 0 &= ic_1 - ic_2 - \frac{2}{3} \end{aligned}$$

which has solution $c_1 = -i/3, c_2 = i/3$. Now we plug this back into our general solution to yield

$$y(t) = -\frac{i}{3}(e^{it} - e^{-it}) - \frac{1}{3}\sin 2t.$$

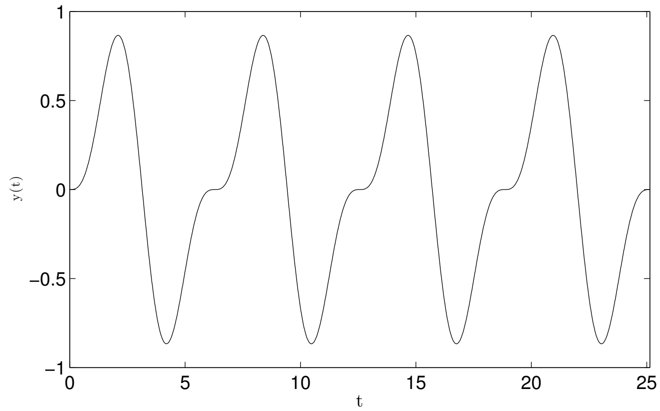
From Euler's trig identities, we know that

$$e^{it} - e^{-it} = 2i\sin t$$

and hence, the solution awesomely simplifies to

$$y(t) = \frac{1}{3}\left(2\sin t - \sin 2t\right).$$

This graph of this function is displayed in the figure below. Notice that it is periodic with period 2π and a maximum displacement of less than 1 (the amplitude of the driving force).



Alternating forces to resonance

With this preliminary example aside, we get to the payoff. Recall the behavior of the car in our spring-mass experiment with periodic forcing. When we applied driving forces with small or large periods, the car's motion was erratic and the amplitude of oscillations were small. For intermediate frequencies, the car exhibited smooth, constant amplitude oscillations and there was one "sweet spot" in the middle at which the amplitude of the oscillations would grow. It is our goal in this section to explain this behavior in an idealized, undamped (sometimes called *harmonic*) oscillator by studying the response amplitude as a function of the input frequency.

To this end, suppose that y satisfies the DE

$$my'' + ky = \sin(\omega t) \tag{2}$$

with initial conditions $y(0) = 0$, $y'(0) = 0$. Here ω is a parameter representing the driving frequency. To solve the SOCCNE in (2), we start with the complementary equation

$$my_c'' + ky_c = 0$$

which has characteristic equation

$$m\lambda^2 + k = 0$$

yielding two complex roots $\lambda = \pm i\sqrt{k/m}$. We call $\sqrt{k/m}$ the *natural (angular) frequency* of the system in radians [1] and denote it by ω_0. In the absence of forcing, it describes the natural frequency and period at which the mass will oscillate. The general solution to the complementary equation is

$$y_c(t) = c_1 e^{i\omega_0 t} + c_2 e^{-i\omega_0 t}.$$

Moving onto the particular solution, we take

$$y_p(t) = A\sin\omega t + B\cos\omega t$$

as our trial solution. Since

$$\begin{aligned} y_p'(t) &= \omega A\cos\omega t - \omega B\sin\omega t \\ y_p''(t) &= -\omega^2 A\sin\omega t - \omega^2 B\cos\omega t \end{aligned}$$

[1] To get the resonant frequency in hertz, divide by 2π. Here we won't worry about this conversion and just talk about resonant frequencies in radians.

plugging y_p in for y in the DE $my'' + ky = \sin \omega t$ yields the equality

$$-\omega^2 Am \sin \omega t - \omega^2 Bm \cos \omega t + Ak \sin \omega t + Bk \cos \omega t = \sin \omega t$$

After equating coefficients of like terms, we obtain the system of equations

$$\begin{aligned} A(k - m\omega^2) &= 1 \\ B(k - m\omega^2) &= 0 \end{aligned}$$

Therefore, $A = 1/(k - m\omega^2)$, $B = 0$, making our particular solution

$$y_p(t) = \frac{1}{k - m\omega^2} \sin \omega t$$

and our general solution

$$y(t) = c_1 e^{i\omega_0 t} + c_2 e^{-i\omega_0 t} + \frac{1}{k - m\omega^2} \sin \omega t$$

Finally, we plug in our initial conditions $y(0) = y'(0) = 0$ and simplify to obtain

$$y(t) = \frac{\omega}{\omega_0(k - m\omega^2)} \sin \omega_0 t + \frac{1}{k - m\omega^2} \sin \omega t$$

But wait! There are two values of ω for which the above procedure does not work: the values of ω for which

$$k - m\omega^2 = 0.$$

Solving for ω gives $\omega = \pm\sqrt{k/m}$, which match the natural frequencies of the system! To find a formula for the solution in this case, we rewrite y as

$$y(t) = \frac{(\omega/\omega_0) \sin \omega_0 t + \sin \omega t}{\omega_0(k - m\omega^2)}$$

and then let $\omega \to \omega_0$, using L'Hopitals Rule (taking the derivative with respect to ω) to calculate the limit:

$$\begin{aligned} y(t) &= \lim_{\omega \to \omega_0} \frac{(\omega/\omega_0) \sin \omega_0 t + \sin \omega t}{\omega_0(k - m\omega^2)} \\ &= \lim_{\omega \to \omega_0} \frac{t \cos \omega t}{(-2\omega_0 m\omega)} \\ &= \frac{t \cos \omega_0 t}{-2\omega_0^2 m} \\ &= -\frac{t}{2k} \cos \omega_0 t. \end{aligned}$$

This may have seemed insane, but look at the outcome: y is a cos function, but with an amplitude $t/2k$ that grows over time. This is exactly what we were hoping to find!

To review, if we force a harmonic oscillator at the harmonic frequency ω_0, the resulting solution will oscillate with increasing amplitude, just like in our experiment. The harmonic frequency is the "sweet spot" where we are driving the system at its natural rate, hereby magnifying the oscillations. For this reason, the harmonic frequency is sometimes referred to as the *resonant frequency.* Failing to take into account resonant frequencies in the design of structures can lead to disasters like the Tacoma Narrow bridge collapse or the swaying of the Millenium bridge (try googling YouTube videos of both for an entertaining study break). On the other hand, resonance can also be used to achieve positive results, for example, in designing seismographs which need to amplify vibrations.

Philosophizing aside, let's review what happened here.

(i) If we consider a harmonic oscillator of the form $my'' + ky = \sin(\omega t)$ where $\omega \neq \sqrt{k/m}$ (the natural frequency), we can apply MUC as before to determine a solution to the equation.

(ii) If we consider a harmonic oscillator of the form $my'' + ky = \sin(\omega t)$ where $\omega = \sqrt{k/m}$, then we need to solve the system first for $\omega \neq \omega_0$ and then let $\omega \to \omega_0$ at the end to obtain a solution.

There is a work around to the application of L'Hopitals above. If ω matches the natural frequency of the homogeneous system, we obtained a solution of the form

$$y(t) = y_c(t) + t(A\cos(\omega t) + B\sin(\omega t))$$

so we could just start out using $y_p(t) = t(A\cos(\omega t) + B\sin(\omega t))$ as our trial solution. This works (check it!) and can be either more or less tedious depending on whether you prefer using the product rule or L'Hopitals Rule. The important takeaway here is that we have mathematically verified resonance in a harmonic oscillator and furthermore, have a specific formula ($\sqrt{k/m}$) for computing the resonant frequency.

Frequency-Response Functions

Of course, the previous section doesn't quite explain what we saw in class during our spring-mass demo because that system involved a non-zero damping coefficient γ (as evidenced by the fact that left to its own devices, the car experienced damped oscillations). When damping is involved, *pure resonance* does not occur, but one can still encounter what we call *practical resonance*: the occurrence of local maxima in frequency-response functions.

Recall from Section 1.5 that frequency response functions show the gain in amplitude of the output/solution to a differential equation as a function of the input/driving frequency ω. Deriving formulas for frequency-response functions with second order equations is much simpler if we represent our periodic inputs as complex exponentials instead of sine or cosine functions. So instead of calculating steady state solutions to

$$my'' + \gamma y' + ky = a\sin\omega t$$

we calculate steady state solutions to

$$my'' + \gamma y' + ky = ae^{i\omega t}.$$

We know from the Principle of Superposition that the general solution to this system will be

$$y(t) = y_p(t) + y_c(t).$$

Furthermore, one can show (see exercises below) that $y_c(t) \to 0$ as $y \to \infty$ if $m, \gamma, k > 0$ so that it is sufficient to find $y_p(t)$. A trial solution to the SOCCNE will be

$$y_p(t) = Ae^{i\omega t}$$

and after plugging this in to solve for A, we get

$$A = \frac{a}{m(i\omega)^2 + \gamma(i\omega) + k}$$

therefore, the frequency-response function (FRF) will be given by

$$f(\omega) = \left|\frac{A}{a}\right| = \frac{1}{|m(i\omega)^2 + \gamma(i\omega) + k|}.$$

This is a really nice and easy to remember formula. Recalling that the characteristic equation for a SOCCHE is $m\lambda^2 + \gamma\lambda + k = 0$, you just need to plug in $i\omega$ for λ and take the absolute value of the reciprocal of the result. However, a word of caution: the denominator could very well be a complex number in which case "absolute value" means modulus. To illustrate, let's look at two examples: one with and one without damping. For the former, consider the system $y'' + y = 0$. The FRF is

$$f(\omega) = |(i\omega)^2 + 1)^{-1}| = 1/(1 - \omega^2)$$

which gets larger and larger as $\omega \to 1$, the resonant frequency. In fact, the resonant frequencies ± 1 are what we call poles of this function (places where the denominator is 0) which will always be the case with pure resonant frequencies (see exercises). For the latter, Consider a spring mass system with $m = 1, \gamma = 2, k = 1$. The FRF is now

$$\begin{aligned} f(\omega) &= \frac{1}{|(i\omega)^2 + 2i\omega + 1|} \\ &= \frac{1}{|(1-\omega^2) + (2\omega)i|} \\ &= \frac{1}{\sqrt{(1-\omega^2)^2 + (2\omega)^2}} \\ &= \frac{1}{\sqrt{(1-\omega^2)^2 + 4\omega^2}}. \end{aligned}$$

The graph of this function has a peak of 1 at $\omega = 0$ and decays as $\omega \to \pm\infty$. Therefore, $\omega = 0$ is what we call a practical resonant frequency because it is the value of forcing frequency which gives you "the most bang for your buck."

Another application of FRFs is to look at how the amplitude gain changes as one increases the damping coefficient γ. For example, suppose we consider the system $y'' + \gamma y' + y = 0$. When $\gamma = 0$, we get the pure resonant frequencies of ± 1. When $\gamma > 0$, then the FRF is

$$f(\omega)\frac{1}{\sqrt{(1-\omega^2)^2 + \gamma^2\omega^2}}.$$

When γ is small, the denominator will be almost 0 at the resonant frequencies so the practical resonant frequencies will be close to the pure resonant frequencies for the undamped system, but as we increase γ, we can see that the FRF gets smaller everywhere except at $\omega = 0$ so that we essentially start to universally kill off the systems' response to all nontrivial driving frequencies. This idea is used in civil and mechanical engineering to reduce the effects of resonance in structures. For example, washing machines use dashpots to damp the system when resonant spinning occurs. This is also why sitting on a washing machine can reduce resonant spinning when the dashpots fail!

Exercises

(1) Suppose that $m = 1, \gamma = 6, k = 5$ and $g(t) = A$ for some positive constant A in Equation (1). Determine the general solution $y(t)$ in terms of A. What happens to your solution as $t \to \infty$?

(2) Compute the solution to the IVP

$$y'' + y = \sin 2t,\ y(0) = 1,\ y'(0) = 0$$

and compare the solution with what we got in our warm-up example (same system with $y(0) = 0$).

(3) Compute the solution $y(t)$ to Equation (1) with $m = 1,\ \gamma = 0,\ k = 4,\ g(t) = \sin t$ and initial conditions $y(0) = y'(0) = 0$.

(4) A periodic external forcing of $g(t) = \sin(\omega t)$ Newtons is applied to an undamped spring-mass system with a mass of 4 kg and a spring constant of $k = 1$ kg/s^2. Suppose that the mass starts from rest with zero initial velocity and let $y(t)$ denote its displacement after t seconds.

(a) Determine a real valued expression for $y(t)$ if $\omega \neq 1/2$. Your final answer will have ω in it, but please use Matlab to validate that you get the right solution if you plug in $\omega = 1$.

(b) Create a plot comparing your solutions in (a) for $\omega = 0.51, 0.505, 0.501$. What appears to be happening as $\omega \to 1/2$?

(c) Calculate the frequency-response function for this system and verify that $\omega = 1/2$ is the only positive resonant frequency. Include a plot of the frequency-response function for $1/2 < \omega < 1$ and discuss the connection with your plots from part (b).

(d) Use L'Hopitals Rule to actually take the limit of your solution from (a) as $\omega \to 1/2$ and discuss how your answer is connected to your plots from (b).

(5) Calculate the frequency response function for a generic, harmonic oscillator with mass m and spring constant k. Show that the function has poles at $\pm$ the resonant frequency.

(6) Suppose we have a spring mass system with $m = 1, \gamma = 2, k = 5$. Determine the frequency (or frequencies) at which the frequency-response function achieves its maximum value.

(7) Which frequency or frequencies $\omega \geq 0$ will maximize the amplitude gain in the second order equation

$$y'' + 2y' + 4y = ae^{i\omega t}?$$

(8) Show that if $m, \gamma, k > 0$, then all solutions to $my'' + \gamma y + ky = 0$ converge to 0 as $t \to \infty$. Hint: Consider separately the three cases of underdamped, critically damped, and overdamped systems.

Section 2.5: RLC Circuits and Fourier Series

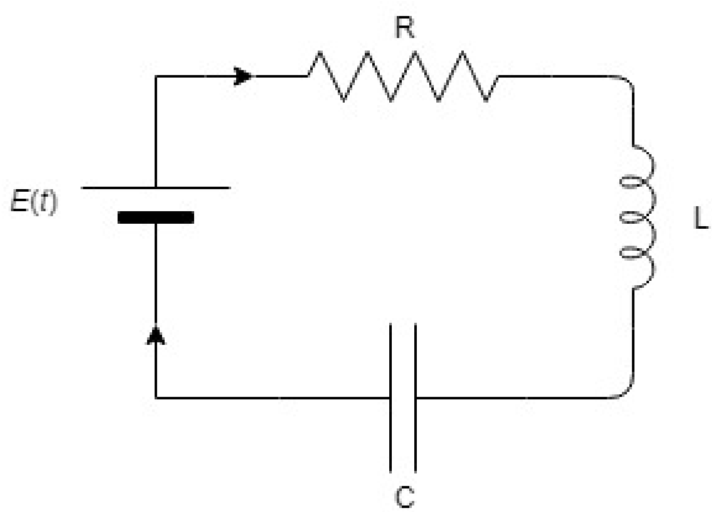
Section objectives:

- Explain analogies between RLC circuits and spring-mass systems
- Represent and integrate piecewise defined functions
- Compute Fourier coefficients and series approximations for periodic functions
- Approximate solutions to DEs with discontinuous periodic forcing functions using Fourier series

Note: At this point in the course, you have two paths. One is to take the red pill and continue reading this section. The second is to take the blue pill (more standard for differential equations courses) and skip to Section 2.6 for a coverage of Laplace transforms. Take your time. Hurry up. The choice is yours. Don't be late.

RLC Circuits

While we have spent most of this chapter thus far studying SOCCNE in the context of spring-mass systems, they also show up in the analysis of RLC circuits. An RLC circuit consists of a closed circuit with a resistor (R in the figure above), an inductor (L), and a capacitor (C) along with an electromotive driving force as shown in the following diagram:



As in the Section 1.5 notes, Kirchhoff's voltage law yields the equation

$$V_L + V_R + V_C = E(t)$$

where V_L, V_R and V_C are the voltage drops across the inductor, resistor, and capacitor, respectively, and $E(t)$ is the emf. If we let $q(t)$ denote the total charge on the capacitor at time t and $i(t) = q'(t)$ the corresponding current, then we know from the last notes that $V_R = iR$ and $V_C = q/C$ for some constants R, C. V_L is given by the following law:

The voltage drop across an inductor is proportional to the rate of change of the current.

Mathematically, this can be expressed as

$$V_L = L\frac{di}{dt}$$

for some positive constant L. Plugging all the voltage drops into our equation from Kirchoff's Law yields

$$Lq'' + Rq' + \frac{1}{C}q = E(t) \quad (1)$$

If we are given the initial charge $q(0) = q_0$ and initial current $i(0) = i_0$, this system is equivalent to the forced spring-mass system

$$my'' + \gamma y' + ky = g(t)$$

with initial conditions $y(0) = y_0$ and $y'(0) = v_0$. That's the power of math: two completely different applications, one ring to rule them all. You can use this connection to apply the intuition of mechanical spring-mass oscillators to electrical oscillators and calculate things like resonant frequencies or classify regimes of behavior (underdamped, etc) - some examples are explored in the exercises below. However, you can do so much more if you want to study the effects of discontinuous forcing functions.

Square and Sawtooth Waves

One particularly interesting example of a differential equation driven by a discontinuous periodic input is an LC circuit driven by a square wave as shown on the left in Figure 1 below. Notice that the ouput in this scenario (voltage across the capacitor) exhibits some wild oscillatory behavior even though the input is a piecewise constant function. A sawtooth input (shown on the right in Figure 1) is also an interesting phenomenon to behold.

To write out definitions of discontinuous functions such as the above inputs, we shall use the piecewise notation common to precalculus textbooks. For example, the notation

$$h(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

corresponds to a function which returns a value of 0 if the input t is less than 0 and returns 1 if the input t is greater than or equal to 0. A graph of this function, often referred to as the Heaviside function, is shown on the left side of Figure 2. We will focus our attentions in this section on periodic functions, that is those which repeat themselves after some amount of minimal amount of time p which we call the period of the function. For example, a standard sinusoidal function has period 2π while the function $\sin(2\pi t)$ has period 1. For periodic, piecewise continuous functions, such as the square wave, we usually just give the function's period along with a description of the function over one period. For example, a standard square wave function alternates back and forth between values of 1, over the first half of

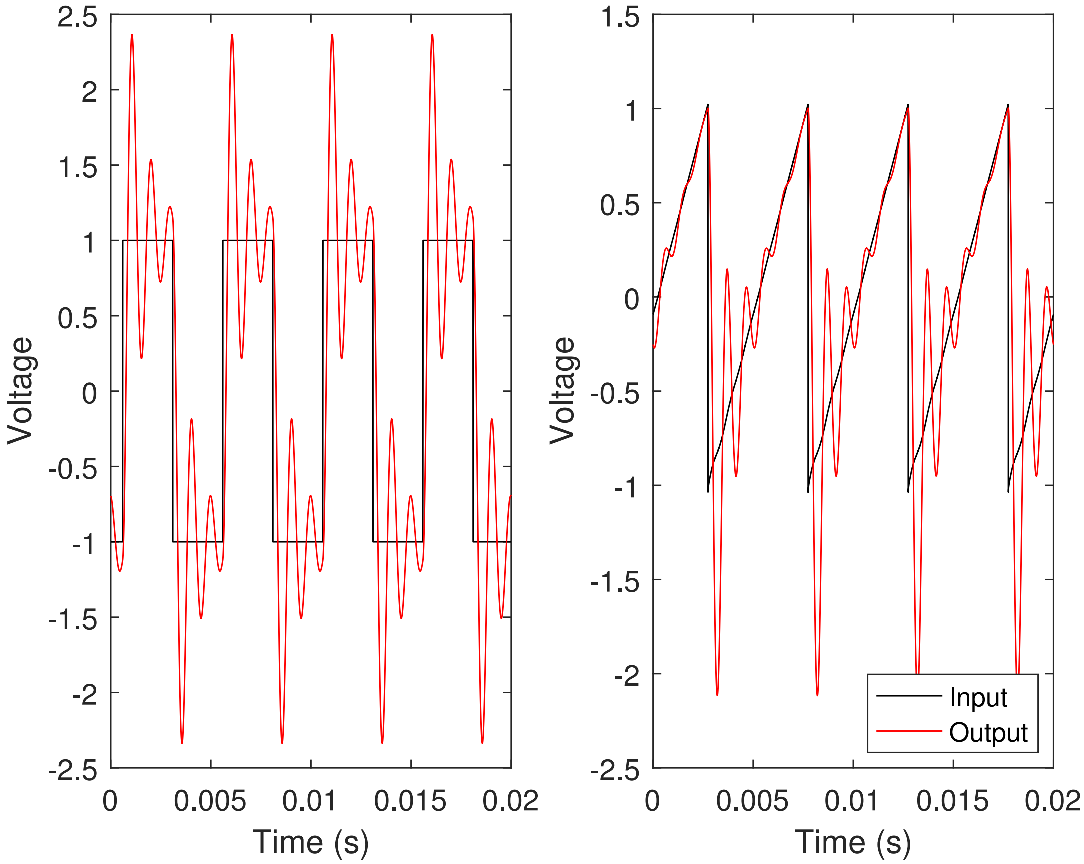


Figure 1: Voltage across the capacitor in an LC circuit with square wave emf (left) and sawtooth wave emf (right), each with a frequency of approximately 200 hz.

each period, and -1 over the second half of each period. In piecewise notation, we can write this as

$$s(t) = \begin{cases} 1 & 0 \leq t < p/2 \\ -1 & p/2 \leq t < p \end{cases}.$$

A *sawtooth wave* ramps up from a minimum of -1 at the start of one period to a maximum of 1 at the end of a period. If the period is p, this means that the slope is $2/p$ so in this case we would just have

$$w(t) = \frac{2}{p}t - 1 \tag{2}$$

over the interval $0 \leq t < p$.

There is confusion on all fronts when you read about the various definitions of the sawtooth and square wave functions because different websites and textbooks like to start and end their periods at different places. For example, some choose periods which are symmetric around 0 as a baseline. We will stick here with $[0, p)$ as the baseline and consider any other example of a square, sawtooth, or other wave as a phase shifted version of this standard. You can also define the endpoints of intervals in different ways, but at the end of the day it doesn't really matter. We will try to stick with the convention of using left-closed/right-open intervals, but make no promises. The right side of Figure 2 demonstrates these conventions by showing one period of a period-one square wave function.

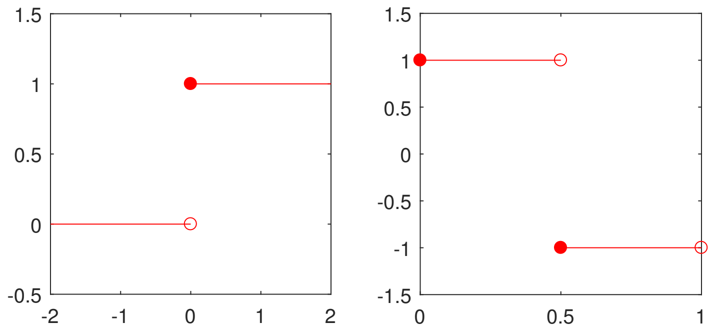


Figure 2: Graph of the Heaviside function over the interval $[-2, 2]$ (left) and one period of the period-1 square wave (right)

Although you probably learned this in Calc 1 or 2, its probably a good idea at this point to review how to integrate piecewise functions as well. The key is to break up integrals based on the different intervals in the definition of the piecewise function and substitute in the appropriate values. For example, if $s(t)$ is the 2π periodic square wave

$$s(t) = \begin{cases} 1 & 0 \leq t < \pi \\ -1 & \pi \leq t < 2\pi \end{cases}$$

then

$$\begin{aligned} \int_0^{2\pi} s(t)\,dt &= \int_0^{\pi} s(t)\,dt + \int_{\pi}^{2\pi} s(t)\,dt \\ &= \int_0^{\pi} 1\,dt + \int_{\pi}^{2\pi} (-1)\,dt \\ &= \pi - \pi = 0. \end{aligned}$$

Of course in this example, the result should have been obvious from the graph as well!

To solve systems of equations with discontinuous forcing, we have to take a quick detour into the beautiful world of Fourier Series and trigonmetric polynomials. To complement the discussion in the next section, we also recommend you check out one (or more) of the following links for some nice visualizations:

- 3Blue1Brown's video on Fourier Series in relation to the heat equation: `https://www.youtube.com/watch?v=r6sGWTCMz2k`

- Smarter Every Day's video on a Visual Introduction to Fourier Series: `https://www.youtube.com/watch?v=ds0cmAV-Yek`
- Jez Swanson's interactive introduction website: `http://www.jezzamon.com/fourier/index.html`

Fourier Series

We have already seen how a sum of a sin and cos wave of the same frequency can be combined into a single phase-shifted sin function, but more interesting things can happen when you combine multiple sin functions with different frequencies. One example (based off Exercise 3 at the end of Matlab Supplement 1) is shown below where we graph the four functions

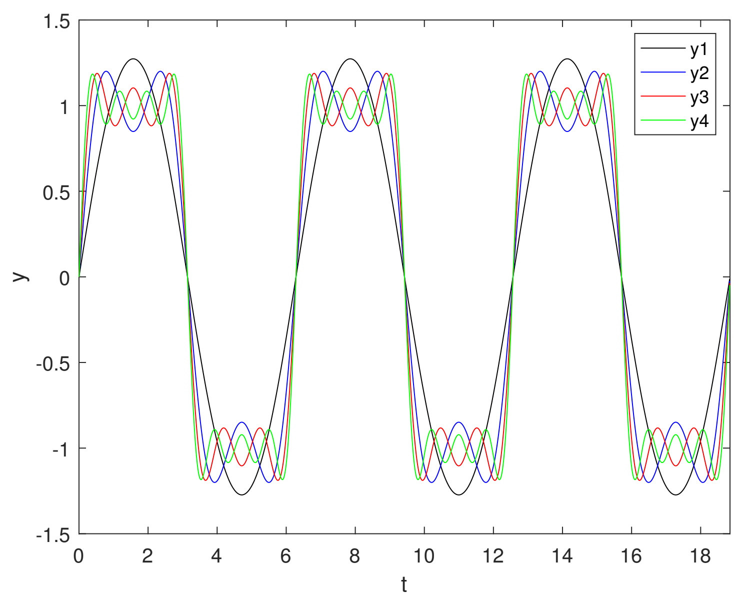
(a) $y_1(t) = (4/\pi)\sin(t)$

(b) $y_2(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t)$

(c) $y_3(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t) + (4/(5\pi))\sin(5t)$

(d) $y_4(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t) + (4/(5\pi))\sin(5t) + (4/(7\pi))\sin(7t)$

Notice that the lowest frequency term y_1 provides an underlying trend for oscillations, but higher frequency terms add additional peaks and troughs on top of the main wave. Interestingly, the graph starts to look like a period-2π square wave, doesn't it? The reason for this sort of behavior is due to a powerful idea developed by Fourier in the 1800s: periodic function can be approximated[1] by sums of sinusoidal functions!



[1]At least under suitable integrability conditions and with a suitable definition of "approximate". That line of reasoning is beyond the scope of this text though.

To develop this idea, recall from Euler's formula that the complex exponential provides a clean way of writing sinusoidal functions. So let's define a trigonometric polynomial with period 2π and degree N as any function of the form

$$f_N(t) = \sum_{n=-N}^{N} a_n e^{int} = a_0 + a_1 e^{it} + a_{-1}e^{-it} + \cdots + a_N e^{iNt} + a_{-N}e^{-iNt}.$$

The functions e^{int} are called harmonics with e^{it} representing the first/fundamental harmonic, the term e^{2it} representing the second harmonic, and so on. Recall that the frequency of e^{int} is $n/2\pi$ so the nth harmonic has a frequency of n times the frequency of the fundamental harmonic. The term $a_0 = a_0 e^{0it}$ is called the constant or zeroth order term

As an example of a trig polynomial, consider the function

$$y_2(t) = (4/\pi)\sin(t) + (4/(3\pi))\sin(3t)$$

from above. Recalling that $\sin y = (e^{iy} - e^{-iy})/(2i)$ from Euler's Formula, we can rewrite y_2 as

$$\frac{2}{\pi i}(e^{it} - e^{-it}) + \frac{2}{3\pi i}(e^{i3t} - e^{-i3t}).$$

Therefore, y_2 is a third order trig polynomial with coefficients

$$\begin{aligned} a_0 &= 0 \\ a_1 &= \frac{2}{\pi i} \\ a_{-1} &= -\frac{2}{\pi i} \\ a_2 &= 0 \\ a_{-2} &= 0 \\ a_3 &= \frac{2}{3\pi i} \\ a_{-3} &= -\frac{2}{3\pi i} \end{aligned}$$

and $a_n = 0$ for all other n. Therefore, y_2 is a third degree trig polynomial.

But of course many periodic functions, like the square or sawtooth wave, are not trigonometric polynomials. In such situations, the goal is to find a trig polynomial which provides an approximation. More specifically, given a periodic function f with period 2π, we would like to be able to find coefficients a_n such that

$$f(t) \approx f_N(t) = \sum_{n=-N}^{N} a_n e^{int} \tag{3}$$

for sufficiently large N. If such an approximation is possible, then we call the righthand side of Equation (3) an Nth order *Fourier series* approximation of f. The a_n are called the corresponding *Fourier coefficients.*

Finding Fourier series approximations (like finding Taylor polynomial approximations) boils down to finding formulas for the coefficients on the right side of Equation (3). To find such formulas, let's start by considering the constant term a_0. Intuitively, it seems like this number should represent some sort of average value for f over $[0, 2\pi)$. (The other terms oscillate so we need to choose the right midline). You may recall from calculus that the average value of a function g over an interval $[a, b]$ is given by

$$\frac{1}{b-a} \int_a^b g(t)\, dt.$$

so let's conjecture that

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, dt.$$

To confirm, we can integrate both sides of Equation (3) and use the linearity of integrals to obtain

$$\int_0^{2\pi} f(t)\ dt \approx \sum_{n=-N}^{N} \int_0^{2\pi} a_n e^{int}\, dt.$$

In an exercise from Section 2.3, you showed that

$$\int_0^{2\pi} e^{ikt}\, dt = 0 \tag{4}$$

for any $k \neq 0$ (if you skipped this exercise, go back and try it now). Therefore, the integrals above reduce to

$$\int_0^{2\pi} f(t)\ dt \approx \int_0^{2\pi} a_0\, dt = 2\pi a_0$$

and solving for a_0 confirms our guess.

The next step is to get a formula for the coefficients a_n with $n \neq 0$. Thinking about what happened with the constant term, we need to multiply f by something which will kill off all the other terms a_m with $m \neq n$ after integration, but keep a_n in tact. In other words, we want to shift the exponents so that a_n is now the constant term. To this end, let's try multiplying each term in Equation (3) by e^{-int}, the inverse of e^{int}. This results in the series

$$f(t)e^{-int} \approx \sum_{m=-N}^{N} a_m e^{imt} e^{-int} = \sum_{m=-N}^{N} a_m e^{i(m-n)t}$$

Notice that the term with $m = n$ now has an exponent of 0 while all the other terms have nonzero exponents $k = m - n$. Therefore, when we integrate and apply Equation (4), we will be left with

$$\int_0^{2\pi} f(t)e^{-int}\, dt \approx \sum_{m=-N}^{N} a_m \int_0^{2\pi} e^{i(m-n)t}\, dt = a_n 2\pi$$

Dividing both sides by 2π gives us a nifty formula for the nth Fourier coefficient

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(t)e^{-int}\, dt \tag{5}$$

that is consistent with our earlier formula for a_0.

As an example of how to work with this formula, let's compute the Fourier coefficients for the square wave function $s(t)$ with period $p = 2\pi$. This function is 1 over the interval $[0, \pi)$ and -1 over the interval $[\pi, 2\pi)$ so for $n = 0$, we get the coefficient

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} s(t)\,dt = 0$$

as shown earlier in this section. For $n = 1$, we get

$$\begin{aligned} a_1 &= \frac{1}{2\pi}\int_0^{2\pi} s(t)e^{-it}\,dt \\ &= \frac{1}{2\pi}\left(\int_0^{\pi}(1)e^{-it}\,dt + \int_{\pi}^{2\pi}(-1)e^{-it}\,dt\right) \\ &= -\frac{1}{i2\pi}\left((e^{-i\pi} - 1) - (e^{-i2\pi} - e^{-i\pi})\right) \\ &= \frac{2}{\pi i} \end{aligned}$$

where in the last line we used the facts that $e^{-i\pi} = -1, e^{-i2\pi} = -1$. We could keep going in this manner and compute a_{-1}, a_2 and so on, but it is more efficient to instead see if we compute a general formula for a_n for $n \neq 0$. We start with the formula in Equation (5) and break it up into two integrals based on the definition of $s(t)$.

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} s(t)e^{-int}\,dt = \frac{1}{2\pi}\left[\int_0^{\pi}(1)e^{-int}\,dt + \int_{\pi}^{2\pi}(-1)e^{-int}\,dt\right].$$

At this point we have to look at two cases: when n is odd and when n is even. When n is even,

$$e^{-in\pi} = \cos(in\pi) - i\sin(in\pi) = 1$$

so

$$\int_0^{\pi} e^{-int}\,dt = -\frac{1}{in}\left(e^{-in\pi} - 1\right) = 0$$

and

$$\int_{\pi}^{2\pi} e^{-int}\,dt = -\frac{1}{in}\left(e^{-2\pi ni} - e^{-in\pi}\right) = 0.$$

Now isn't that special! If n is odd, I'll leave it to you to show that $e^{-in\pi} = -1$ so

$$\int_0^{\pi} e^{-int}\,dt = -\frac{1}{in}\left(e^{-in\pi} - 1\right) = \frac{2}{in}$$

and

$$\int_{\pi}^{2\pi} e^{-int}\,dt = -\frac{1}{in}\left(e^{-2\pi ni} - e^{-in\pi}\right) = -\frac{2}{in}$$

so

$$a_n = \frac{1}{2\pi}\left[\frac{2}{in} + \frac{2}{in}\right] = \frac{2}{in\pi}.$$

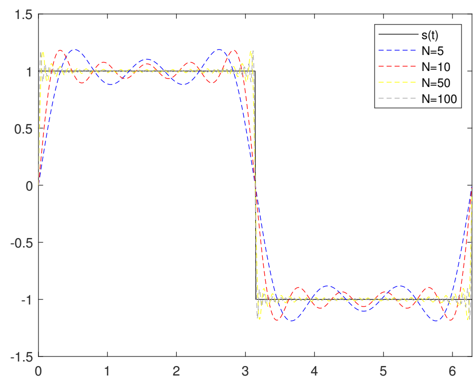
We therefore arrive at the following formula for the Fourier coefficients of the square wave

$$a_n = \begin{cases} 0, & \text{if } n \text{ is even} \\ \frac{2}{in\pi}, & \text{if } n \text{ is odd.} \end{cases}$$

Recall that a_n is the coefficient of the nth term in the Fourier series approximation of s so our work thus far implies that

$$s(t) \approx s_N(t) = \sum_{n=-N}^{N} a_n e^{int} = \sum_{|n|\leq N, n \text{ odd}} \frac{2}{in\pi}\left(e^{int} - e^{-int}\right)$$

for sufficiently large N. The figure below validates this calculation by comparing $s_N(t)$ with $s(t)$ for increasing values of $N = 5, 10, 50, 100$. Even though there is always some funkiness near the discontinuity, the values of $s_N(t)$ do seem to be converging to $s(t)$ elsewhere in the interval as we increase N (Google Gibbs Phenomenon for more on this).



Before getting back to DEs, we also want to note that this and other complex exponential Fourier series approximation can be rewritten in terms of sines and cosines via Euler's trig identities. Some expositions on Fourier series start with expansions of this form and ignore complex exponentials, but the formulas get a bit messier because you need to calculate both sin and cos coefficients. In the case of the square wave, however, the situation is simplified by the lack of even terms. When the smoke clears, we end up with the following expansion that depends only on sin's:

$$s(t) \approx \sum_{n=1,3,\ldots,N} \frac{4}{n\pi} \sin(nt).$$

Bringing this discussion full circle, note that plugging in $N = 1, 3, 5, 7$ yields the functions $y_1, \ldots y_4$ graphed at the beginning of this section (surprise, those weren't just arbitrary functions chosen to torture you).

Solving Differential Equations with Periodic forcing

Now that we have a method for approximating periodic functions with trig polynomials, our next step is to find approximate solutions to differential equations with periodic forcing. We shall illustrate with a square wave forcing term although the exercises will explore other examples. We are also going to focus on finding a particular solution and not worry about tuning our solution to meet some specific initial conditions although the PSI in this latter situation can be done with some time and patience (which I don't have right now).

Whether we use $R, L, 1/C$ or m, γ, k doesn't make much difference mathematically, but since we started this section talking about RLC circuits, we'll stick with that notation below. So the DE we will be solving is

$$Lq'' + Rq' + \frac{1}{C}q = s(t) \tag{6}$$

The idea is to rewrite $s(t)$ as a sum of complex exponential like in the previous section:

$$s(t) = \sum a_n e^{int}$$

and then find particular solutions to each of the simpler equations

$$Lq_n'' + Rq_n' + \frac{1}{C}q_n = a_n e^{int}. \tag{7}$$

The principle of superposition tells us that we can then add these solutions together to obtain an approximate solution

$$q_p(t) = \sum_{n=-N}^{N} q_n(t)$$

to Equation (6). But solving Equation (7) is easy. Plugging in the trial solution $q_n(t) = A_n e^{int}$ and solving for A_n yields

$$A_n\left(-Ln^2 + iRn\frac{1}{C}\right)e^{int} = a_n e^{int} \Longrightarrow A_n = \frac{a_n}{1/C - Ln^2 + iRn}.$$

Therefore, the particular solution[2] to Equation (6) is

$$q_p(t) \approx \sum_{n=-N}^{N} \frac{a_n}{1/C - Ln^2 + iRn} e^{int}.$$

From the last section, we know that

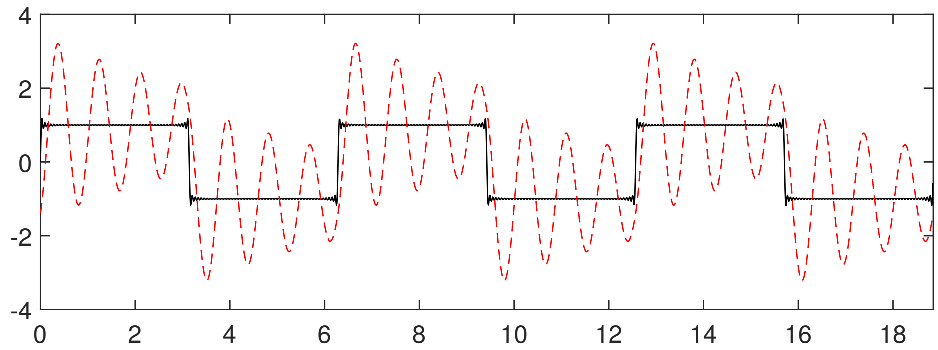
$$a_n = \frac{2}{in\pi}$$

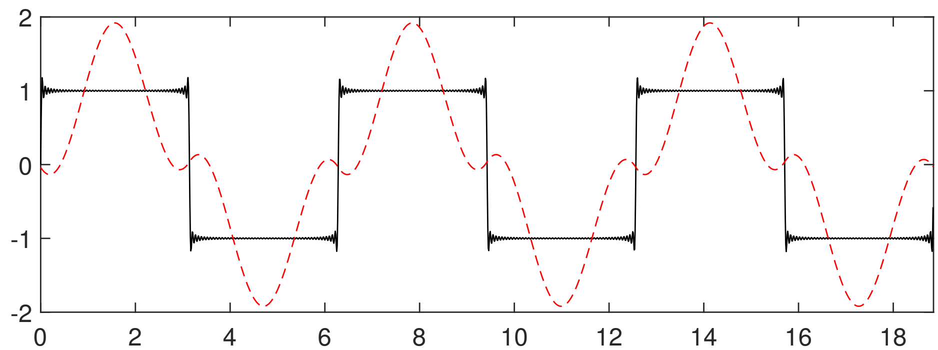
for n odd and $a_n = 0$ for n even so after again applying Euler's trig identities, the result "simplifies" to

$$\begin{aligned} q_p(t) \quad &\approx \quad \sum_{n=1,3,\ldots,N} \left[\frac{2}{in\pi(1/C - Ln^2 + iRn)}e^{int} - \frac{2}{in\pi(1/C - Ln^2 + iRn)}e^{-int}\right] \\ &= \quad \sum_{n=1,3,\ldots,N} \frac{4}{n\pi(1/C - Ln^2 + iRn)}\sin(nt). \end{aligned}$$

Two examples of such solutions are displayed in the graph below. In both examples, $R = 0.01, 1/C = 1.1$ and $N = 100$. The difference is in the choice of L which is set to 0.021 in the top figure and $L = 0.21$ in the bottom. Since the natural frequency of a pure LC circuit is $\sqrt{1/(LC)}$, the top figure oscillates more rapidly and more closely resembles the experiment shown at the beginning of the section.

[2]Note there is a problem if $R = 0$ and $1/C - Ln^2 \neq 0$ for any n. In this case, one term in the fourier series is hitting the resonant frequency for the systems and the series fails to converge. In practice, this shouldn't happen, but in simulations, you need to be careful avoiding such choices.





Using Matlab to find the Fourier Coefficients

For functions more complicated than a square wave, calculating the integrals in (5) by hand can be at best a frustrating exercise in Calc 2 and at worst, an impossible task. Instead, one can use numerical integration to calculate these coefficients. This is a topic for Math 110 (Numerical Analysis) so we won't get into the details, but Matlab's "integral" command can calculate definite integrals of function handles over predefined interval using a technique called Adaptive Quadrature. For example, if we wanted to perform the integral

$$\int_0^1 x^2\,dx$$

the code would be

```
integral(@(x) x.^2,0,1)
```

The lines below show how to compute the Nth order Fourier series approximation of the standard square wave function using the integral command:

```
N=50;                            % Number of terms to use
f=@(t) 1*(t<pi)+(-1)*(t>pi);     % Function - standard square wave here

t=0:0.01:(2*pi);                 % Time interval to view solution
c0=integral(f,0,2*pi)/(2*pi);    % Constant term in Fourier Series
```

```
fN=c0;                         % Start by adding constant .
for n=1:N
    % Compute coefficient for exp(int)
    cp =integral(@(t) f(t).*exp(-n*1i*t),0,2*pi)/(2*pi);
    % Compute coefficient for exp(-int)
    cn =integral(@(t) f(t).*exp(n*1i*t),0,2*pi)/(2*pi);
    % Add the new terms to the Fourier series
    fN=fN+(cp*exp(n*1i*t)+cn*exp(-n*1i*t));
end

plot(t,f(t),'k',t,fN,'r--')
xlabel('t')
xlim([0,2*pi])
legend('Actual',['Fourier, N=',num2str(N)])
```

Feel free to use modified versions of this code as needed for the exercises below!

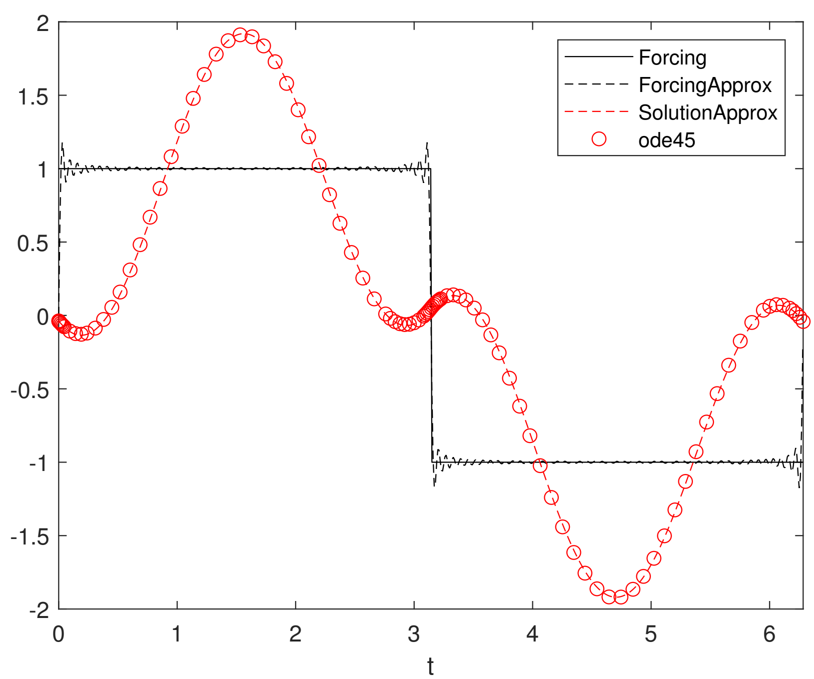
Using Matlab to validate solutions

Of course like anything we do, it is helpful to validate our rather nasty hand calculations with ode45. So to conclude this function, we'll include code and plot which will (i) compute the Fourier series approximation of a forcing function f, (ii) compare the approximation with f, (iii) compute the Fourier series approximation of the solution y to a given spring-mass system (or circuit), and (iv) validate that solution with ode45. (Phew)

```
% Define paramaters
m=0.21;
gamma=0.01;
k=1.1;

% Define time interval, forcing function, N
N=100;
t=0:0.01:(2*pi);
f=@(t) 1*(t<pi)+(-1)*(t>pi);

% Compute constant terms and initialize series
c0=integral(f,0,2*pi)/(2*pi);
a0=c0/k;
fN=c0;
yN=a0;

% Compute other terms
for n=1:N
    cp=integral(@(t) f(t).*exp(-n*1i*t),0,2*pi)/(2*pi);
    cn=integral(@(t) f(t).*exp(n*1i*t),0,2*pi)/(2*pi);
    ap=cp/(-m*n^2+gamma*n*1i+k);
    an=cn/(-m*n^2-gamma*n*1i+k);
```

```
    fN=fN+cp*exp(n*1i*t)+cn*exp(-n*1i*t);
    yN=yN+ap*exp(n*1i*t)+an*exp(-n*1i*t);
end

% Now we solve with ode45.
% Need to use initial value and slope of yN as ICs
y0 = [yN(1),(yN(2)-yN(1))/0.01];
[t2,y]=ode45(@(t,y) [y(2);f(t)/m-(k/m)*y(1)-(gamma/m)*y(2)],[0,2*pi],y0);

% Final Compare
plot(t,f(t),'k',t,fN,'k--',t,yN,'r--',t2,y(:,1),'ro')
xlabel('t')
xlim([0,2*pi])
legend('Forcing','ForcingApprox','SolutionApprox','ode45')
```

Worth it though for this beauty:



Closing Remarks

This section was probably a lot to take in, but if you made it through, congratulations! You now have a basic understanding of Fourier analysis, an advanced mathematical topic that may just be one of the most important concepts covered thus far in the course. The rest of this chapter can theoretically be skipped without any loss of continuity, but we will include it here to break down a topic that provides an alternative perspective on solving second order equations and one you may very well encounter in advanced engineering coursework. The topic I speak of is the infamous Laplace Transform.

Exercises

(1) Find a formula for the resonant frequency of a pure LC circuit $Lq'' + \frac{1}{C}q = 0$ (no resistor).

(2) Assume that there is no initial charge or current across the capacitor in an LC circuit with $L = 1, C = 1/4$ and $E(t) = \sin(\omega t)$. Determine a formula for the charge after t seconds in terms of ω and calculate the limit as $\omega \to 2$.

(3) Find conditions on L, R, C in an RLC circuit which guarantee that the corresponding characteristic equation will have

   (a) Two real roots.
   (b) One real root
   (c) Two complex roots.
   (d) Two purely imaginary roots (no real part)

   For each scenario, describe the asymptotic behavior of $q(t)$ as $t \to \infty$.

(4) Suppose that the charge, $q(t)$, across a capacitor (in kilocoulombs) after t seconds in a series LC circuit with alternating current source, satisfies a second order equation of the form

$$\begin{aligned} q'' + 9q &= 5\cos(2t) \\ q(0) &= 0 \\ q'(0) &= v_0 \end{aligned}$$

   where $v_0 > 0$.

   (a) Determine a formula for $q(t)$. The answer will depend on v_0.
   (b) Create a Matlab plot which compares $q(t)$ for the three choices of $v_0 = 0, 1, 2$. Discuss how the behavior of the solution depends of v_0.

(5) Define the function

$$f(t) = \begin{cases} 0 & 0 \le t < \pi/2 \\ 1 & \pi/2 \le t < 3\pi/2 \\ 0 & 3\pi/2 \le t < 2\pi \end{cases}.$$

   (a) Calculate the first three Fourier coefficients a_0, a_1, a_{-1}.
   (b) Sketch a plot of both f and its first order trig approximation over the interval $[-2\pi, 2\pi]$.

(6) Calculate all Fourier coefficients for a modified square wave of period 2π whose values over the interval $[0, 2\pi)$ are given by

$$\tilde{s}(t) = \begin{cases} 1 & 0 \le t < \pi \\ 0 & \pi \le t < 2\pi \end{cases}.$$

   Sketch a plot of the first, fifth, and 25th order Fourier series approximations along with $\tilde{s}(t)$ to validate your calculations.

(7) Determine a formula for the Nth order approximation of a particular solution to the differential equation

$$my'' + ky = \tilde{s}(t)$$

where $\tilde{s}(t)$ is the modified square wave function from the last exercise. Create several Matlab plots showing your solution alongside $\tilde{s}$ with different values of k, m and $N = 100$. Discuss how m, k affect the solution and which values (if any) lead to problems.

(8) Calculate and plot an approximate solution to the second order equation

$$0.025y'' + 0.01y' + 1.1y = w(t)$$

where $w(t)$ is a sawtooth wave with period 2π (see Equation (2) from earlier in this section). Use at last 100 terms in your approximation and include a plot which compares your solution to $w(t)$.

Section 2.6 Notes :
Introduction to Laplace Transforms

Section objectives:

- Calculate Laplace transforms of basic functions (sin, cos, exp, polynomials) using the definition, linearity, and exponential shifts.
- Determine Inverse Laplace Transforms using algebraic manipulations and comparisons with known transforms.

Motivation

Although most differential equations which we will solve using the method of *Laplace transforms* can also be solved using the methods of Section 2.2-2.5, we are going to discuss it anyways for three reasons. First, you will use the Laplace transform and its cousin the Fourier transform in many upper division engineering, math, and physics courses (PDEs, systems, signals, and circuits to name a few). It will help to get a brief introduction to the underlying concepts now. Second, once you get used to the main idea behind Laplace transforms, they offer a quicker and elegant alternative to our previous methods, especially if you want to avoid "complex" arithmetic (pun intended). As you progress into more complicated examples of interconnected masses and circuits, solutions to problems are much easier to work with and formulate if you work in Laplace land[1]. Finally, and perhaps most importantly of all, we have hopefully convinced you at this point that there is something truly magical about exponential functions. One can get all sorts of behavior out of the function $f(z) = e^z$ depending on whether z is complex, pure imaginary, positive, negative, or zero. Laplace transforms will essentially give us a method for breaking down potentially non-exponential functions of time into exponential components and describing them in terms of what exponential functions they "most closely represent".

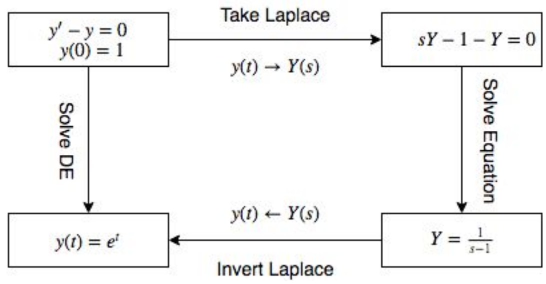
Before we define the Laplace transform, here is some further perspective on *WHY* it can make the study of differential equations so much simpler. The key idea is to take a differential equation with independent variable t and convert it to an algebraic equation in a new variable s. We then solve the algebraic equation and convert back to the t domain to determine the solution to our original equation of interest. This process is diagrammed in the figure below. We will start with some basic definitions and examples that on the surface, have nothing to do with differential equations.

Definition of Laplace transforms and some basic examples

We start by defining the **Laplace transform**, $\mathcal{L}\{f\}(s)$, of a function $f(t)$ as the value of the improper integral

$$\mathcal{L}\{f\}(s) = \int_0^\infty e^{-st} f(t)dt = \lim_{A\to\infty} \int_0^A e^{-st} f(t)dt. \tag{1}$$

[1]For example, in circuits, one can generalize Ohm's Law $V = IR$ to impedance relationships $V(s) = Z(s)I(s)$ in terms of Laplace transforms.



We will sometimes use the uppercase letter notation $F(s) = \mathcal{L}\{f\}(s)$ to denote the Laplace transform of the function f. It is important to remember that the Laplace transform is a "function of functions": it takes a real valued function f (with independent variable t) as input and returns a different function F (with independent variable s) as output. Basically, one can think of the Laplace transform as a portal through the multiverse of functions: it takes us from a world where inputs are represented by t (sometimes called the "t-domain" or our world) to a world where inputs are represented by s (the "s-domain" or Laplace land).

Technically, the s-domain is further restricted to the set of all values of s for which the improper integral in (1) converges[2]. To illustrate the use of the definition of the Laplace transform and the role that s plays in the calculation, we will calculate the Laplace transform of $f(t) = 1$:

$$\begin{aligned}\mathcal{L}\{1\}(s) &= \lim_{A\to\infty} \int_0^A 1e^{-st}\,dt \\ &= \lim_{A\to\infty} -\frac{e^{-st}}{s}\Big|_0^A \\ &= \lim_{A\to\infty} \frac{1-e^{-sA}}{s}.\end{aligned}$$

One of the common mistakes newcomers to the Laplace transform make is to perform the integral in the second line above with respect to s instead of t. But s is a constant as far as the integral is concerned. It is only now, as we evaluate

$$\lim_{A\to\infty} \frac{1-e^{-sA}}{s} = \begin{cases} \frac{1}{s} & s > 0 \\ \infty & s \leq 0 \end{cases}$$

that the role of s as a variable comes into play. Since the limit above is finite if and only if $s > 0$, the domain of the $\mathcal{L}\{1\}(s)$ is $s > 0$ and for any $s > 0$, we have

$$\mathcal{L}\{1\}(s) = \frac{1}{s}.$$

Exercises 1 below will ask you to verify the following Laplace Transforms as well:

(a) $\mathcal{L}\{t\}(s) = \frac{1}{s^2}$ with domain $s > 0$.

(b) $\mathcal{L}\{e^{(a+ib)t}\}(s) = \frac{1}{s-(a+ib)}$ with domain $s > a$

[2]For now, s can be taken to be a real number, but pay attention to when and where the results will also make sense of complex s.

The formula in (b) is the one we will use the most and in some ways, encapsulates everything important about Laplace transforms. First, it contains many other LT's as special cases. For example, if we plug in $b = 0$, we get

$$\mathcal{L}\{e^{at}\}(s) = 1/(s-a).$$

If we further set $a = 0$, we get out previous result that $\mathcal{L}\{1\}(s) = 1/s$. The next section will show how it can also be used to find the Laplace transform of sin and cos.

Linearity of the Laplace transform

Linearity is a reoccurring theme in this course (recall the Principle of Superposition) and it is linearity which makes the Laplace transform nice to work with. The linearity property of Laplace transforms can be stated as follows: if $f_1(t)$ and $f_2(t)$ are two functions and c_1, c_2 are two scalars then

$$\mathcal{L}\{c_1 f_1 + c_2 f_2\}(s) = c_1\mathcal{L}\{f_1\}(s) + c_2\mathcal{L}\{f_2\}(s) \tag{2}$$

for any value of s in the domain of $\mathcal{L}\{f_1\}$ and $\mathcal{L}\{f_2\}$. To see why this holds, we start with two arbitrary functions $f_1(t)$ and $f_2(t)$ and two arbitrary constants c_1 and c_2 and let s be in the domains of f_1 and f_2. As usual in these type of proofs, our goal is to start with the expression on the left side of the equality in equation (2) and manipulate it to obtain the expression on the right side. Now the definition of the Laplace transform implies that

$$\mathcal{L}\{c_1 f_1 + c_2 f_2\}(s) = \int_0^\infty (c_1 f_1(t) + c_2 f_2(t))e^{-st}\,dt.$$

However, the integral itself is a linear operator: you are well acquainted with the fact that

$$\int (2t + t^2)\,dt = 2\int t\,dt + \int t^2\,dt$$

for example. Therefore, after distributing the e^{-st}, we can break up the integral[3] above as

$$\begin{aligned}\int_0^\infty (c_1 f_1(t) + c_2 f_2(t))e^{-st}\,dt &= \left(c_1\int_0^\infty f_1(t)e^{-st}\,dt + c_2\int_0^\infty f_2(t)e^{-st}\,dt\right)\\ &= c_1\mathcal{L}\{f_1\}(s) + c_2\mathcal{L}\{f_2\}(s)\end{aligned}$$

by definition of $\mathcal{L}\{f_1\}(s)$ and $\mathcal{L}\{f_2\}(s)$. This final expression was the one appearing on the right side of equation (2) so we have reached our goal!

Now that you (hopefully!) understand why the linearity property holds, we can see how it can be used as a time saving device. Let start by calculating the Laplace transform of $f(t) = 2 - 3t + 5e^{2t}$. We know from the previous section that

$$\begin{aligned}\mathcal{L}\{1\}(s) &= \frac{1}{s}\\ \mathcal{L}\{t\}(s) &= \frac{1}{s^2}\\ \mathcal{L}\{e^{2t}\}(s) &= \frac{1}{s-2}\end{aligned}$$

[3]Assuming that all integrals converge which is true as long as s is in the domain of all Laplace transforms

so the linearity property implies that

$$\mathcal{L}\{f\}(s) = 2\mathcal{L}\{1\}(s) - 3\mathcal{L}\{t\}(s) + 5\mathcal{L}\{e^{2t}\}(s) = \frac{2}{s} - \frac{3}{s^2} + \frac{5}{s-2}.$$

We can also use the linearity property to calculate the Laplace transform of $e^{at}\sin bt$ and $e^{at}\cos bt$ (To further appreciate Linearity, try calculating those Laplace transforms directly from the definition!) The trick is to recall that

$$e^{(a+ib)t} = e^{at}\cos bt + ie^{at}\sin bt.$$

Therefore linearity implies that

$$\mathcal{L}\{e^{(a+ib)t}\}(s) = \mathcal{L}\{e^{at}\cos bt\}(s) + i\mathcal{L}\{e^{at}\sin bt\}(s).$$

But we also know from Exercise 1 that

$$\mathcal{L}\{e^{(a+ib)t}\}(s) = \frac{1}{s-(a+ib)} = \frac{s-a}{(s-a)^2+b^2} + i\frac{b}{(s-a)^2+b^2}$$

So matching up the real and imaginary parts of the two expressions above yields

$$\begin{aligned} \mathcal{L}\{e^{at}\cos bt\}(s) &= \frac{s-a}{(s-a)^2+b^2} \\ \mathcal{L}\{e^{at}\sin bt\}(s) &= \frac{b}{(s-a)^2+b^2}. \end{aligned}$$

Also notice that plugging in $a = 0$ yields the formulas

$$\begin{aligned} \mathcal{L}\{\cos bt\}(s) &= \frac{s}{s^2+b^2} \\ \mathcal{L}\{\sin bt\}(s) &= \frac{b}{s^2+b^2}. \end{aligned}$$

So to summarize, Exercise 1b is some useful shit. If you can remember it, then you can remember most of the other important Laplace transforms.

Exponential shifts

Calculating the Laplace transform of functions like te^{at} directly involves integration by parts. But our result below will allow us to get around these issues by yielding a general formula for $e^{at}f(t)$ for any function f. We can then apply this trick with the f of our choosing! To derive our formula, let's suppose that f is some function and s is in the domain of the Laplace transform of f. Then

$$\begin{aligned} \mathcal{L}\{e^{at}f(t)\}(s) &= \int_0^\infty e^{at}f(t)e^{-st}\,dt \\ &= \int_0^\infty f(t)e^{-(s-a)t}\,dt \\ &= \mathcal{L}\{f\}(s-a). \end{aligned}$$

(Provided $s - a$ is still in the domain[4]) This yields the important formula

$$\mathcal{L}\{e^{at} f(t)\}(s) = \mathcal{L}\{f(t)\}(s - a). \tag{3}$$

That is, multiplying a function by e^{at} performs a shift in Laplace transform space. To illustrate the usefulness of this trick, suppose that we want to compute the Laplace transform of te^{2t}. Apply the above result with $f(t) = t$ and $a = 2$. We know that f has Laplace transform

$$\mathcal{L}\{t\}(s) = \frac{1}{s^2}$$

so the Laplace transform of te^{2t} is just

$$\mathcal{L}\{t\}(s - a) = \frac{1}{(s - a)^2}.$$

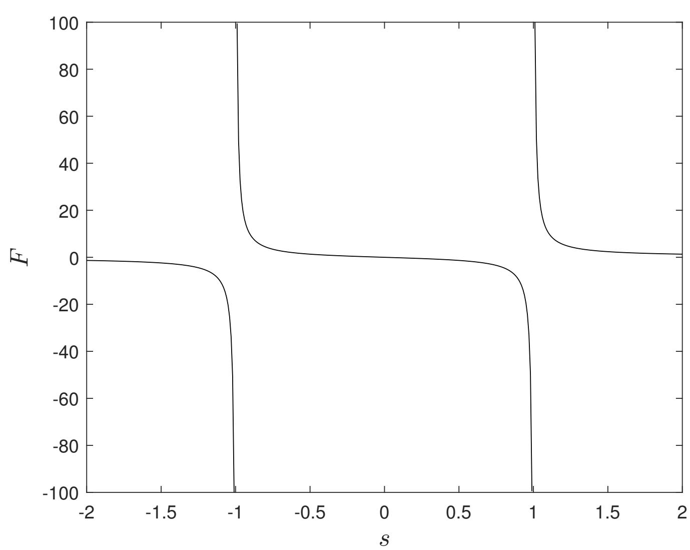
In other words, we just plug $s - a$ in for s in the Laplace transform formula.

The above trick can also be useful in remembering the formulas for $e^{at} \sin bt$ or $e^{at} \cos bt$. For example, if we remember that $\cos bt$ has Laplace transform $s/(s^2 + b^2)$, then

$$\mathcal{L}\{e^{at} \cos bt\}(s) = (s - a)/((s - a)^2 + b^2).$$

Interpretation of the previous results

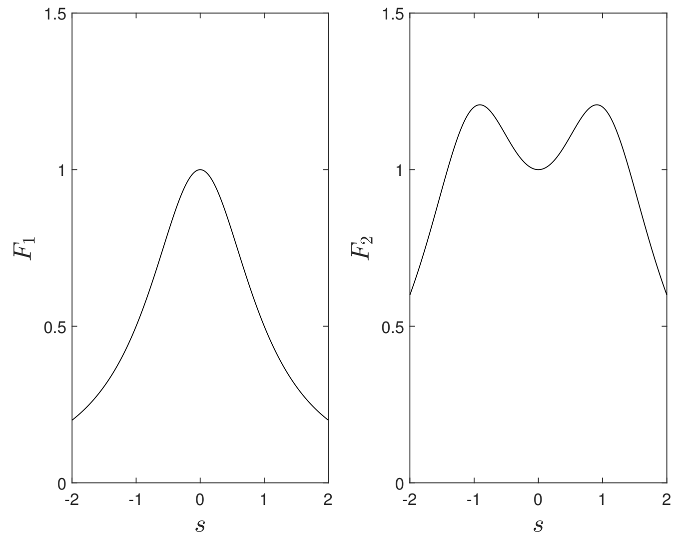
Besides providing useful tools for computing and memorizing the LTs of important functions, Exercise 1b, along with linearity and exponential shifts, can help justify our earlier statement that the Laplace transform tells you which exponentials are dominant in a function. Take the function $f(t) = e^t + e^{-t}$ which we now know has LT $F(s) = 1/(s-1) + 1/(s+1)$ plotted in the figure below.



[4]For further discussion on this topic, see Section 5.1 in Brannan and Boyce (2010) *Differential Equations: An Introduction to Modern Methods and Applications, 1st Edition.* Wiley.

Note that the Laplace transform blows up at the two exponents $s = \pm 1$. Anytime your input function f is a sum of exponentials this will always happen: the *poles* of the Laplace transform tell you the exponents of the input.

Now consider the effect of multiplying other functions by exponentials. The graph below compares the Laplace transform of $f_1(t) = \sin t$ (left) with the Laplace transform of $f_2(t) = \sin t(e^t + e^{-t}))$ (right):

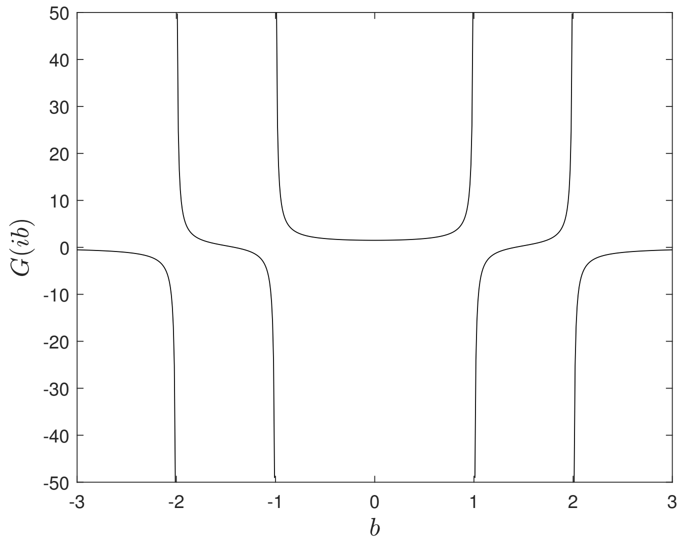


Multiplication by the exponentials has "shifted" the peak of F_1 from $s = 0$ to $s = \pm 1$. So even though it no longer blows up at these values, the LT still clearly shows the exponents dominant in the input.

What about complex exponentials? If we allow s to be a complex number, then yes, the Laplace transform tells us some important stuff about them too. For example, suppose that we wish to look at the Laplace transform of $g(t) = \sin t + \sin 2t$, a combination of two sin functions of different frequencies. Sin functions are combinations of exponentials with pure imaginary exponents, in the case of g: e^{it}, e^{-it}, e^{2it}, and e^{-2it}. Now the replace transform of g is:

$$G(s) = \frac{1}{s^2 + 1} + \frac{2}{s^2 + 4}.$$

If we replace s with ib and look at the Laplace transform $G(ib)$ as a function of b, then by analogy with what we did above, we should see spikes at $b = \pm 1, \pm 2$, the exponents. Indeed we do:



We'll return to this discussion at the end of the next section when we talk about transfer functions. But first, let's see what its like riding backwards.

Inverting the Laplace transform

Inverse functions are one of those important ideas that often get lost in the pile of precalculus laundry, but the idea is simple: inverses undo actions. I recently went to Universal Studios with my kids and they forced me to go on the Mummy ride with them. It started off fun, a quick acceleration into a forward facing roller-coaster. But then we got to the end and had to traverse the whole track in reverse to get back to where we started. In other words, we had to invert our original path.

As a part of our process for using LTs to solve DEs, we will need to solve the following type of problem: given a Laplace transform (function in s-domain), can we find the corresponding function in t-domain? This process is called ***inverting the Laplace transform***. We use the notation $\mathcal{L}^{-1}$ to denote the inverse Laplace transform. For example,

$$\mathcal{L}^{-1}\left(\frac{1}{s-1}\right) = e^t$$

because the Laplace transform of e^t is $1/(s-1)$ and

$$\mathcal{L}^{-1}\left(\frac{s}{s^2+1}\right) = \cos t$$

because the Laplace transform of $\cos t$ is $s/(s^2+1)$.

Linearity is your friend in inverting Laplace transforms. For example, we know that the inverse Laplace transform of

$$\frac{2}{s} + \frac{3}{s+2}$$

is

$$2 + 3e^{-2t}$$

because by linearity,

$$\mathcal{L}\{2+3e^{-2t}\}(s) = 2\mathcal{L}\{1\}(s) + 3\mathcal{L}\{e^{-2t}\}(s) = 2/s + 3/(s+2).$$

In fact, the linearity of the Laplace transform implies the linearity of the inverse[5].

As another example, suppose we wish to find the function $f(t)$ with Laplace transform

$$F(s) = \frac{4}{(s-2)^2+1}$$

This looks similar to the inverse Laplace transform of $e^{2t}\sin t$ (which is $1/((s-2)^2+1)$) with one difference: the 4 in the numerator. But by linearity, we can just pull that out to obtain

$$f(t) = 4e^{2t}\sin t.$$

Other manipulations of such expressions are explored in the exercises.

Inverting Laplace transforms often involves the use of partial fraction decomposition (pfd). Pfd is basically the reverse process of "finding a common denominator". In elementary school you learn that

$$\frac{1}{2} + \frac{1}{4} = \frac{4+2}{8}.$$

If we wanted to instead start with 6/8 and break it up as $A/2 + B/4$ for some constants A, B, that would be Pfd. As an example, suppose we wish to find the inverse LT of

$$\frac{1}{(s-1)(s+2)}$$

we know the inverse of $1/(s-1)$ and $1/(s+2)$ so we search for constant A, B such that

$$\frac{1}{(s-1)(s+2)} = \frac{A}{s-1} + \frac{B}{s+2}.$$

This is kind of like the Method of Undetermined Coefficients in the sense that we have identified a trial decomposition with undetermined coefficients A, B. To solve for them, now find a common denominator on the right and combined the fractions:

$$\frac{1}{(s-1)(s+2)} = \frac{A(s+2)+B(s-1)}{(s-1)(s+2)}.$$

We then group together like terms in the numerators:

$$\frac{0s+1}{(s-1)(s+2)} = \frac{(A+B)s+2A-B}{(s-1)(s+2)} \iff A+B=0,\ 2A-B=1.$$

The solution to this system of equations is $A = 1/3, B = -1/3$. Therefore,

$$\frac{1}{(s-1)(s+2)} = \frac{1/3}{s-1} - \frac{1/3}{s+2}.$$

[5] At least when the inverse exists - see Section 7.1 of Edwards and Penney for details on Existence and Uniqueness.

We can now invert each term on the right to obtain

$$\mathcal{L}^{-1}\left(\frac{1}{(s-1)(s+2)}\right) = \frac{1}{3}\left(e^{t} - e^{-2t}\right).$$

Pfd is algebraically quite tedious (and not my personal favorite mathematical algorithm), but necessary for working with LTs. In addition to the above example in which we had a product of two first order polynomials in the denominator, there are several other situations which require the use of pfd, all with different "trial solutions" to attempt. We invite you to review your favorite calculus text, visit Khan academy, or google "partial fraction decompositions" for a more comprehensive review. Additional practice will be given in the exercises below, some of which will be given as in class group works to explore together.

Exercise

(1) Calculate the Laplace transform for each of the following functions and state the domain of definition.

- (a) $f(t) = t$
- (b) $f(t) = e^{(a+ib)t}$
- (c) $f(t) = 3e^{-t}\cos t - e^{3t} + 1$
- (d) $f(t) = 2\cos 3t + 4\sin t$
- (e) te^{3t}

(2) Determine inverse Laplace transforms for the following functions.

- (a) $2s/(s^2 + 16)$
- (b) $1/(s^2 + 9)$.
- (c) $1/[(s-1)(s+2)]$.
- (d) $1/[s(s^2 + 1)]$
- (e) $3/(s^2 + 2s + 5)$
- (f) $1/[s(s^2 + 2s + 5)]$

(3) Determine inverse Laplace transforms for the following functions.

- (a) $4s/(s^2 + 4)$.
- (b) $3/(s^2 + 4)$.
- (c) $5/(s+1) + s/[(s+1)^2 + 4]$.
- (d) $1/[s(s-4)]$.
- (e) $3/(s^2 + 4s + 13)$

(4) Determine a formula for the Laplace transform of $f(t) = t^2$ and then generalize your result to find a formula for the Laplace transform of t^n which applies to any integer $n \geq 1$.

(5) Calculate the Laplace transform $Y(s)$ of the function

$$y(t) = \cos 3t + \cos 5t$$

and create a plot of $Y(ib)$ vs b over the interval $-10 \leq b \leq 10$. Where are the spikes? How could you have guessed their locations ahead of time by looking at y?

(6) Calculate the Laplace transform $Y(s)$ of the function

$$y(t) = e^{-t/10} \sin t$$

and then plot $|Y(ib)|$ vs b over the interval $-2 \leq b \leq 2$ where $|\cdot|$ represents the magnitude or modulus of $|Y(ib)|$. Discuss what you see in the graph and explain why can't we just plot $Y(ib)$ in this case.

Section 2.7 :
Solving Differential Equations with Laplace Transforms

Section objectives:

- Determine solutions to first and second order constant coefficient linear equations using Laplace Transforms.

In the last section, we defined the Laplace Transform (LT) $\mathcal{L}\{f\}(s)$ of a function $f(t)$ by

$$\mathcal{L}\{f\}(s) = \int_0^\infty e^{-st} f(t)dt \tag{1}$$

Today, we will connect this apparent red herring back to the topic of this course. We start with an example.

Solving first order equations with Laplace transforms

Consider the initial value problem

$$\begin{aligned} y' &= y \\ y(0) &= 1. \end{aligned}$$

We already know the solution is

$$y(t) = e^t,$$

but let's solve this equation using Laplace transforms just for shits and giggles. First, take the Laplace transform of both sides of the equation $y' = y$ to obtain the identity

$$\mathcal{L}\{y'\}(s) = \mathcal{L}\{y\}(s). \tag{2}$$

If there is a way of relating $\mathcal{L}\{y'\}(s)$ to $\mathcal{L}\{y\}(s)$, we may be able to solve this equation to obtain a formula for $\mathcal{L}\{y\}(s)$. In fact, there is a general way of relating these two expressions using integration by parts:

$$\begin{aligned} \mathcal{L}\{y'\}(s) &= \int_0^\infty y'(t)e^{-st}\, dt \\ &= y(t)e^{-st}\Big|_0^\infty + s\int_0^\infty y(t)e^{-st}\, dt \\ &= -y(0) + s\mathcal{L}\{y\}(s). \end{aligned}$$

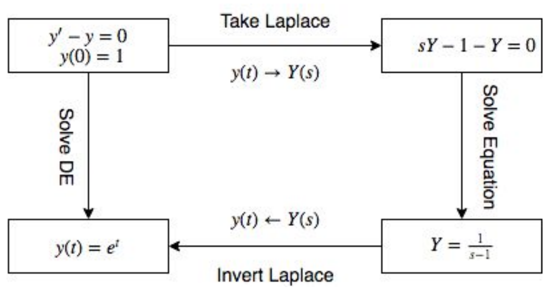
(Note: in the last line, we need the assumption that s is in the domain of y so that $e^{-st}y(t) \to 0$ as $t \to \infty$). The result of this calculation is the formula

$$\mathcal{L}\{y'\}(s) = s\mathcal{L}\{y\}(s) - y(0). \tag{3}$$

Substituting this expression back into (2) and letting $Y = \mathcal{L}\{y\}(s)$ to avoid the overly awkward notation yields

$$\begin{aligned} sY - y(0) &= Y \\ (s-1)Y &= y(0) \\ Y &= \frac{1}{s-1}. \end{aligned}$$

The formula in (3) is pretty powerful: it allows us to convert a first order differential equation (with independent variable t and dependent variable y) into a linear algebraic equation for the unknown function Y and this algebraic equation is much easier to solve. This is the value of Laplace transforms! In our picture from last lecture notes,



we have now gone from the upper left corner to the lower right corner. The last step which remains is to "invert" the Laplace transform and obtain y from Y; this is the step in which we go from the bottom right to the bottom left in the figure above. In our example, this is relatively easy: we need to find a function y whose Laplace transform is $1/(s-1)$. Reviewing our examples from last class shows that e^{at} has Laplace transform $1/(s-a)$. Here, our Laplace transform looks like this with $a = 1$ so $y(t) = e^t$.

To summarize: our work:

- $Y(s) = 1/(s-1)$ was the solution obtained in s (LT) domain
- $y(t) = e^t$ was the solution obtained in t (real time) domain.

Now let's see how this plays out in a second order example.

Second order example

Consider the SOCCHE equation $y'' + y = 0$ with initial conditions $y(0) = 1, y'(0) = 0$. Applying the Laplace transform to both sides of the SOCCHE and setting $Y = \mathcal{L}\{y\}(s)$ yields

$$\mathcal{L}\{y''\}(s) + Y = 0. \tag{4}$$

To use the process from the previous section, we need to relate the LT of y'' back to the LT of y. This can be done by iterating the formula in Equation (3):

$$\begin{aligned}\mathcal{L}\{y''\}(s) &= s\mathcal{L}\{y'\}(s) - y'(0)\\ &= s(s\mathcal{L}\{y\}(s) - y(0)) - y'(0)\\ &= s^2Y - sy(0) - y'(0).\end{aligned} \tag{5}$$

Using this new formula, Equation (4) reduces to

$$s^2Y - sy(0) - y'(0) + Y = 0.$$

or after plugging in our initial conditions and solving for Y,

$$(s^2 + 1)Y - s = 0 \Longrightarrow Y = s/(s^2 + 1).$$

Reviewing our notes from last time shows that $s/(s^2 + 1)$ is the Laplace transform of $\cos t$ so that $y(t) = \cos t$ is the solution to the SOCCHE.

LT vs Characteristic Equations

Equations (3) and (5) tell us how to convert any first or second order constant coefficient *differential* equation in y into a first or second order[1] *algebraic* equation in Y. Once this task is completed, we solve for Y and use algebraic trickery, partial fraction, whatever it takes to rewrite the formula for Y as the Laplace transform of something we recognize. This method has some advantages over the method of characteristic equations. First of all, the initial conditions are automatically worked into the equations so there is no PSI step. Second, you bypass the need to deal with complex numbers in second order examples. Third, it can be applied to both homogeneous and nonhomogeneous equations and hence, there is no need for the whole "$y_p + y_c$" stuff. The disadvantage is that using Laplace transforms to solve first and second order equations often just turns into a not too exciting exercise in applying complicated partial fraction decompositions.

An an example of the advantages and disadvantages, consider the IVP

$$y'' + y = 0,\ y(0) = 1,\ y'(0) = 0.$$

Taking the LT of both sides in the DE and applying (5) gives us the algebraic equation

$$s^2Y - sy(0) - y'(0) + Y = 0$$

for $Y(s) = \mathcal{L}\{y\}(s)$. Solving for Y and plugging in initial conditions then yields

$$Y = \frac{s}{s^2+1}.$$

which we recognize as the LT of $y(t) = \cos t$. That was quicker than mucking around with Euler's Formula! But now suppose that we have the nonhomogeneous equation

$$y'' + y = \sin(2t),\ y(0) = 0,\ y'(0) = 0.$$

which we previously solved using MUC. This time, taking the LT of both sides and applying (5) leads to the algebraic equation

$$s^2Y + Y = \frac{2}{s^2+4}.$$

Now when we solve for Y, we need to look for a partial fraction decomposition of the form

$$Y = \frac{2}{(s^2+1)(s^2+4)} = \frac{As+B}{s^2+1} + \frac{Cs+D}{s^2+4}. \tag{6}$$

Finding a common denominator and matching like terms yields the systems of four equations

$$\begin{aligned} A + C &= 0 \\ B + D &= 0 \\ 4A + C &= 0 \\ 4B + D &= 2. \end{aligned}$$

[1]The exercises investigate formulas for higher order differential equations as well.

for the four unknowns A, B, C, D. Noting that the first and third equations yields $A = C = 0$, we can use the second and fourth to get $B = 2/3, D = -2/3$. Plugging this all into equation (6) implies that

$$Y = \frac{2}{3}\left(\frac{1}{s^2+1} - \frac{1}{s^2+4}\right).$$

which we can now invert to obtain the solution

$$y(t) = \frac{2}{3}\left(\sin t - \frac{1}{3}\sin 2t\right).$$

Was this last example easier then our previous procedure using MUC? Maybe slightly if you don't mind all the partial fractions and are good at solving systems of equations. In general, here are my personal "quick and dirty" opinions about how I would approach second order equations:

- If you only want to know about the general solution or asymptotic behavior, use the method of characteristic equations/undetermined coefficients and don't bother with Laplace.
- If you need to solve a homogeneous IVP, use the method of characteristic equations if the roots are real and LT if the roots are complex.
- If you need to solve a nonhomogeneous IVP, both methods are tedious.
  - Choose MUC if you prefer calculus and complex numbers.
  - Choose LT if you prefer partial fractions and linear equations.

Of course you will undoubtedly encounter others with different perspectives, maybe even someone who requires you to use their preferred method, so its useful to be comfortable with both classes of approaches.

Transfer Functions

More important than arguable gains in calculation speed are the gains in simplicity we get by applying LTs to study input-output type problems[2]. For example, consider the generic SOCCNE

$$my'' + \gamma y' + ky = g(t)$$

with initial conditions $y(0) = y_0, y'(0) = v_0$. Taking the Laplace transform of both sides gives the equation

$$ms^2Y - mv_0 - msy_0 + \gamma sY - \gamma y_0 + kY = G(s)$$

and solving for Y yields the solution

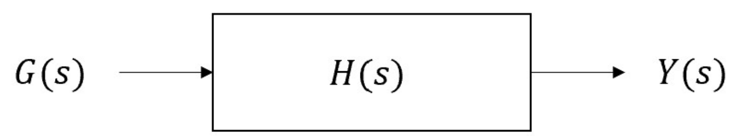
$$Y(s) = \overbrace{H(s)[(ms+\gamma)y_0 + mv_0]}^{\text{Free Response}} + \overbrace{H(s)G(s)}^{\text{Forced Response}} \tag{7}$$

[2]One particularly important application is designing controller for feedback loops, but you'll have to take Control Systems to find out more about that.

where

$$H(s) = \frac{1}{ms^2 + \gamma s + k} = \frac{1}{\text{Characteristic Polynomial}}.$$

Looking at the two terms on the right side of (7), the first term is often called the *free response* of the system because it involves the initial conditions and the second part is called the *forced response* because it involves the reaction the external forcing. H is called the *transfer* function because it describes how the system responds to both initial conditions (free response) and external motivation (forced response). It essentially describes the way in which are system converts inputs into outputs:



The analog of Equation (7) in the time-domain would be our old friend.

$$y(t) = y_c(t) + y_p(t)$$

However, the formulation in Laplace land incorporates y_0, v_0 into the free response directly and hence, instead of just any old y_p we have a very special one given by the inverse LT of $H(s)G(s)$. The advantage of using this new formulation is that we don't have to go through all the hassle of finding a particular solution in order to perform an input-output analysis of the system: we just need to study the transfer function H.

To see why, suppose we start a system off with zero initial conditions so that the solution in Laplace land is just the forced response

$$Y(s) = H(s)G(s)$$

and we would like to determine how sensitive our system is to different forcing frequencies ω like we did in our previous frequency-response type analyses. In Section 2.5, we showed that studying the graph of $Y(i\omega)$ can reveal important information about spiking frequencies in y. Since the transfer function H is the ratio of input to output:

$$|H| = \left|\frac{Y}{G}\right| = \frac{\text{output}}{\text{input}},$$

graphing the magnitude of $H(i\omega)$ vs ω will generate a frequency-response curves whose peaks identify the resonant frequencies for the system.

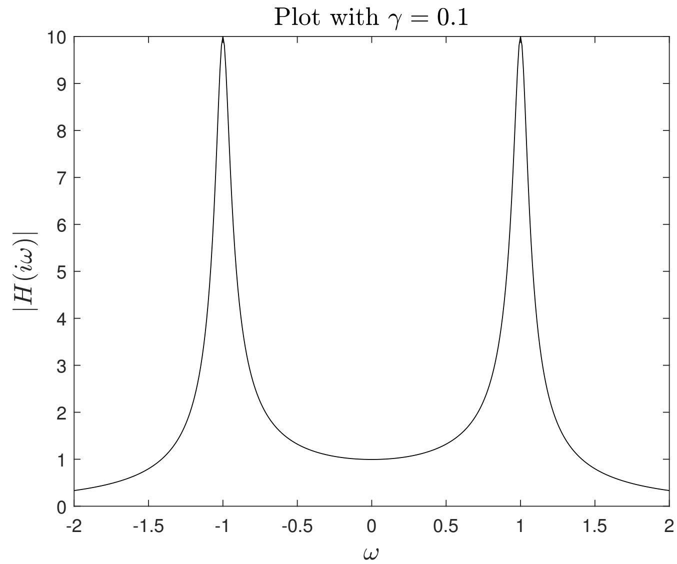
As an example, consider the spring-mass system $y'' + \gamma y' + y$. The transfer function will be

$$H(s) = \frac{1}{s^2 + \gamma s + 1}$$

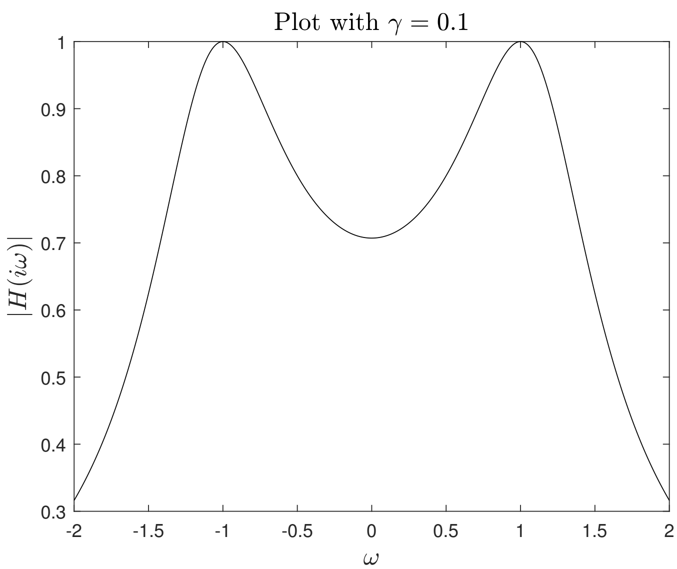
and the frequency-response function will be

$$|H(i\omega)| = \left|\frac{1}{(1-\omega^2) + i\gamma\omega}\right|$$

If $\gamma = 0$, then we get $1/(1-\omega^2)$ which blows up at the resonant (angular) frequencies of $\omega = \pm 1$. When $\gamma > 0$ but small, we get peaks at ± 1:



As we increase the damping, we reduce the peaks, but they are still there:



Exercises

(1) Derive a formula of the LT of y''' in terms of the LT of y as in Equations (3) and (5).

(2) Use LTs to solve the following IVPs. Compare with solutions obtained via previous methods.

(a) $y' = -3y,\ y(0) = 2$

(b) $y'' + y = 0,\ y(0) = 2,\ y'(0) = -1$

(c) $y'' = -4y,\ y(0) = 2, y'(0) = -1$

(d) $y' = 2 + y,\ y(0) = 1$

(e) $y' = e^{-2t} - y,\ y(0) = 1$

(f) $y'' + 9y = 0, y(0) = 1, y'(0) = 1$

(g) $y'' + y' - 2y = 0, y(0) = 0, y'(0) = 1$

(h) $y'' + y = 1, y(0) = 0, y'(0) = 0$

(i) $y'' + 2y' + 5y = 0, y(0) = 0, y'(0) = 1$

(j) $y'' - y = 0,\ y(0) = 0,\ y'(0) = 2.$

(k) $y'' + 6y' + 25y = 0,\ y(0) = 0,\ y'(0) = 1.$

(l) $y'' + 4y = 4,\ y(0) = 0,\ y'(0) = 0.$

(m) $y'' + 4y = \sin t,\ y(0) = 0,\ y'(0) = 0.$

(3) Find a formula for the solution y to the initial value problem $y'' + \omega^2 y = 0$ with initial conditions $y(0) = y_0, y'(0) = v_0$ in terms of ω, y_0, v_0. Compute the amplitude, period, and frequency of the solution and describe how they depend on y_0, v_0, and ω.

(4) Consider a spring mass system with $m = 1, \gamma = 2, k = 1$.

(a) What is the transfer function for this system?

(b) Derive a formula for and plot the frequency response function for this system. Interpret the resulting graph.

Section 2.8: Laplace Transforms and Discontinuous Forcing Functions

Section objectives:

- Setup up DEs involving piecewise defined forcing functions
- Derive the time-shift formula for Laplace transforms
- Use Laplace transforms to solve DEs involving piecewise constant functions and interpret the solution

Motivation for discontinuous forcing

So far in this course, we have looked at several applications which lead to non-homogeneous equations

- Newton's Law of Cooling
- Mixing Problems
- Circuits with an DC/AC power source
- Spring-mass systems with a forcing mechanism

In Section 2.5, we looked at some examples of the latter two types of problems in which the forcing term was discontinuous but periodic. However, if instead we want to turn on (off) a force after a specified amount of time and leave it on (off), this leads to a slightly different type of modeling scenario. In a mixing problem, we could also imagine decontaminating (or contaminating) a water source after several days. In Newton's Law of Cooling, we could imagine turning on an ac unit during the hours of 9 am and 6 pm. All of these would lead to discontinuous forcing functions in the corresponding IVPs.

We will demonstrate this modeling procedure first using second order equations since that is the theme of this chapter, but will show an example incorporating mixing problems later in this section as well. Suppose that we have a standard spring-mass system with mass m, damping coefficient γ, spring constant k, initial displacement y_0, and initial velocity v_0 m/s^2. Suppose that initially there is no external forcing applied, but after 1-second, we turn on a constant force of 1 N. Then the IVP for the displacement $y(t)$ after t seconds can still be given by the IVP

$$\begin{aligned} my^{''} + \gamma y' + ky &= g(t) \\ y(0) &= y_0 \\ y'(0) &= v_0 \end{aligned}$$

but now we need to represent g by the *piecewise constant* function

$$g(t) = \begin{cases} 0 & 0 \leq t \leq 1 \\ 1 & t > 1 \end{cases}.$$

Taking the LT of both sides of our SOCCNE and using the formulas from last section, we obtain the equation

$$m(s^2Y - sy(0) - y'(0)) + \gamma(sY - y(0)) + kY = G$$

where $Y = \mathcal{L}\{y\}(s), G = \mathcal{L}\{g\}(s)$. Solving for Y gives

$$Y = \frac{my_0 s + \gamma y_0 + mv_0}{ms^2 + \gamma s + k} + \frac{G}{ms^2 + \gamma s + k}.$$

So to be able to invert this, we first need to be able to calculate the LT of g. This can be done by breaking up the integral in the Laplace transform definition into two separate regions and plugging in the value of g on each:

$$\begin{aligned} \mathcal{L}\{g\}(s) &= \int_0^\infty e^{-st} g(t)\, dt \\ &= \int_0^1 e^{-st} g(t)\, dt + \int_1^\infty e^{-st} g(t)\, dt \\ &= \int_0^1 e^{-st}(0)\, dt + \int_1^\infty e^{-st}(1)\, dt \\ &= e^{-s}/s. \end{aligned}$$

This allows us to simplify the above expression to

$$Y = \frac{my_0 s + \gamma y_0 + mv_0}{ms^2 + \gamma s + k} + \frac{e^{-s}}{s(ms^2 + \gamma s + k)}.$$

Now the first term is a rational function so we should be able to invert it, but the second term has this strange e^{-s} in the numerator which makes it like nothing we have seen thus far. So before we continue with some full examples, we will take a quick detour to discuss the inversion of terms of the form $e^{-sc}F(s)$ where F is a LT we know how to invert.

The time shift formula

The key to inverting terms of the form $e^{-sc}F(s)$ is to write problems in terms of the function

$$u_c(t) = \begin{cases} 0 & 0 \le t < c \\ 1 & t \ge c \end{cases}$$

which we often call the *unit step* function. In the last section, we calculated $u_1(t)$. To find a general formula for the Laplace transform $U_c(s)$ of $u_c(t)$, we use a similar procedure and break up the integral in the definition of the LT between the intervals $(0, c)$ and (c, ∞):

$$\begin{aligned} \mathcal{L}\{u_c(t)\}(s) &= \int_0^\infty u_c(t) e^{-st}\, dt \\ &= \int_0^c 0e^{-st}\, dt + \int_c^\infty 1e^{-st}\, dt \\ &= 0 + \frac{-1}{s} e^{-st}\Big|_c^\infty \\ &= \frac{e^{-cs}}{s} \end{aligned}$$

provided that $s > 0$. So

$$U_c(s) = \frac{e^{-sc}}{s}.$$

How does this help us? Well, suppose that we encounter a product of the form $e^{-sc}F(s)$ where $c > 0$ is a constant and we recognize F as the Laplace transform of some function $f(t)$. Then, using the definition of the Laplace transform

$$\begin{aligned} e^{-sc}F(s) &= \int_0^\infty e^{-sc} f(t) e^{-st}\, dt \\ &= \int_0^\infty f(t) e^{-s(t+c)}\, dt \\ &= \int_c^\infty f(v-c) e^{-sv}\, dv \qquad (1) \end{aligned}$$

where in the last line, we made a change of variables $v = t + c$ (notice that this changed the lower limit of integration from 0 to c). This almost looks like the Laplace transform of f, but there are a few discrepancies. First, we are integrating with respect to v instead of t, but that doesn't matter since the integration variable is just a dummy variable anyways. Second, the limits are from c to ∞ instead of 0 to ∞. However, we can take care of this by replacing $f(v-c)$ with the product $f(v-c)u_c(v)$: if $v < c$, then $u_c(v) = 0$ so that $f(v-c)u_c(v) = 0$ while if $v > c$, $u_c(v) = 1$ so that $f(v-c)u_c(v) = f(v-c)$. These newfound revelations imply that

$$\int_c^\infty f(v-c) e^{-sv}\, dv = \int_0^\infty f(v-c) u_c(v) e^{-sv}\, dv$$

and by definition, the right side of the above equation is the Laplace transform of the function $g(t) = f(t-c)u_c(t)$. Combining this with the calculation in (1) leads to the ***time shift*** formula:

$$\mathcal{L}\{f(t-c)u_c(t)\}(s) = e^{-sc}F(s).$$

For application, this formula is more useful in its inverted form:

$$f(t-c)u_c(t) = \mathcal{L}^{-1}\{e^{-sc}F(s)\}.$$

In words, if we encounter a term of the form:

$$e^{-sc}F(s)$$

and we recognize $F(s)$ as the Laplace transform of some function $f(t)$, then the inverse Laplace transform of $e^{-cs}F(s)$ is $u_c(t)f(t-c)$.

Examples and Applications.

With the time shift formula in hand, we can now solve SOCCNE with discontinuous forcing. We first consider the concrete example $y'' + y = u_1(t)$ with initial conditions $y(0) = 1, y'(0) = 0$. Taking LT of both sides of the DE yields

$$s^2Y - sy(0) - y'(0) + Y = e^{-s}/s \Longrightarrow Y = s/(s^2+1) + e^{-s}/[s(s^2+1)].$$

Now we easily recognize $s/(s^2+1)$ as the LT of $\cos t$. For the second term, we ignore the e^{-s} and use partial fractions to break apart

$$F(s) = \frac{1}{s(s^2+1)} = \frac{A}{s} + \frac{Bs+C}{s^2+1}.$$

After some algebra, we get $A = 1, B = -1, C = 0$ so that

$$F(s) = \frac{1}{s} - \frac{s}{s^2+1}$$

which has inverse LT $f(t) = 1 - \cos t$. Using the time-shift formula implies that the inverse LT of $e^{-s}F(s)$ is therefore

$$u_1(t)f(t-1) = (1 - \cos(t-1))u_1(t).$$

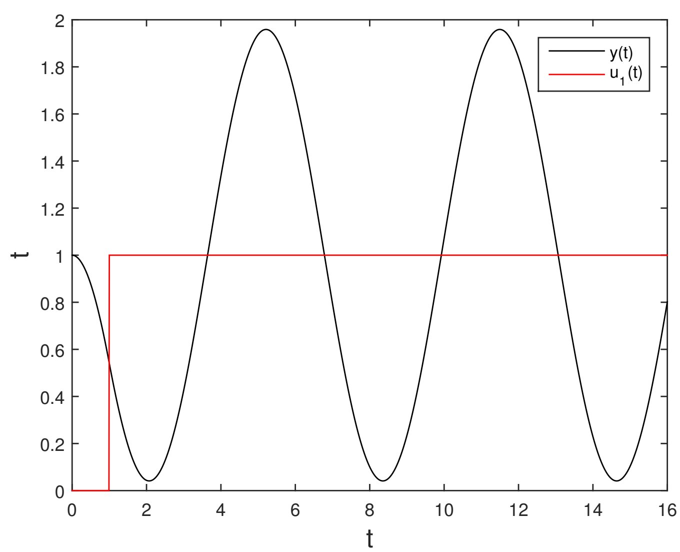
Combining this with the inverse of $s/(s^2+1)$ yields the solution

$$y(t) = \cos t + (1 - \cos(t-1))u_1(t).$$

To get a better idea of what this solution means, plug in $u_1(t) = 0$ for $0 \le t < 1$ and $u_1(t) = 1$ for $t \ge 1$. Then

$$y(t) = \begin{cases} \cos t & 0 \le t < 1 \\ \cos t + 1 - \cos(t-1) & t \ge 1 \end{cases}.$$

This solution is plotted below against the unit-step function. Note that the solution y is continuous even though the forcing is not!



Other types of piecewise constant forcing functions can be handled by taking linear combinations of unit step functions. For example, suppose in the previous example that we want to apply a force of 2 N after 1 seconds, but then shut off the force after an additional second. We can write this as a piecewise function

$$g(t) = \begin{cases} 0 & 0 \le t < 1 \\ 2 & 1 \le t < 2 \\ 0 & t > 2 \end{cases}$$

but this definition is equivalent to using the forcing function $2(u_1(t) - u_2(t))$. Such examples will be further explored in the exercises.

Mixing Example

We would like to close with one example of a discontinuous forcing function in a mixing problem. Xavier's School for the Gifted Youngsters has a 200 gal tank which supplies drinking water for its students and this tank is initially filled with pure water. The tank is supplied from a nearby stream at a rate of 40 gal per hour and water flows from the tank into the school's drinking fountains at a rate of 40 gal per hour as well so that the amount of water in the tank stays constant over time. One day, Worthington Industries begins polluting the stream with its new mutant "cure" in a concentration of 1/10 oz per gal, however, the mutants catch this mistake after one hour and pure water once again begins to flow into their reservoir. If Xavier deems the water drinkable once the concentration of contaminant is less than 1/100 oz per gal, when can the children quench their thirst?

To solve this problem, we begin by setting up an initial value problem for the unknown function, $y(t)$ = amount of cure in the tank after t hours. Let

$$c(t) = \begin{cases} 0.1 & 0 \le t < 1 \\ 0 & t \ge 1 \end{cases}$$

be the function giving the concentration of cure flowing into the tank at time t. Then our standard rate in/rate out analysis implies that

$$\frac{dy}{dt} = 40c(t) - 0.2y(t).$$

Taking the Laplace transform of each side of our equation and rearranging tells us that y has Laplace transform Y given by

$$Y(s) = \frac{4(1 - e^{-s})}{s(s + 0.2)} = (1 - e^{-s})F(s)$$

where $F(s) = 4/[s(s + 0.2)]$ is the Laplace transform of $f(t) = 20(1 - e^{-0.2t})$. Therefore, we can apply the Time-Shift formula with $c = 1$ to conclude that

$$\begin{aligned} y(t) &= f(t) - f(t-1)u_1(t) \\ &= 20(1 - e^{-0.2t}) - 20(1 - e^{-0.2(t-1)})u_1(t) \\ &= \begin{cases} 20(1 - e^{-0.2t}) & 0 \le t < 1 \\ 20e^{-0.2t}(e^{0.2} - 1) & t \ge 1 \end{cases} \end{aligned}$$

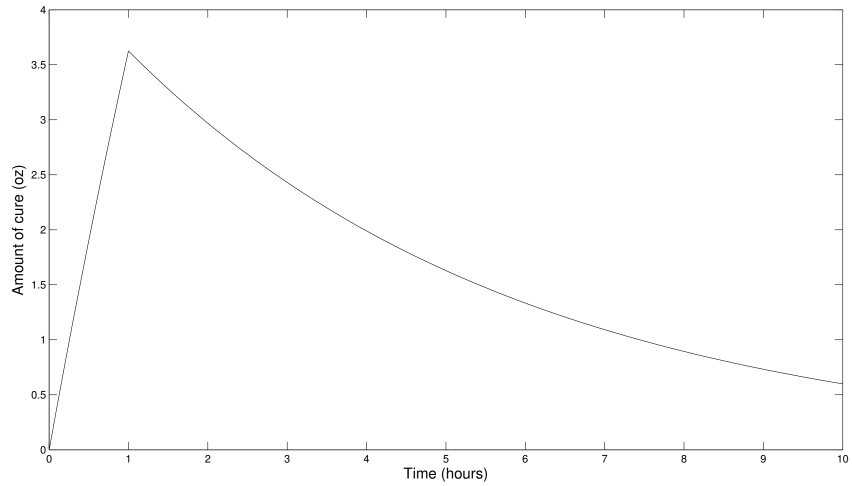


Figure 1: Solution $y(t)$ to the mutant-cure problem

The solution is plotted in Figure 1. Notice that the solution is continuous even though the forcing term is not, however, the solution is not differentiable at the discontinuity of the forcing (at time $t = 1$). Also notice that the amount of cure rises towards the concentration of chemical in the stream ($1/10 \times 200 = 20$ oz) up until time 1 and then exponentially decays after the cure is detected. To answer our original equation of interest, the concentration of cure in the reservoir will reach a safe level of 1/100 oz per gal when the amount in the reservoir reaches 2 oz and the time at which this occurs can be found by solving

$$20e^{-0.2t}(e^{0.2} - 1) = 2$$

to obtain $t \approx 3.97$ hours.

Exercises

(1) Calculate the Laplace transform of the function

$$g(t) = \begin{cases} 0 & 0 \leq t < 1 \\ 2 & 1 \leq t < 2 \\ 0 & t \geq 2 \end{cases}.$$

(2) Calculate the LT of the following forcing function: a constant force of 1 N applied from time 1 until time 2, a constant force of 4 N applied from time 2 until time 4 and no force applied otherwise.

(3) Solve the IVP $y'' + 4y = 4u_5(t), y(0) = 0, y'(0) = 1$.

(4) For each of the following scenarios, write down and solve an IVP for the mass's displacement after t seconds.

(a) Mass of 1 kg, no damping, spring constant $k = 4$, initial displacement of 1 meters, initial velocity of 0 m/s, and a constant external force of 1 N applied after 3 seconds.

(b) Mass of 1 kg, no damping, spring constant $k = 4$, initial displacement of 1 meters, initial velocity of 0 m/s, and a constant external force of 4 N applied after 3 seconds.

(c) Mass of 1 kg, no damping, spring constant $k = 4$, initial displacement of 1 meter, initial velocity of -1 m/s, and a constant external force of 1 N applied after 3 seconds.

(d) Mass of 1 kg, damping coefficient of $\gamma = 2$, spring constant $k = 5$, initial displacement of 1 meter, initial velocity of 0 m/s, and a constant external force of 1 N applied after 1 seconds.

(5) Suppose that a 1 kg mass in an undamped spring mass system with spring constant $k = 9$ is initially displaced 1 meter from equilibrium and released. After 2 seconds, a constant external force of 3 Newtons is applied. Determine a formula for the displacement of the mass after t seconds and use your formula to calculate the displacement at the following times: (i) $t = \pi/2$, and (ii) $t = \pi + 2$.

Part III
Systems of Equations

Linear Algebra Supplement

Introduction

Linear algebra is one of the pillars of mathematics and we cannot do justice to the subject in a few pages. Our goal here is just to help clarify some important definitions, terminology, and notation that will be used throughout this chapter. I highly recommend you take an introduction to linear algebra course (Math 75 here at Pacific) at some point during your education to fully appreciate the subject. I also recommend videos 1-6 of the "Essence of Linear Algebra" series on the YouTube channel 3Blue1Brown which can be viewed at the following url:

`http://www.3blue1brown.com/essence-of-linear-algebra-page/`.

They have some beautiful geometric interpretations and visualizations which may be helpful in better understanding the algebraic operations defined below.

We will be mainly using linear algebra as a computational and organizational tool in this course[1], a way to efficiently represent large systems of equations and solve them using tried and true techniques. There are a few key skills we hope you will get out of this brief introduction and which will be important moving forward:

- Representing linear systems of equations in matrix-vector form
- Perform basic matrix-vector calculations like adding/scaling vectors or multiplying matrices/vectors
- Calculating and interpreting determinants of matrices
- Defining and interpreting inverse matrices
- Using MATLAB to solve linear systems of equations

But first...

The Big Picture

Linear algebra as the name suggests is about linear equations. To begin, please take a moment to do the following artistic venture:

(i) Draw two lines that intersect in exactly one point and write down their equations

(ii) Draw two lines that intersect in no points and write down there equations

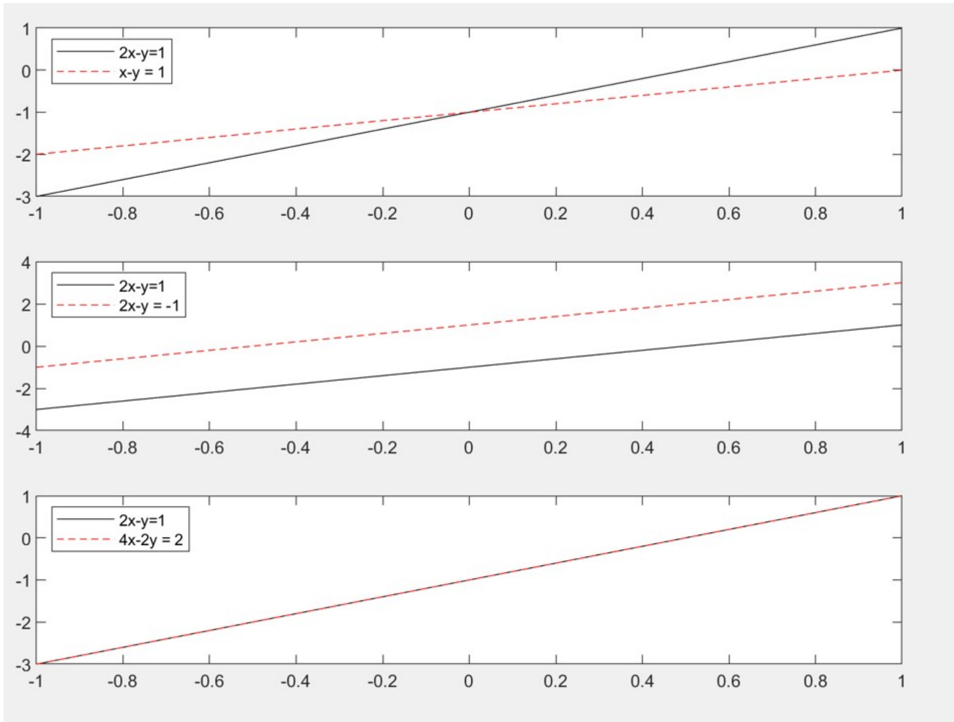
[1] 3blue1brown calls this the "computer science" perspective

(iii) Draw two lines that intersect in infinitely many points and write down their equations

(iv) Can you think of any other possibilities besides 1-3 above?

I will start a new page so you aren't tempted to cheat and look at the answer too quickly.

Ok, done thinking about it? Hopefully, you came up with something similar to the trio of possibilities below (and said no to question four):



The three pictures above correspond respectively to (top) lines with different slopes that intersect in exactly one point, (middle) lines with the same slope but different intercepts that never intersect, or (iii) lines with the same slope and same intercept that intersect in infinitely many points. All three plots above are visual representations of the following type of problem: find all pairs of values x, y which simultaneously satisfy the two equations

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned} \tag{1}$$

This is called a *linear system of equations* in the two unknowns x, y and any pair of values x, y which makes both equations work is called a ***solution*** to the system. The constants a_{ij} are called *coefficients* and they are "double subscripted" in such a way that the first subscript tells you which equation the constant is in (1 or 2) and the second tells you which variable the coefficient is in front of (1 for x and 2 for y). For this reason, we will often use x_1, x_2 instead of x, y in our equations[2]. We have seen many examples of this type of problem when solving IVPs and we have typically solved them using elimination of variables with substitution. For example, if we were given the system

$$\begin{aligned} x + y &= 2 \\ x - y &= 0 \end{aligned}$$

we would add the two equations together to get $2x = 2$, solve for $x = 1$, and then sub this back into either equation to get $y = 1$ as well.

[2]Which is especially helpful in higher dimensions when you start to run out of letters. x,y,z,w, and then what?

Linear algebra begins with finding more efficient methods for carrying out this process, algorithms that apply in higher dimensions as well. You are probably getting tired of doing elimination by this point. There should be a way to routinize the process and avoid such tedium in the future. In addition, we noted above that linear systems in two unknowns can either have 1, 0, or infinitely many solutions. This trio of possibilities in fact holds for larger systems as well[3], but it is much harder to tell from looking at the equations which situation we are in. So another goal of linear algebra is to come up with simple methods for determining which of the three possibilities holds for a given system of equations.

The key to achieving both goals is organizing information in linear systems into matrices and vectors. When we are solving a linear equation in one variable

$$ax = b$$

where a, b are constants, we could divide both sides by a, provided $a \neq 0$, and get the solution $x = b/a$. Is there an analog of this process for linear system of two or more equations? Furthermore, what is the purpose of the condition $a \neq 0$? The next few sections will explore answers to both these questions while covering the five core tools discussed in the introduction along the way.

Matrix-Vector Form of Systems

To make an analogy between the system of two equations in two unknowns like in (1) with a single equation in one unknown, $ax = b$, we need to identify what the a, x, b are in the larger system. Well, x is the unknown in our 1D example and now we have two unknowns x_1, x_2 so it makes sense to organize them into a vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

We'll call this the ***solution vector***. In the 1d example, a is the coefficient of x. Now we have four coefficients in two rows so let's slap some brackets around them and call the resulting object A:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

The object A is what we call the ***coefficient matrix***. It is a 2×2 matrix with two rows (going across) and two columns (going down). The first subscript now refers to the row in which we find a number and the second refers to the column. Now maybe, with the right definition of a matrix-vector product, we could write the left hand side of the equations in (1) as:

$$A\mathbf{x}.$$

We are getting closer to $ax = b$; we just need the righthand side now. Since there are two numbers on the right, let's also organize them into a vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

[3] A fact, at risk of sounding blasphemous that I often refer to as the "Holy Trinity of Linear Algebra".

Now we have $A\mathbf{x} = \mathbf{b}$; objective achieved.

But hold on a minute. We said "if we define matrix-vector products in the right way". What is that right way? The next section will tackle that question. But first, a couple examples. If our system of equations is

$$\begin{aligned} x + y &= 2 \\ x - y &= 0 \end{aligned}$$

then $A\mathbf{x} = \mathbf{b}$ with coefficient matrix, solution vector, and constant vector:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

However, if our system of equations is

$$\begin{aligned} x_2 &= 2 \\ 2x_1 &= -1 \end{aligned}$$

then $A\mathbf{x} = \mathbf{b}$ with coefficient matrix, solution vector, and constant vector:

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Matrix-Vector Multiplication

First, let's provide more information on what we mean by matrices and vectors. A ***matrix*** is, for the purposes of this class, any $m \times n$ array of numbers, arranged into m rows and n columns. A matrix with only a single column and n rows is called a ***vector***. We refer to the element in the ith row and jth column of a matrix A as a_{ij} and call this the ijth ***element*** of the matrix. If $\mathbf{y}$ is a vector, we usually just say the first, second, third element (or sometimes component), etc, and write y_j. We will use capitol letters (A, B, C, etc) to denote matrices and boldface letters ($\mathbf{y}, \mathbf{v}$, etc) to denote vectors. Most of the examples below consider only 2×2 matrices and vectors of length 2, but the same arguments hold for higher dimensional matrices as well. For vector of length 2, we often just use the familiar (x, y) notation for referring to its components. The ***diagonal*** of a matrix will always refer to the main diagonal consisting of all elements whose row and column indices are identical; that is, all elements of the form a_{ii}.

We define the ***vector space*** $\mathbb{R}^2$ as the set of all two element vectors equipped with the following operations:

(Scalar Multiplication) If c is a real (or complex) number (scaler) and $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is a vector then

$$c\mathbf{v} = \begin{bmatrix} cx \\ cy \end{bmatrix}$$

(Vector Addition) If $\mathbf{v_1} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ are two vectors, then

$$\mathbf{v_1} + \mathbf{v_2} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}.$$

In other words, multiplying a vector by a real number "scales" each component while adding two vectors corresponds to component by component addition. Note that the rules above apply to real and vector valued functions as well. For example,

$$e^{2t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} e^{2t} \\ -e^{2t} \end{bmatrix}$$

and

$$\begin{bmatrix} e^t \\ e^t \end{bmatrix} + \begin{bmatrix} e^{2t} \\ -e^{2t} \end{bmatrix} = \begin{bmatrix} e^t + e^{2t} \\ e^t - e^{2t} \end{bmatrix}.$$

These observations will be useful when we get back to discussing systems of equations.

Given a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

we can define a function or *transformation* from $\mathbb{R}^2$ into itself as follows: if $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ is any vector in $\mathbb{R}^2$, then $A\mathbf{y}$ is the vector whose first element is the dot product of the first row of A and $\mathbf{y}$ and whose second element is the dot product of the second row of A and $\mathbf{y}$. In symbols:

$$A\mathbf{y} = \begin{bmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{bmatrix}.$$

For example, if

$$A = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

then

$$A\mathbf{y} = \begin{bmatrix} 1 - 1 \\ 2 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In some ways, the matrix-vector product $A\mathbf{y}$ is similar to the familiar scalar product ay. For example, we can distribute A across linear combinations of vectors:

$$A(c_1\mathbf{y}_1 + c_2\mathbf{y}_2) = c_1A\mathbf{y}_1 + c_2A\mathbf{y}_2 \tag{2}$$

for any scalars c_1, c_2 and vectors $\mathbf{y}_1, \mathbf{y}_2$. However, there are also some ways in which matrix-vector products are NOT like scalar products:

(i) It is not true that $A\mathbf{y} = \mathbf{y}A$ since the term on the right doesn't even make sense. In other words, when multiplying matrices and vectors, order matters!

(ii) Even if we know that $A\mathbf{x} = A\mathbf{y}$ for some vectors $\mathbf{x}, \mathbf{y}$, this does not necessarily imply that $\mathbf{x} = \mathbf{y}$. For example, if

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

then

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = A\mathbf{y}$$

even though clearly, $\mathbf{x} \neq \mathbf{y}$.

Because of remarks (i) and (ii), we have to use caution when solving equations involving matrices: one cannot "divide both sides" of an equation by a matrix or rearrange the order of terms. This is the analog of not being able to divide by a in the equation $ax = b$ when $a = 0$.

Back to Solving Systems of Equations

Okay with this matrix-vector multiplication notation in hand, let's confirm it matches our $A\mathbf{x} = \mathbf{b}$ jazz from earlier. We claimed that

$$\begin{aligned} x + y &= 2 \\ x - y &= 0 \end{aligned}$$

was equivalent to

$$A\mathbf{x} = \mathbf{b}$$

with

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Now we can check that

$$A\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ x - y \end{bmatrix}$$

Therefore,

$$A\mathbf{x} = \mathbf{b}$$

means

$$\begin{bmatrix} x + y \\ x - y \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

and equating the first components on each side with each other and then the second, we get our starting equations.

Before we move on to discuss determinants, one brief connection with the course. A system of equations of the form $A\mathbf{x} = \mathbf{0}$ (where $\mathbf{0}$ means a vector of all zeros) is called ***homogeneous*** whereas a system of equations $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$ is called nonhomogeneous. Homogeneous equations represent equations of lines with a y-intercept of 0. Such lines always intersect in at least one points, namely $(0, 0)$. This is called the trivial solution. For homogeneous equations, there are only two possibilities: (i) they have only the trivial solution or (ii) they have infinitely many solutions[4]. We'll know take up the task of how to distinguish between these scenarios.

Determinants

Suppose we have a generic system of two linear, *homogeneous* algebraic equations

$$\begin{aligned} a_{11}x + a_{12}y &= 0 \\ a_{21}x + a_{22}y &= 0 \end{aligned} \qquad (3)$$

[4]One member of the trinity has been abandoned.

with coefficient matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

It shouldn't be too hard to convince yourself based off our earlier discussion that there will only be non-trivial solutions if the second equation has the same slope as the first, or in other words if

$$\frac{a_{11}}{a_{12}} = \frac{a_{21}}{a_{22}}$$

(provided $a_{12}, a_{22} \neq 0$; otherwise, we can do some other equivalent comparisons to get the result that will follow). If we rearrange this expression, we obtain a formula which is valid for all a_{ij}, namely that

$$a_{11}a_{22} - a_{21}a_{12} = 0.$$

We call this number the ***determinant*** of the coefficient matrix for the system:

$$\det(A) = a_{11}a_{22} - a_{21}a_{12}.$$

For example, if

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

then

$$\det(A) = (1)(-1) - (1)(1) = -2.$$

The argument we used to get here shows that if, for a particular system of homogeneous equations, $\det(A) = 0$, then the system will have an infinite number of solutions. We call matrices with zero-determinant ***singular***. They are the matrix analog of 0 because we can't solve $A\mathbf{x} = \mathbf{0}$ via "division by A" and hence represent singularities in the universe of matrices. If $\det(A) \neq 0$, then the matrix A is called nonsingular and we can solve equations like $A\mathbf{x} = \mathbf{b}$ via "division by A".

But hold on a second. What do you mean "divide by A"? Can we divide by a matrix. Well, more precisely, we will actually multiply both sides of the equation by a special matrix called A^{-1} that will cancel out A (similar to multiplication by the reciprocal of a nonzero real number). Only matrices with $\det(A) \neq 0$ will have such inverse. The next section will define them after first covering matrix by matrix multiplication.

Matrix Multiplication and Inverse Matrices

Recall that the product $A\mathbf{x}$ for

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

can be found by taking the dot product of the successive rows of A with $\mathbf{x}$ to get a new vector

$$A\mathbf{x} = \begin{bmatrix} (1)(1) + (-1)(-1) \\ (2)(1) + (0)(-1) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Similarly, if we want to take $A\mathbf{y}$ for the same matrix A, but a different vector $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, we get a different vector

$$A\mathbf{y} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

If we form a matrix B by appending $\mathbf{x}, \mathbf{y}$ together:

$$B = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

then we define the product AB as the new 2×2 matrix obtained by appending $A\mathbf{x}, A\mathbf{y}$:

$$AB = \begin{bmatrix} A\mathbf{x} & A\mathbf{y} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$$

.

In general, if A, B are two $n \times n$ matrices, then the product AB is a new $n \times n$ matrix whose ijth element is computed as follows: take the dot product of row i in A and column j in B. For example, let's suppose that

$$A = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then AB will have a total of four dot products obtained by taking different combinations of rows, columns in A, B, respectively. For example,

$$\begin{aligned} (AB)_{11} &= (0, -1) \cdot (3, 1) \\ &= (0)(3) + (-1)(1) \\ &= -1 \end{aligned}$$

while

$$\begin{aligned} (AB)_{21} &= (2, 1) \cdot (3, 1) \\ &= (2)(3) + (1)(1) = 7. \end{aligned}$$

You can visualize the full answer and play around with examples at this online calculator `http://www.calcul.com/show/calculator/matrix-multiplication`.

To give a more rigorous definition of the matrix inverse requires the definition of matrix multiplication along with a look back at scalar multiplication from a different perspective. 1 is a very special number in scalar multiplication: multiplying any number by 1 does not change its value. For this reason, 1 is often called the multiplicative identity. Furthermore, for any real number a except $a = 0$, there is always some number a^{-1}, often called the reciprocal or inverse of a, such that $aa^{-1} = 1$.

Is there an analog for matrices? Well, to replace 1, we need a matrix I which does not change other matrices upon multiplication. If you were trying to guess a 2×2 matrix with this property, your first try might be the matrix with all 1's. But this doesn't work because, for example,

$$\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}.$$

What goes wrong? The 1s in the off-diagonal have the effect of adding the columns of A together instead of just keeping the values fixed. So instead we 0 those entries out:

$$\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}.$$

That's better! We have found our matrix analog of 1.

We always use I to denote this matrix with 1s on the diagonal and 0s elsewhere and call it the **identity** matrix. We can use I to come to an official definition of the matrix inverse: the **inverse** of a matrix A is a matrix A^{-1} such that $AA^{-1} = I$. We previously saw that the inverse of a matrix does not always exist, but when it does, we say that A is **invertible**.

Some matrices are easy to invert and multiply: those whose only nonzero entries are on the diagonal. For example, if

$$M = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

then

$$M^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix}$$

because $MM^{-1} = I$. Also note that when we multiply a matrix A by a diagonal matrix M it has the effect of scaling each row of A by the diagonal entry of M:

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 6 & 3 \end{bmatrix}.$$

By the same reasoning, multiplying by M^{-1} divides each row of A by the corresponding diagonal element in M.

There is actually a general formula for the inverse of a 2×2 matrix: if A is given by

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The division by $\det(A)$ corresponds with our earlier observation that we can really only invert matrices with nonzero determinant. The formula for A^{-1} is usually derived from scratch in an introduction to linear algebra course or textbook (again, plug for Math 75), but you can easily verify it works by showing that $AA^{-1} = I$. As an example of how to use the formula, consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Then $\det(A) = 5$ and so

$$A^{-1} = \frac{1}{5} \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{bmatrix}.$$

Solutions to $A\mathbf{x} = \mathbf{b}$

The matrix-inverse offers an alternative way to solve systems of equations. Starting with

$$A\mathbf{x} = \mathbf{b},$$

we multiply both sides on the left by A^{-1} to get

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}.$$

On the left, A^{-1} and A now cancel out so we end up with the solution vector

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

For example, to solve the system

$$\begin{aligned} x - 2y &= 3 \\ 2x + y &= 1 \end{aligned}$$

which has coefficient matrix

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

and constant vector

$$\mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

we first calculate A^{-1}:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

Then we mutliply $A^{-1}\mathbf{b}$:

$$\frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3+2 \\ -6+1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

That's actually more tedious to by hand than elimination so don't do it. Use MATLAB. First enter A and b as arrays:

```
>> A=[1 -2;2 1]

A =

     1    -2
     2     1

>> b=[3;1]

b =

     3
     1
```

Then you can either use linsolve:

```
>> x=linsolve(A,b)

x =

     1
    -1
```

the matrix-inverse:

```
>> x=A^(-1)*b

x =

     1
    -1
```

or the backslash operator (not recommended for beginners):

```
>> x=A\b

x =

     1
    -1
```

Regardless of which pill you choose, we get the right answer.

And now, back to our main story arch.

Section 3.1 : Introduction to Systems of Equations

Class objectives:

- Set up systems of first order equations to describe compartmental models.
- Solve diagonal and triangular systems of equations via substitution.
- Write systems of equations in matrix-vector form.

Introduction

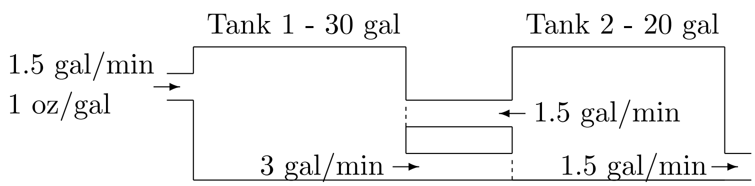
The first two chapters focused on first and second order equations describing how a single unknown quantity $y(t)$ changes over time. In this module, we will model scenarios which involve multiple dynamically evolving variables. A ***system of first order equations***, as the name suggests, is a set of two or more first order differential equations describing the rates of change for two interrelated quantities. You have already encountered systems of first order equations in Chapter 2 when we used ode45 for solve second order equations. For example, to solve the second order equation $y'' + 4y = 0$, we defined new variables $y_1 = y, y_2 = y'$ and used the fact that $y'' = y'$ to obtain two equivalent first order equations

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -4y_1 \end{aligned}$$

This is a system of first order equations in the unknowns y_1, y_2. However, conversion of second order equations is into the only important application of first order systems.

Compartmental Models

Many important models in biology, physiology, epidemiology, physics, information technology and other branches of applied math rely on our ability to set up compartmental models in which various components of a system are split up into "compartments" that receive content from each other and the outside environment. For example, suppose that we have a system consisting of two interconnected tanks. The first tank has a volume of 30 gal and the second tank has a volume of 20 gal. Both tanks are initially filled with pure water. Water containing a concentration of 1 oz/gal of salt is pumped into Tank 1 at a rate of 1.5 gal/min. The well stirred mixture flows from Tank 1 to Tank 2 at a rate of 3 gal/min. The mixture flows from Tank 2 back to Tank 1 at a rate of 1.5 gal/min and is pumped out of Tank 2 at a rate of 1.5 gal/min. The diagram below shows a graphical depiction of this information. We would like to know how the amount of salt in the two tanks changes over time and if these amounts will eventually converge to some sort of equilibrium values.



Let $y_1(t)$ and $y_2(t)$ denote the amount of salt (oz) in Tanks 1 and 2, respectively, after t minutes. Then the balancing principle yields the system of two equations

$$\frac{dy_1}{dt} = (1.5)(1) + (1.5)\left(\frac{y_2}{20}\right) - (3)\left(\frac{y_1}{30}\right)$$
$$\frac{dy_2}{dt} = (3)\left(\frac{y_1}{30}\right) - (1.5)\left(\frac{y_2}{20}\right) - (1.5)\left(\frac{y_2}{20}\right).$$

The three terms in the equation for $y_1(t)$ represent the rate at which salt flows in from the outside environment, the rate at which salt flows in from Tank 2, and the rate at which salt flows out into Tank 2. The three terms in the equation for $y_2(t)$ represent the rate at which water flows into Tank 2 from Tank 1, the rate at which water flows out from Tank 2 into Tank 1, and the rate at which water flows out from Tank 2 into the outside environment. After some minor simplifications and rearrangement of terms, we can rewrite the above system as

$$\frac{dy_1}{dt} = -0.1y_1 + 0.075y_2 + 1.5$$
$$\frac{dy_2}{dt} = 0.1y_1 - 0.15y_2. \tag{1}$$

Note that we are also given the initial conditions $y_1(0) = y_2(0) = 0$ since both tanks initially contain pure water.

Equation (1) is our first example of a *system of first order, constant coefficient nonhomogeneous equations* (SOFOCCNE): first order because they involve first derivatives of the unknown functions y_1 and y_2; constant coefficient because the coefficients of y_1', y_2', y_1, y_2 in the equations are constants; and nonhomogeneous because of the "+1.5" term in the first equation. In general, a SOFFOCNE is any system of equations which can be written in the form

$$\frac{dy_1}{dt} = a_{11}y_1 + a_{12}y_2 + b_1(t)$$
$$\frac{dy_2}{dt} = a_{21}y_1 + a_{22}y_2 + b_2(t) \tag{2}$$

for some constants $a_{11}, a_{12}, a_{21}, a_{22}$ and some functions $b_1(t), b_2(t)$. When $b_1 = b_2 = 0$, we call the system homogeneous and get a SOFOCCHE. Notice the double subscript indexing on the a's follows the pattern "equation number" followed by "variable number". For example, a_{12} is the coefficient of variable 2 in equation 1. The use of such indexing is to foreshadow the matrix-vector developments below (queue the dramatic music).

Our goals for SOFOCCNE equations are similar to our goals for first and second order equations except that now, our solution involves two functions $y_1(t), y_2(t)$ so we will be asking questions about BOTH unknowns. In the above example, for instance, it is altogether possible that the unknown concentrations in the two tanks will converge to different equilibrium values or oscillate or not converge at all. We will return to the analysis of this example in Section 3.3.

As a second example of a system of first order equations, consider the following model for combat. Two forces, which we shall call Gondor and Mordor for the sake of concreteness, are at war. If we assume that Gondorian forces kill Mordor forces at a rate proportional to the current size of the Gondor army and Mordor forces kill Gondorians at a rate proportional to the size of the Mordor army then we obtain a system of two differential equations

$$\frac{dG}{dt} = -aM$$
$$\frac{dM}{dt} = -bG$$

where $G(t)$ is the number of Gondorians (say in thousands) after t hours and $M(t)$ is the number of Mordorians (in thousands). a and b are positive constants which depend on each force's killing effectiveness: increasing a increases the killing efficiency of Mordor's soldiers while increasing b increases the effectiveness of Gondor's. You can think of b as the average number of Mordorians killed per day by the average Gondorian solider and similarly for a with Mordor and Gondor swapped. In other words, a, b are the "per capita" kill rates of an average Mordor, Gondor soldier and the overall kill rates have the form

$$\text{per capita rate} \times \text{population size.}$$

The question we will seek to answer is: which army will die out first in a battle to the death? We will also want to investigate how the answer to this question depends on a, b, and the initial numbers of Gondorians and Mordorians. An example is shown in the figure below. We shall refer to this type of figure as a **time-series** plot because it illustrates how the solutions evolve over time. In this example, the time-series plot shows that the Gondor army will die off after just 12 hours of fighting.

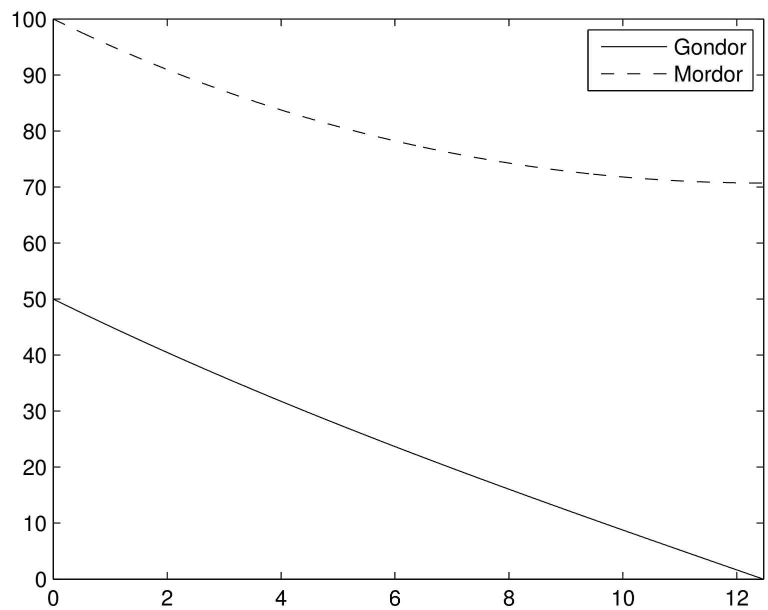
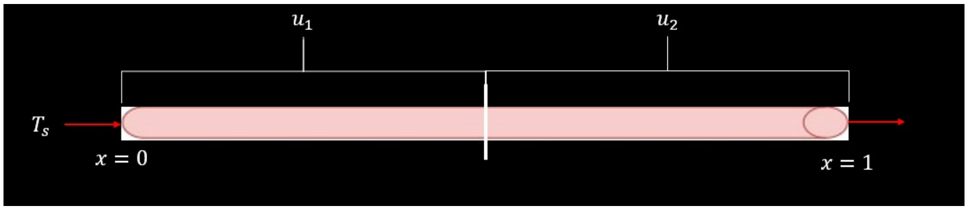


Figure 1: An example of the outcome of the Gondor/Mordor battle in the case where $G_0 = 50$, $M_0 = 100$, $a = 0.05$, and $b = 0.1$. Doubling the kill rate of Gondor's soldiers was not enough to save the humans from impending doom.

As a final example of the usefulness of a compartmental model, we turn to the world of heat transfer. Suppose we have an insulated rod, of length 1 unit, and place it in a medium with temperature T_s so that heat can only enter and leave the rod through the edges as indicated in the figure below.



To model the distribution of heat along the rod, we would technically need a function $u(t, x)$ which depends on both position along the rod x and time t. Studying evolving functions of two variables is the topic of Partial Differential Equations as we have mentioned before and when you take Math 157, you will see the heat equation in all its glory. However, we can derive a simplified version by splitting the rod into two compartments[1] labeled 1 (left) and 2 (right) below. The goal will then be to set up equations for u_1, u_2, the average temperatures in each compartment. To do so we apply Newton's Law of Cooling at the left and middle boundaries to obtain:

$$u_1' = -k(u_1 - T_s) - k(u_1 - u_2) = -2ku_1 + ku_2 + kT_s$$

and then at the middle and right boundaries to obtain:

$$u_2' = -k(u_2 - u_1) - k(u_2 - T_s) = ku_1 - 2ku_2 + kT_s$$

where $k > 0$ is a constant related to the diffusivity of the material used to make the rod. So once again, we get a first order system of equations.

A diagonal system of equations

All three examples in the previous section are complicated by the fact that they consist of two inter-related differential equations: the rate of change of unknown 1 depends on the value of unknown 2 and vica-versa. However, in order to motivate what will come in future sections and to think a bit more about what exactly we mean by "a solution" to a system of equations, consider the simpler problem of finding functions $y_1(t), y_2(t)$ such that

$$\begin{aligned} y_1' &= 2y_1 \\ y_2' &= 2y_2 \end{aligned}$$

with initial conditions $y_1(0) = 1, y_2(0) = -1$. In this case, the two differential equations are "uncoupled" and we can solve them as individual FOCCHE equations to obtain $y_1(t) = e^{2t}, y_2(t) = -e^{2t}$.

The last exercise below explores how to solve a triangular system of equations in which the first equation depends only on y_1 and can be solved directly while the second can then be solved by substitution. You'll notice that we once again obtain solutions that are exponential functions, but with two different exponents λ_1, λ_2 and potentially different coefficients.

[1] or more

However, as soon as we completely couple the equations so that the differential equation for y_1 depends on y_2 and vica-versa, such direct substitutions are not possible and we will have to dig a little deeper to uncover the right framework for discussing solutions to SOFOCCNE. Linear algebra provides that framework. Before reading the next subsection, please review the linear algebra supplement for additional terminology and notation.

Matrix-Vector Form of Systems

In Equation (2), we used that mysterious double index notation to represent an arbitrary SOFOCCNE. The mystery will now be explained as we show how to formulate such systems of equations in the proper framework of matrices and vectors.

In the matrix-vector formulation of SOFOCCNE, we treat our two solutions as components of a single ***solution vector***

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}.$$

In other words, we are now thinking of our solution as a function $\mathbf{y} : \mathbb{R} \to \mathbb{R}^2$. Its derivative $\mathbf{y}'(t)$ is the vector-valued function we get by taking derivatives of each component with respect to t:

$$\mathbf{y}'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix}$$

Using the definitions of vector addition and the matrix-vector product (see Linear Algebra Supplement) , we can now rewrite (2) in the more compact form

$$\mathbf{y}' = A\mathbf{y} + \mathbf{b}(t) \tag{3}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is called the ***coefficient matrix*** and

$$\mathbf{b}(t) = \begin{bmatrix} b_1(t) \\ b_2(t) \end{bmatrix}$$

is called the forcing vector. Of course, if our system is homogeneous, $\mathbf{b} = \mathbf{0}$ and we simply have $\mathbf{y}' = A\mathbf{y}$.

Matrix-vector notation is nice and convenient (which satisfies my innate sense of laziness), but don't forget that this single matrix-vector equation is actually shorthand for the system of equations in (2). For example, if we say that the vector

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

is a solution to the matrix-vector equation

$$\mathbf{y}' = A\mathbf{y} \tag{4}$$

with

$$A = \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix}$$

then what we actually mean is that $y_1(t)$ and $y_2(t)$ satisfy the system of equations

$$\begin{aligned} y_1' &= y_1 + 3y_2 \\ y_2' &= 5y_1 + 3y_2 \end{aligned}$$

because if we simplify the righthand side of (4), we obtain the vector

$$A\mathbf{y} = \begin{bmatrix} y_1 + 3y_2 \\ 5y_1 + 3y_2 \end{bmatrix}$$

and for this to equal the vector

$$\mathbf{y}' = \begin{bmatrix} y_1' \\ y_2' \end{bmatrix},$$

we must have $dy_1/dt = y_1 + 3y_2$ and $dy_2/dt = 5y_1 + 3y_2$.

As a second example, we can write our Gondor-Mordor system in the form $\mathbf{y}' = A\mathbf{y}$ with solution vector and coefficient matrix

$$\mathbf{y}(t) = \begin{bmatrix} G(t) \\ M(t) \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} 0 & -a \\ -b & 0 \end{bmatrix}.$$

Note that we can (and should) also denote the initial conditions as a vector. For example, if we assume that the initial sizes of the Gondor and Mordor armies are 50 and 100 respectively, then

$$\mathbf{y}(0) = \begin{bmatrix} 50 \\ 100 \end{bmatrix}.$$

We'll leave it to you to do the same for the heat transfer example. In the next Matlab Supplement, we will explore how the matrix-vector representation of systems allows us to obtain approximate solutions in Matlab. We will then discuss one of my favorite definitions of the course in Section 3.2: the notorious eigenvalue-eigenvector problem.

Exercises

(1) Water with a chemical concentration of 1 g/L flows into Tank 1 at a rate of 1 L/s and out of Tank 2 at a rate of 1 L/s. Water also flows from Tank 1 to Tank 2 at a rate of 2 L/s and from Tank 2 back to Tank 1 at a rate of 1 L/s. Tank 1 initially contains 2 L of water with a concentration of 2 g/L while Tank 2 initially contains 1 L with a concentration of 0 g/L. Write down a system of equations for the amounts of chemical in Tanks 1, 2 after t seconds.

(2) Compartmental models also arise in the study of age structured populations. For example, suppose a researcher is interested in modeling the future progression of a species of fish in which juveniles reach adulthood in, on average, 5 days. Adults gives birth to new juveniles at a per capita rate of 1 every two days and die after an average of 4 days. Write a system of differential equations describing how the number of juvenile and adult fish evolves over time. Express your system in matrix-vector form.

(3) A health body cell type divides into two new cells on average once every five days, but mutates into a tumorous cell type at an average rate of once every 10 days. Tumorous cells types divide on average once every two days, but also only have an average lifespan of 4 days. Write out a system of differential equations for modeling the quantities $h(t)$ = number of healthy cells and $s(t)$ = number of "sick" tumorous cells after t days and express your system in matrix-vector form.

(4) Determine the solutions $x_1(t), x_2(t)$ to the system of equations

$$\begin{aligned} x_1' &= 2x_1 \\ x_2' &= x_1 + x_2 \end{aligned}$$

with initial conditions $x_1(0) = 1, x_2(0) = -1$ by solving the first equation for $x_1(t)$ and then substituting your answer into the equation for $x_2(t)$.

Matlab Supplement 4: Solutions to Systems using ode45

Euler's Method has an analog for systems of equations, but since you are probably sick of that dead white dude by now, let's focus more on how to use Matlab's ode45 function in this new context. Its easier to use and more accurate at the end of the day even if it is a bit more of a black box!

We will demonstrate the techniques using the Mordor-Gondor combat model $G' = -aM, M' = -bG$ with initial conditions $G(0) = 50, M(0) = 100$ (i.e. the M army has twice as many troops as the G army) and investigate how changing a,b, effects the outcome of the battle. We will start by assuming that $a = b = 1$ which will obviously result in a Mordor victory, but will be good for a first demonstration. As stated in the last section, this system can be written in the form $y' = Ay$ with coefficient matrix

```
A = [0 -1;-1 0]


A =

     0    -1
    -1     0
```

We also enter in the initial conditions

```
y0=[50;100]


y0 =

    50
   100
```

Recall from Chapter 1 that the first three outputs of ode45 are the rate function for the DE, the time interval to solve over, and the initial conditions. With the introduction of our matrix-vector notation, these conventions still hold for systems: the rate function is now just a matrix-vector product

```
f=@(t,y) A*y


f =

  function_handle with value:

    @(t,y)A*y
```

Note that we MUST include t as an input even though our rate function does not explicitly depend on t. If we use the time interval $[0, 1]$ then we call ode45 as follows

```
[t,y]=ode45(f,[0,1],y0);
```

Before plotting the solution, let's examine the outputs t and y in more detail. t is the vector of times used by ode45. Its length tells us how many times were used

```
length(t)
```

```
ans =

    41
```

Its entries tell us the actual times used. For example, the first five are

```
t(1:5)
```

```
ans =

         0
    0.0250
    0.0500
    0.0750
    0.1000
```

The output y, on the other hand, is a matrix whose size is

```
size(y)
```

```
ans =

    41     2
```

Each row represents the solution vector at a different time and each column is a different component of the solution (Column 1 = Gondor, Column 2 = Mordor). For example,

```
y(1,:)
```

```
ans =

    50   100
```

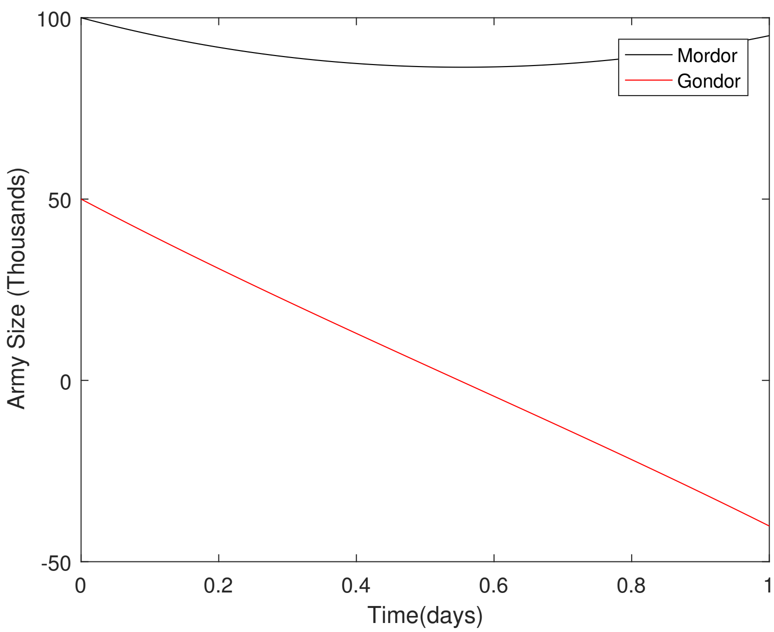
tells you the initial sizes of each army whereas the command

```
y(1:5,1)
```

```
ans =

   50.0000
   47.5154
   45.0604
   42.6337
   40.2335
```

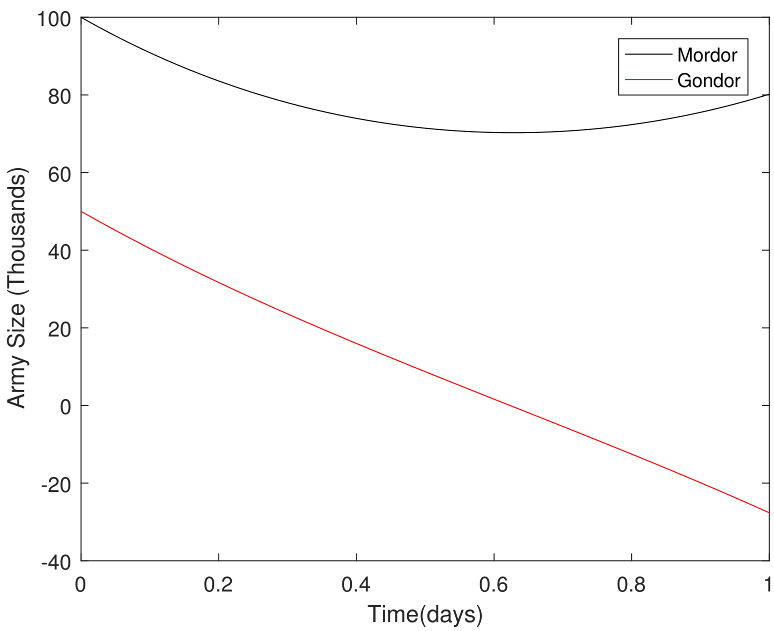
tells you the Gondor army sizes at the first five times. So when we run the command plot(t,y), we will actually see two graphs appear, one for each unknown

```
plot(t,y)
xlabel('Time(days)')
ylabel('Army Size (Thousands)')
legend('Mordor','Gondor')
```



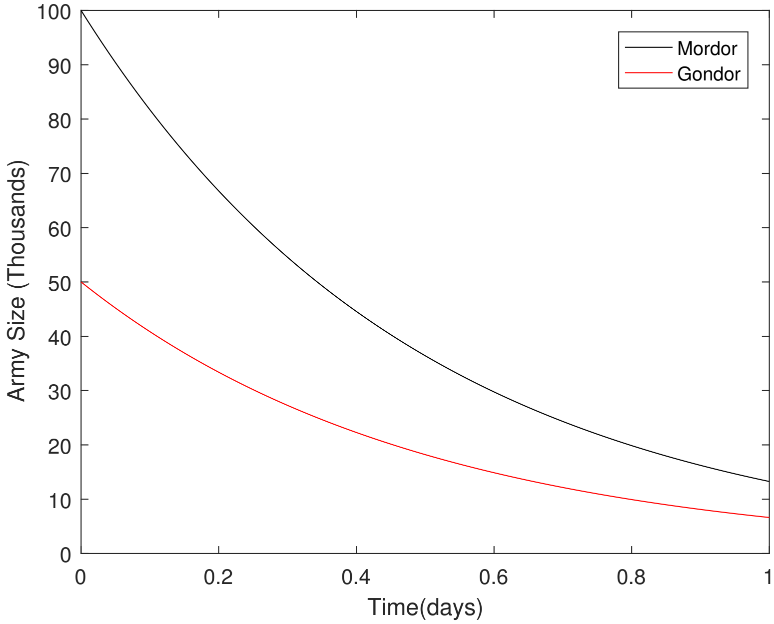
As expected, the Gondor army loses since they start out with less troops, but kill at the same rate. Let's try doubling their kill rate (b=2) in the above code

```
A = [0 -1;-2 0];
f=@(t,y) A*y;
[t,y]=ode45(f,[0,1],y0);
plot(t,y)
xlabel('Time(days)')
ylabel('Army Size (Thousands)')
legend('Mordor','Gondor')
```



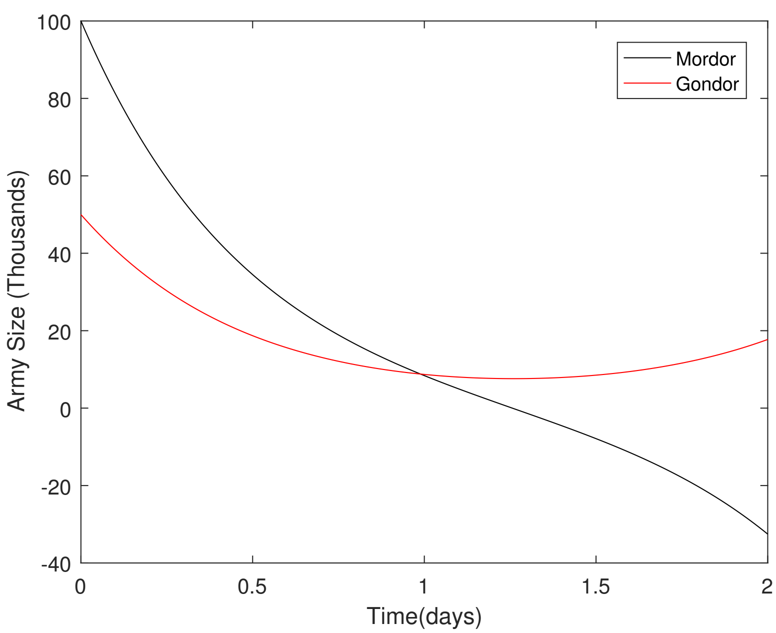
Still not looking good. How about quadrupling their kill rate? (if you get tired of cut and pasting, then write a function as in Supplement 3):

```
A = [0 -1;-4 0];
f=@(t,y) A*y;
[t,y]=ode45(f,[0,1],y0);
plot(t,y)
xlabel('Time(days)')
ylabel('Army Size (Thousands)')
legend('Mordor','Gondor')
```



Hmm, that it looking better. Now it is moving towards a draw. A little bigger kill rate and a longer time scale saves Gondor at last:

```
A = [0 -1;-4.1 0];
f=@(t,y) A*y;
[t,y]=ode45(f,[0,2],y0);
plot(t,y)
xlabel('Time(days)')
ylabel('Army Size (Thousands)')
legend('Mordor','Gondor')
```



Despite the fictional setting for the previous example, it actually illustrates a principle known as Lanchester's Square Law: in order to overcome an initial deficit of $1/r$, killing efficiency must increase by r^2. In our example, the initial size of the Gondor was 1/2 the size of the Mordor Army so they needed to be $4 = 2^2$ timess as efficient of killers to win. Notice also that after the Mordor size becomes negative, they start supplying the Gondor Army with undead warriors. This is another thing to watch out for in modeling with differential equations: models may only be valid up until a certain time. In our example, as soon as one of the armies dies out, the battle is over and the model should be terminated as well. A programming structure known as a while loop could accomplish this task, but that would take us off topic so let's move on for now.

Exercises

(1) Recall the adult-juvenile fish population model from last class (see Section 3.1, Exercise 3). Suppose that initially there are 50 juvenile and 100 adult fish.

  (a) Use ode45to plot the populations sizes of juvenile and adult fish change over time. Does the overall population size appear to be growing, fluctuating, approaching an equilibrium, or dying off?

  (b) There are five parameters in this system: the average juvenile to adult incubation period (5 days in the example), the death rate of adults (4 days), the birth rate (once every two days), and the two initial population sizes. By playing around with these values, see if you can kill the fish. Include a discussion of your thought process as well as plots from some of your trial runs.

(2) Suppose we have a system consisting of two interconnected tanks both of which hold 20 gal of water. Water containing a concentration of 1 oz/gal of salt is pumped into Tank 1 at a rate of 1 gal/min. The well stirred mixture flows from Tank 1 to Tank 2 at a rate of 2 gal/min. The mixture flows from Tank 2 back to Tank 1 at a rate of 1 gal/min and is pumped out of Tank 2 at a rate of 1 gal/min.

  (a) Set up a system of differential equations for how the amounts of salt in each container evolve over time. Write your system as a matrix-vector equation of the form $\mathbf{y}' = A\mathbf{y} + \mathbf{b}$ where $\mathbf{b}$ is a vector.

  (b) Create a plot which shows how the amounts of salt in both tanks evolve over time. What are the limiting values?

Section 3.2: The Eigenvalue-Eigenvector Problem

Class objectives:

- Define and compute eigenvalue-eigenvector pairs of 2×2 matrices.

Solutions to SOFOCCHE

Writing a system of equations in the matrix-vector form

$$\mathbf{y}' = A\mathbf{y} \tag{1}$$

is about more than just convenience: it also suggests a form of solutions. By analogy with the first order equation $y' = ay$ which had solution $y(t) = e^{at}c$, we might also expect that $\mathbf{y}' = A\mathbf{y}$ has a solution of the form $e^{At}\mathbf{v}$ where $\mathbf{v}$ is now a vector. This actually does turn out to be correct in some sense, but we have to start smaller (what does e to a matrix power mean?). As a simple example, suppose that

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Then the system $\mathbf{y}' = A\mathbf{y}$ simplifies to the two equations $y_1' = 2y_1, y_2' = 2y_2$ which we showed in the last section has solution $y_1(t) = c_1e^{2t}, y_2(t) = c_2e^{2t}$ where c_1, c_2 depend on the initial conditions. We can combine y_1, y_2 into the single solution vector

$$\mathbf{y}(t) = \begin{bmatrix} c_1e^{2t} \\ c_2e^{2t} \end{bmatrix} = e^{2t}\mathbf{v}$$

where $\mathbf{v} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. So for this example, the solution really is $e^{\lambda t}\mathbf{v}$ where λ is a constant that takes the place of A (in this case $\lambda = 2$).

Of course, the above example was easy because A had a very specific diagonal structure which lead to two uncoupled equations for y_1, y_2. How does this play out in more complicated situations?

Finding solutions to systems of SOFOCCHE equations

To motivate the next development in our story, recall that we began our study of SOCCHE by plugging $e^{\lambda t}$ into the equation $my'' + \gamma y' + ky = 0$ and using the resulting characteristic equation $m\lambda^2 + \gamma\lambda + k = 0$ to determine the values of λ which correspond to exponential solutions. We shall start at a similar place here. Suppose that we seek a solution to (1) of the form

$$\mathbf{y}(t) = e^{\lambda t}\mathbf{v} = e^{\lambda t}\begin{bmatrix} x \\ y \end{bmatrix} \tag{2}$$

where $\lambda, x,$ and y are constants that we need to determine. If we substitute this candidate solution into (1), we obtain the condition

$$\lambda\mathbf{v}e^{\lambda t} = A\mathbf{v}e^{\lambda t}$$

dividing both sides by $e^{\lambda t}$ (we cannot divide by $\mathbf{v}$ since it is a vector) leads to the equation

$$\lambda \mathbf{v} = A\mathbf{v}$$

This equation is special and gets a German name.

Definition 1. *(Eigenvalue/Eigenvector Problem) Let A be a given matrix. If*

$$A\mathbf{v} = \lambda \mathbf{v} \tag{3}$$

for some nonzero vector $\mathbf{v}$, then λ is called an **eigenvalue** *of A and $\mathbf{v}$ is called the corresponding* **eigenvector**.

A few remarks are in order

- After World War II, many math books refused to use the German name and instead called these objects "characteristic" values and vectors after a loose translation of the word "eigen". I guess we gave up on that.
- Eigenvectors, by definition, must be nonzero. If we did not include this restriction, everything would be an eigenvalue since $A\mathbf{0} = \lambda \mathbf{0}$ for any λ and there is no point in defining something if it is always the same.
- If we can find one eigenvector $\mathbf{v}$ corresponding to an eigenvalue λ, then we can find infinitely more if we multiply $\mathbf{v}$ by any nonzero constant c. To justify this:

  $$A(c\mathbf{v}) = cA\mathbf{v} = c\lambda\mathbf{v} = \lambda(c\mathbf{v})$$

  which means that $c\mathbf{v}$ is an eigenvector of A.
- I love this definition so much that I would actually get it tattooed on my arm if my partner allowed me to get math tattoos.

Solving eigenvalue problems is important in a wide range of different disciplines (in fact, when you tune a guitar you are solving an eigenvalue/eigenvector equation!) In this class, our main concern will be using eigenvalues and eigenvectors for finding solutions to SOFOCCHE. We will return to this task in the next section. First, we will focus on developing an algorithm for computing eigenvalues and eigenvectors for general 2×2 matrices.

Finding eigenvalues and eigenvectors

As a first example, suppose that

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Then the equation $A\mathbf{v} = \lambda \mathbf{v}$ yields the system of two equations:

$$2x = \lambda x$$
$$2y = \lambda y$$

for the x, y components of $\mathbf{v}$. Since at least one of x and y must be non-zero, it must be the case that $\lambda = 2$. But then we obtain the system of equations $x = x$ and $y = y$ which can be satisfied by ANY numbers x and y. Therefore, A only has one eigenvalue $\lambda = 2$, but ANY nonzero $\mathbf{v}$ is an eigenvector.

Computing eigenvalue/eigenvector pairs in the case where A is diagonal reveals some insight into the definition: a diagonal matrix

$$A = \begin{bmatrix} \mathrm{a} & 0 \\ 0 & \mathrm{b} \end{bmatrix}$$

has eigenvalues a, b and if we multiply A by a vector $\mathbf{y} = \begin{bmatrix} x \\ y \end{bmatrix}$, the result is

$$A\mathbf{y} = \begin{bmatrix} ax \\ by \end{bmatrix}.$$

In other words, A scales the x component by a and the y component by b. That is why a and b are the eigenvalues (the scaling factors) and $\mathbf{v}_1$ and $\mathbf{v_2}$ are the corresponding eigenvectors (scaling directions).

The computation of eigenvalues and eigenvectors gets more complicated (and interesting) when A has nonzero off-diagonal entries. For example, suppose that our matrix of coefficients is

$$A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

(A matrix of this form appeared in our combat model from last section). By definition, an eigenvalue/eigenvector pair $(\lambda, \mathbf{v})$ means a solution to the system of equations

$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

or equivalently,

$$\begin{aligned} -y &= \lambda x \\ -x &= \lambda y \end{aligned}$$

Moving all variables to one side, we can rewrite this as the system of two equations

$$\begin{aligned} \lambda x + y &= 0 \\ x + \lambda y &= 0 \end{aligned} \tag{4}$$

which should look somewhat familiar EXCEPT for the fact that we actually have three unknowns here: λ, x, y.

To make sense of this all, note in the language of the linear algebra supplement, we can treat the system of equations in (4) above as a system of linear algebraic equations in the unknowns (x, y) if we define a coefficient matrix

$$A_\lambda = \begin{bmatrix} \lambda & 1 \\ 1 & \lambda \end{bmatrix}$$

which depends on λ. Therefore, we know that if we choose a λ for which $\det(A_\lambda) \neq 0$, then the only solution $\mathbf{v} = (x, y)$ to the system will be the trivial solution $\mathbf{v} = (0, 0)$ which is not allowed to be an eigenvector. We can conclude that the only eigenvalues of the system will be the values of λ for which

$$\det(A_\lambda) = \lambda^2 - 1 = 0.$$

Solving this equation yields two solutions $\lambda_1 = 1, \lambda_2 = -1$ (which we will often write in shorthand as $\lambda_{1,2} = \pm 1$).

Now that we have found the eigenvalues, we need to find the eigenvectors. To do so, we let the eigenvalues take turns playing the role of λ in Equation (4). For example, setting $\lambda = \lambda_1 = 1$ simplifies things to

$$\begin{aligned} x + y &= 0 \\ x + y &= 0 \end{aligned}$$

The solution to this equation is the set of all points (x, y) on the line $y = -x$. So any vector of the form

$$t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where $t \neq 0$ would be an eigenvector corresponding to λ_1.

A similar calculation reveals that the eigenvectors corresponding to $\lambda_2 = -1$ are all vectors of the form

$$t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where $t \neq 0$. We often summarize this information by saying that A has two sets of eigenvalue-eigenvector pairs. In the first set, we pair $\lambda_1 = 1$ with any nonzero multiple of $\vec{v_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and in the second set, we pair $\lambda_2 = -1$ with any nonzero multiple of $\vec{v_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

General procedure for computing eigenvalues and eigenvectors

Let's review the procedure above by seeing how it applies to an arbitrary 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

When solving for eigenvalues and eigenvectors, we are looking for solutions $\lambda, \mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ to the eigenvalue-eigenvector problem

$$A\mathbf{v} = \lambda \mathbf{v}.$$

Writing this out in component form implies that λ, x, y must satisfy the system of equations:

$$\begin{aligned} a_{11}x + a_{12}y &= \lambda x \\ a_{21}x + a_{22}y &= \lambda y. \end{aligned}$$

Combining like terms simplifies this to

$$(a_{11} - \lambda)x + a_{12}y = 0$$
$$a_{21}x + (a_{22} - \lambda)y = 0.$$

From our brief aside into linear systems, we know that this system will have nontrivial solutions if and only if the determinant of the matrix

$$A_\lambda = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$$

is 0. We call the equation $\det(A_\lambda) = 0$ the **characteristic equation** of A.

Notice that the matrix A_λ above can just be obtained by subtracting λ from the diagonal elements of A. Recalling from the supplement that the identity matrix I is defined as

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

we can alternatively write the characteristic equation as

$$\det(A - \lambda I) = 0.$$

This is the notation typically used throughout linear algebra and we will conform to the standard from now on. Once we solve the characteristic equation for λ, we then successively plug any such solutions back into the definition $A\mathbf{v} = \lambda\mathbf{v}$ and solve for the corresponding eigenvectors. To summarize,

(i) Subtract λ from the diagonal elements of A. Call the resulting matrix $A - \lambda I$.

(ii) Determine all values of λ for which $\det(A - \lambda I) = 0$. These are your eigenvalues

(iii) Sequentially substitute each eigenvalue λ into the expression $A\mathbf{v} = \lambda\mathbf{v}$ and solve for $\mathbf{v}$. The set of all nonzero solutions is the set of all eigenvectors corresponding to λ.

Note that if we actually write out the terms in the characteristic equation for an arbitrary matrix A, we get the expression

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0.$$

The reason for taking the time to do so is simply to point out that for 2×2 matrices, the characteristic equation is always a quadratic and hence, will always have either two real, one real, or two complex roots. The practice problems below will explore some of these different scenarios and we highly recommend you do as many of them as possible to ensure that you are comfortable with each option. In the next section, we'll tie all this back into solutions of SOFOCCHE.

Exercises

(1) Calculate all eigenvalue-eigenvector pairs for each coefficient matrix below.

(a)
$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b)
$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(c)
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

(d)
$$A = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix}$$

(e)
$$A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}$$

(f)
$$A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$$

(g)
$$A = \begin{bmatrix} 2 & -5 \\ 1 & -2 \end{bmatrix}$$

(2) Explain why a 2×2 matrix will always have either two real, one real, or two complex conjugate eigenvalues.

(3) A 2×2 skew symmetric matrix is a matrix of the form

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

where a, b are constants. Calculate the eigenvalues and eigenvectors of A in terms of a, b.

(4) Find an expression for the eigenvalues of the matrix

$$A = \begin{bmatrix} -2 & 2 \\ \alpha & 0 \end{bmatrix}$$

in terms of α and state the ranges of α for which

(a) A has two negative real eigenvalues

(b) A has one negative and one positive real eigenvalue

(c) A has two complex eigenvalues

Section 3.3: Eigenvalues, Eigenvectors, and SOFOCCHE

Class objectives:

- Use eigenvalues and eigenvectors to find solutions of SOFOCCHE
- Describe the relationship between eigenvalues/eigenvectors and the behavior of solutions to SOFOCCHE

In the last section, we showed that finding solutions $\mathbf{y}(t) = e^{\lambda t}\mathbf{v}$ to the SOFOCCHE

$$\mathbf{y}' = A\mathbf{y} \tag{1}$$

involves finding eigenvalue/eigenvector pairs $(\lambda, \mathbf{v})$ for the coefficient matrix A and we discussed an algorithm for computing these German-named gems. In this section, we will tie all this back in with the meat and potatoes of this chapter: solving initial value problems for SOFOCCHE. After discussing a few examples, we will see how the asymptotic behavior of solutions to SOFOCCHE can be guessed just by looking at the resulting eigenvalues.

Solutions SOFOCCHE via Eigen-shit

The idea is similar to the Method of Characteristic Equations for solving second order equations[1]. There, we obtained specific exponential solutions $e^{\lambda t}$ by finding roots to the quadratic equation

$$m\lambda^2 + \gamma\lambda + k = 0;$$

For systems, we saw in the last notes how to find specific exponential solutions by solving the quadratic equation

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0.$$

If this quadratic has two real or two complex roots, then we obtain two linearly independent[2] solutions $\mathbf{v}_i e^{\lambda_i t}$ with different exponents λ_i. Appealing to the principle of superposition[3] yields the general solution

$$\mathbf{y}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t}$$

and we can then plug in initial conditions to solve for c_1, c_2 in a specific IVP[4]. The situation of one real root is more complicated and will not be dealt with in generality here[5].

[1]In fact, later we will see that they are the same!

[2]Recall that linearly independent solutions were defined for scalar valued functions as two solutions which are not scalar multiples of each other; the same definition applies to vector valued functions

[3]Although we only justified this principle for single first and second order equations, it can be extended to systems of linear equations using the linearity of matrix multiplication - see Section 5.1 in Edwards and Penney for a detailed proof.

[4]Yes, we are leaving out some details here such as "are there situations where we cannot solve for c_1, c_2?". One can show as long as the two eigenvalues are distinct, you can always solve the resulting system of equations - we again refer the reader to Section 5.1 of Edwards and Penney for details.

[5]We will tell you that the general solution ends up having the form $\mathbf{y}(t) = c_1\mathbf{v}_1 e^{\lambda t} + c_2(\mathbf{u}t + \mathbf{w})e^{\lambda t}$ for some $\mathbf{u}, \mathbf{v}$ where λ in the single eigenvalue. We will refer you to Section 5.4 in (surprise) Edwards and Penney for details. And yes, we realize we are going crazy with footnotes here

Let's look at an example. Suppose that our matrix of coefficients is

$$A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

and we wish to determine the corresponding solution $\mathbf{y}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ to the homogeneous system (1) with initial conditions $x(0) = 1$ and $y(0) = 2$. In Section 3.2, we showed that A has eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = -1.$$

with the corresponding representative eigenvectors:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Therefore, we know that two solutions to our system of differential equations are

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t, \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

We now combine our two solutions using the principle of superposition to obtain a general solution

$$\mathbf{y}(t) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

to (1). Remember, this means that in component form, we have two solutions

$$\begin{aligned} x(t) &= c_1 e^t + c_2 e^{-t} \\ y(t) &= -c_1 e^t + c_2 e^{-t}. \end{aligned}$$

To solve for c_1, c_2, we plug the initial conditions $x(0) = 1$ and $y(0) = 2$ to yield a system of two equations

$$\begin{aligned} 1 &= c_1 + c_2 \\ 2 &= -c_1 + c_2 \end{aligned}$$

for the unknowns c_1 and c_2 which we can solve to obtain $c_1 = -1/2$ and $c_2 = 3/2$. Therefore,

$$\mathbf{y}(t) = -\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t + \frac{3}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

is the vector valued solution to our differential equation of interest.

When first exposed to this idea, many aspiring students confuse $y_1(t)$ with the first part of the solution

$$-\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t$$

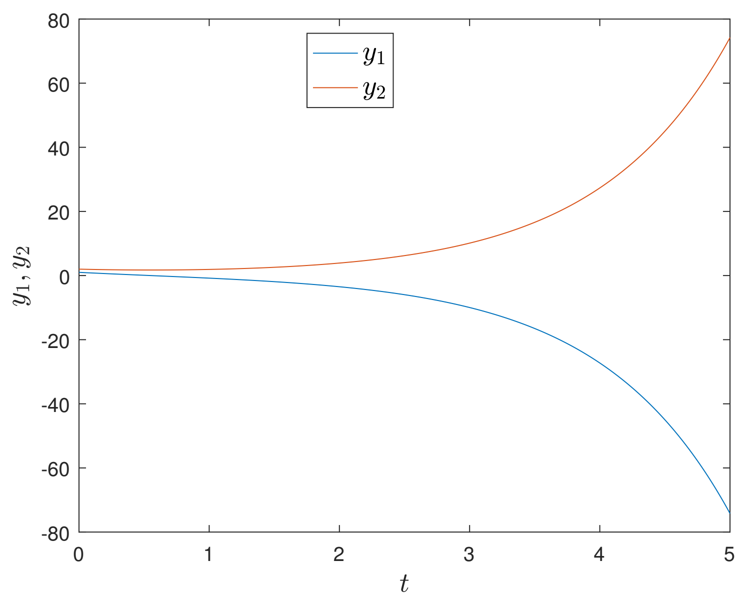
or sometimes, with the first eigenvalue-eigenvector pair. But this is thinking about the equation for $\mathbf{y}(t)$ in the wrong direction. $\mathbf{y}$ has two **rows** and the first **row** combines information from both eigenvectors to give $y_1(t)$ as a function[6]. So in our example, we actually have

$$y_1(t) = -\frac{1}{2}e^t + \frac{3}{2}e^{-t}$$

and

$$y_2(t) = \frac{1}{2}e^t + \frac{3}{2}e^{-t}$$

as the two components of the solution are shown in the figure below. Notice that y_1 diverges towards $-\infty$ while y_2 diverges towards ∞ due to the signs of e^t in both formulas. The astute reader may recognize this as an example of the combat model from lab in which both armies have equal kill rates $a = b = 1$, but army two starts out with twice as many troops. The graph shows the destruction of Army 1 in this scenario; the more general scenario is included in the supplementary section that follows this one.



The best way to start getting some feel for what's going on here is to do some examples so we recommend completing Exercise 1 below before reading the next section. As you progress, think about how the behavior of the solutions depends on the eigenvalues of the matrix A. For instance, in the last example, the eigenvalues of $1, -1$ contributed terms of the form e^t and e^{-t} to the solution components $y_1(t)$ and $y_2(t)$. The e^{-t} terms died out as t increased, but the e^t terms exploded quickly. So we would see similar behavior in the above example for other initial conditions as well (as long as they don't correspond to $c_1 = 0$ in the general solution - See Exercise 2 below).

Eigenvalues and Asymptotic Behavior

One of our (at least my) goals is always to develop simple rules for describing the general behavior of equations without too much tedious algebra. In the FOCCHE $y' = ay$, we can

[6]We will redundantly point this out again in the Matlab Supplement.

simply look at the sign of a to determine if the solution is going to grow, decay, or stay constant over time. For SOCCHE, we can look at the roots of the characteristic equation

$$m\lambda^2 + \gamma\lambda + k = 0$$

to determine if the solution is overdamped (exponential decay), critically damped (almost exponential decay), underdamped (decaying oscillations), or undamped (constant oscillations).

Eigenvalues provide a similar method of classification for solutions to SOFOCCHE. To see why, suppose that a matrix A has two distinct eigenvalues λ_1 and λ_2 so that all solutions are of the form

$$\mathbf{y}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t}.$$

Written out in component form, this becomes:

$$\mathbf{y}(t) = \begin{bmatrix} a_1 e^{\lambda_1 t} + b_1 e^{\lambda_2 t} \\ a_2 e^{\lambda_1 t} + b_2 e^{\lambda_2 t} \end{bmatrix}$$

for some constants a_1, a_2, b_1, b_2 that will depend on the initial conditions and the eigenvectors. Using the "largest exponent wins" rule, we can determine the behavior of solutions from this expression. For example, suppose that $\lambda_1 > \lambda_2 > 0$ (i.e. A has two positive eigenvalues). Then the behavior of $y_i(t)$ will be determined by the sign of a_i:

- If $a_i > 0$, then $y_i(t) \to \infty$.
- If $a_i < 0$, then $y_i(t) \to -\infty$.
- If $a_i = 0$, then $y_i(t)$ goes to 0 or $\pm\infty$ depending on the whether the sign of b_i is 0 or $\pm$.

So there are only three possible types of behavior for each component in this scenario. If we wanted a one sentence summary, we would say that "typical" solutions tend to diverge at exponential rates when the eigenvalues of A are positive.

In contrast, suppose that A was two complex conjugate eigenvalues $a \pm bi$ where $a, b > 0$. Then, after all the smoke clears, the solution vector will be something of the form

$$\mathbf{y}(t) = \begin{bmatrix} e^{at}(a_1 \cos bt + b_1 \sin bt) \\ e^{at}(a_2 \cos bt + b_2 \sin bt) \end{bmatrix}.$$

Therefore, both components will exhibit growing amplitude oscillations as $t \to \infty$. Further possibilities are explored in the exercises below.

So in the context of differential equations, the eigenvalues are the growth exponents associated with a given matrix. Just like the scalar valued equation $y' = ay$ with led to solution $y(t) = ce^{at} = e^{at}c$, the vector valued equation

$$\mathbf{y}' = A\mathbf{y}$$

leads to the solution

$$\mathbf{y}(t) = c_1\mathbf{v_1} e^{\lambda_1 t} + c_2\mathbf{v_2} e^{\lambda_2 t}$$

as long as A has two distinct eigenvalues λ_1, λ_2.

To facilitate solutions to more complicated problems, Matlab Supplement 5 will discuss the computation of eigenvalues and eigenvectors in Matlab before we move on to discuss Higher Order Spring-Mass systems in Section 3.4.

Exercises

1. For each of the matrices in Exercise 1 from Section 3.2 (excluding e), determine the solution to equation (1), once with initial conditions $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and once with initial conditions $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Sketch a plot of the resulting solutions and describe the long term behavior.

2. Find an initial condition vector $\mathbf{y}_0$ to the SOFOCCHE with coefficient matrix
$$A = \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix}$$
which will force $\mathbf{y}(t) \to \mathbf{0}$ as $t \to \infty$.

3. Determine the solutions $x(t), y(t)$ to the system of equations
$$\begin{aligned} x' &= -2x + y \\ y' &= x - 2y \end{aligned}$$
with the initial conditions $x(0) = 1, y(0) = -3$.

4. The value of commodity x (which could be negative) decreases at a rate proportional to the value of a second commodity y (which can also be negative), but the value of y increases at a rate proportional to the value of x with the same constant of proportionality k.
   (a) Write down a system of differential equations for the values $x(t)$ and $y(t)$ of each commodity at time t and determine the eigenvalues/ eigenvectors of the coefficient matrix for this. Your answers should depend on k.
   (b) Suppose that $k = 4$ and $x(0) = 1, y(0) = 0$. Find formulas for and sketch graphs of $x(t)$ and $y(t)$. Include Matlab Validation.

5. Determine the solutions $x(t), y(t)$ to the system of first order equations
$$\begin{aligned} x' &= -4x - 3y \\ y' &= 3x - 4y \end{aligned}$$
with initial conditions $x(0) = 1, y(0) = 0$ **and** sketch a graph of $x(t)$.

6. Determine the solution vector $\mathbf{y}(t)$ to the system of equations $\mathbf{y}' = A\mathbf{y}$ where
$$A = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}, \ \mathbf{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
and make a time-series plot of the components.

7. Generalizing the last two exercises, suppose that a SOFOCCHE has coefficient matrix A of the form
$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$
where a, b are constants. Write out the general solution and describe how the behavior depends on a, b.

8. Determine a classification of possible behaviors of solutions to $\mathbf{y}' = A\mathbf{y}$ in the following scenarios:

   (a) The eigenvalues of A are both negative.

   (b) The eigenvalues of A have opposite signs.

   (c) The eigenvalues of A are pure imaginary (i.e. of the form $\pm ib$).

   (d) The eigenvalues of A are complex valued with negative real part.

   (e) The eigenvalues of A are complex valued with positive real part.

   For each scenario, make a plot (sketch or Matlab) of what a "typical" solution might look like.

9. Consider the system of equations

$$\begin{aligned} \frac{dx}{dt} &= -x + \alpha y \\ \frac{dy}{dt} &= -\alpha x \end{aligned}$$

   where α is a constant.

   (a) Determine a formula for the eigenvalues of the coefficient matrix for this system in terms of α.

   (b) For which value of α will the solutions $x(t)$ and $y(t)$ exhibit decaying oscillations?

   (c) Explain why there are no values of α which would lead to constant oscillations.

Matlab Supplement 5: Eigen-shit in Matlab

The eig command

Computing eigenvalues and eigenvectors for 2×2 systems by hand isn't too bad, but the process gets tedious for higher dimensional systems (does anyone like to factor cubic polynomials?) These notes will give you a brief introduction to using Matlab as a tool for computing eigenvalues/vectors and some intuition about dealing with higher dimensional systems. Let's start with the adult-juvenile fish example from Supplement 4. If we let x, y denote the sizes of the juvenile and adult fish populations, then x, y satisfied the DEs

$$x' = -x/5 + y/2, \; y' = x/5 - y/4.$$

This system has coefficient matrix

```
A=[-1/5 1/2;1/5 1/4]

A =

   -0.2000    0.5000
    0.2000    0.2500
```

The initial conditions were given as

```
y0=[100;50]

y0 =

   100
    50
```

The eig command will compute eigenvalues

```
eig(A)

ans =

   -0.3631
    0.4131
```

If you want the eigenvectors as well, then you must save the output as [evecs,evals]

```
[V,L]=eig(A)

V =

   -0.9507   -0.6320
    0.3101   -0.7750


L =

   -0.3631         0
         0    0.4131
```

The eigenvalues now appear on the diagonals of L and the corresponding eigenvectors are stored in the columns of V. For example, the first eigenvalue is

```
L(1,1)

ans =

   -0.3631
```

and a corresponding eigenvector is

```
V(:,1)

ans =

   -0.9507
    0.3101
```

Note that the representative eigenvectors were chosen so that they have length 1.

```
sqrt(V(1,1)^2+V(2,1)^2)

ans =

    1.0000
```

If you do an example by hand and check your answers against Matlab, just keep this in mind.

Now we know that the general solution (using Matlab notation) will be

$$c(1) * V(:,1) * \exp(L(1,1) * t) + c(2) * V(:,2) * \exp(L(2,2) * t).$$

To find $c(1), c(2)$, we plug in $t = 0$ to obtain the system of equations

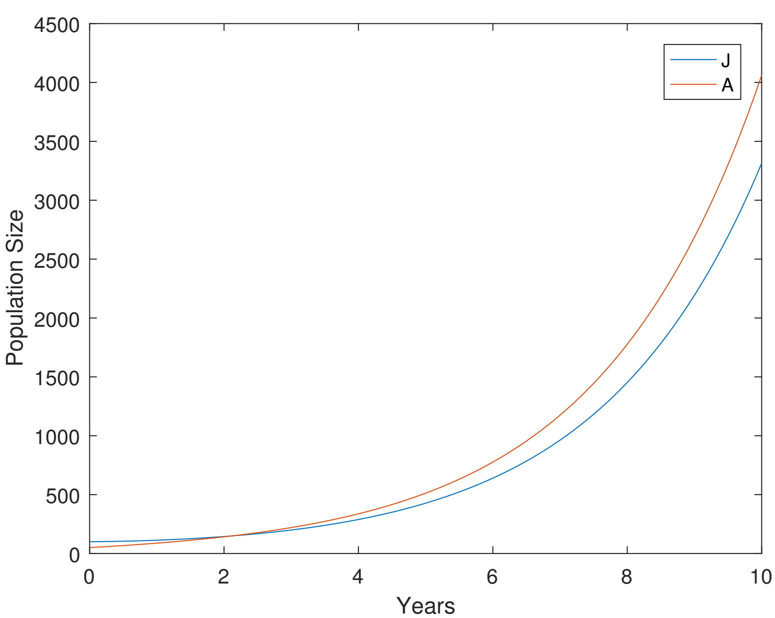
$$y0 = c(1) * V(:,1) + c(2) * V(:,2).$$

In Matlab, you can solve a linear system of equations of this form using the linsolve command:

```
c=linsolve(V,y0)

c =

  -49.2052
  -84.2098
```

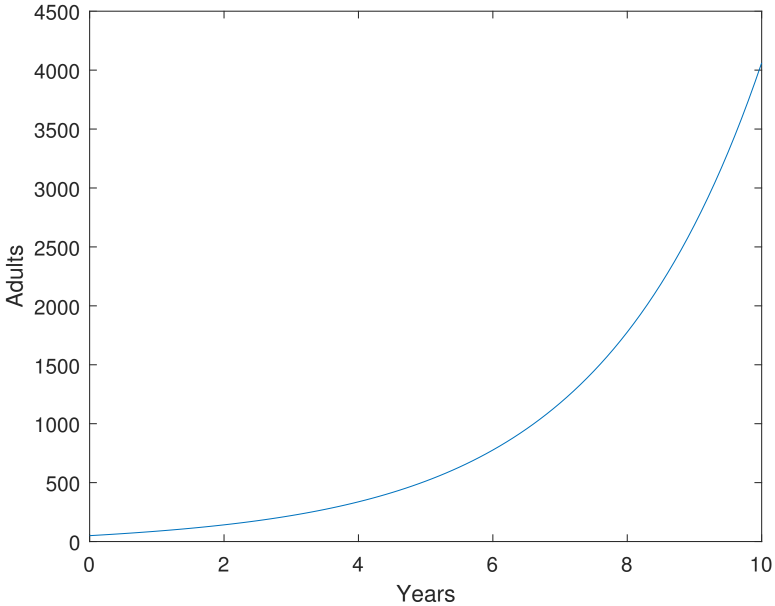
Note that in this output, the first number is $c(1)$ and the second gives $c(2)$. We can then plot out solutions over a time interval of our choosing (I choose you 10):

```
t=0:.1:10;
y=c(1)*V(:,1)*exp(L(1,1)*t) + c(2)*V(:,2)*exp(L(2,2)*t);
plot(t,y)
xlabel('Years')
ylabel('Population Size')
legend('J','A')
```



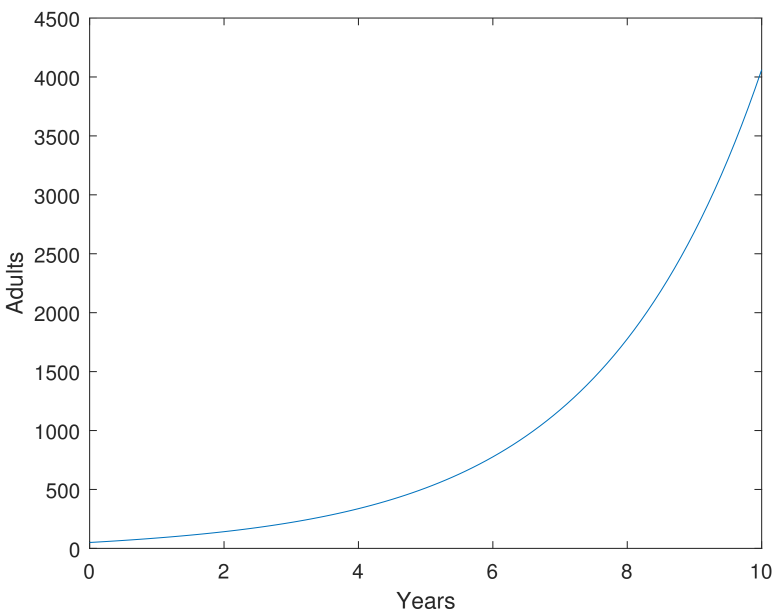
Note that if we want to ONLY plot the adults, say, then we would use the second row of y

```
plot(t,y(2,:))
xlabel('Years')
ylabel('Adults')
```



Or going back a step further, only the second row of V

```
t=0:.1:10;
y=c(1)*V(2,1)*exp(L(1,1)*t) + c(2)*V(2,2)*exp(L(2,2)*t);
plot(t,y)
xlabel('Years')
ylabel('Adults')
```



BOTH eigenvalues and eigenvectors contribute to the adults and the juveniles, but only the second components of the eigenvectors contribute to the juveniles. This may sound like a trivial observation, but many aspiring students have been dashed against the rocks for this mixup.

Higher Dimensional Example

Let's suppose that we add another life stage to a population model. Suppose an owl population can be split up into three life stages: juveniles (who neither mate nor reproduce), subadults (who mate but don't reproduce), and adults (who mate and reproduce). Juveniles become subadults after an average span of six months, subadults turn into adults after an average span of 1 year, and adults have an average lifespan of 10 years. Living female adults give birth to an average number of 2 new juveniles every 5 years. Assuming that about half of all owls adults are female and that initially there are 100 juveniles, 50 subadults, and 200 adults, we can write out an initial value problem for j, s, a = number of each type after t years:

$$j' = -2j + a/5,\ s' = 2j - s,\ a' = s - a/10$$

with $j(0) = 100, s(0) = 50, a(0) = 200$. This system has coefficient matrix and initial condition vector

```
A=[-2 0 1/5;2 -1 0;0 1 -1/10]
y0=[100; 50; 200]
```

```
A =

   -2.0000         0    0.2000
    2.0000   -1.0000         0
         0    1.0000   -0.1000


y0 =

   100
    50
   200
```

The eigen-shit can be found by

```
[V,L]=eig(A)
```

```
V =

  -0.2375 + 0.0621i  -0.2375 - 0.0621i   0.0943 + 0.0000i
   0.8061 + 0.0000i   0.8061 + 0.0000i   0.1749 + 0.0000i
  -0.5356 - 0.0554i  -0.5356 + 0.0554i   0.9801 + 0.0000i


L =
```

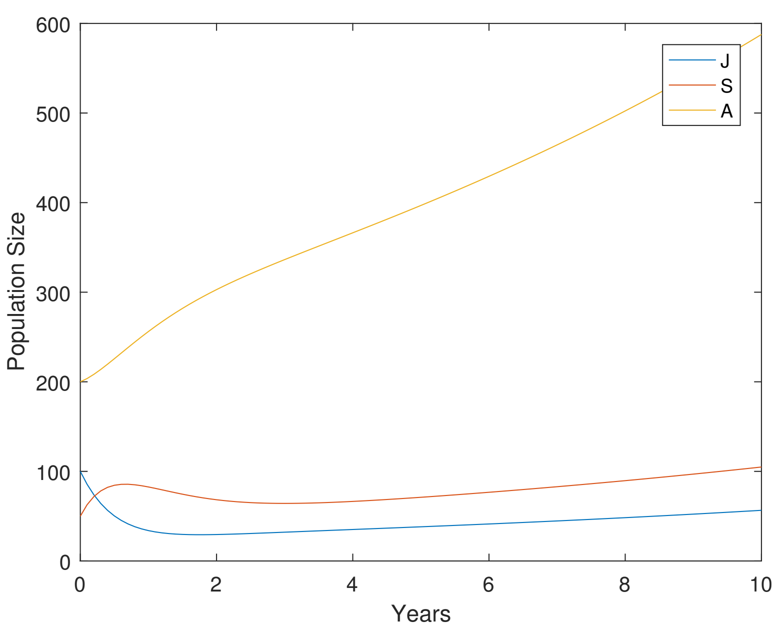
```
  -1.5892 + 0.1540i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i  -1.5892 - 0.1540i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0785 + 0.0000i
```

Its a mess, but you can say a few things right off the bat: A has three eigenvalues, one of which is a small positive number and two of which are complex conjugates with negative real part. Although we have not explicitly stated any results along these lines, you can probably guess by analogy that the general solution is of the form

$$c(1) * V(:,1) * \exp(L(1,1) * t) + c(2) * V(:,2) * \exp(L(2,2) * t) + c(3) * V(:,3) * \exp(L(3,3) * t)$$

The first two terms in this sum will exhibit decaying oscillations while the third will exhibit exponential growth so as a whole, the owl population will grow. To confirm, let's calculate the c coefficients and plot our solution over a 10 year period

```
c=linsolve(V,y0);
t=0:0.1:10;
y=c(1)*V(:,1)*exp(L(1,1)*t)+c(2)*V(:,2)*exp(L(2,2)*t)+
    c(3)*V(:,3)*exp(L(3,3)*t);
plot(t,y)
xlabel('Years')
ylabel('Population Size')
legend('J','S','A')
```



The warning message above appears due to approximation errors within Matlab. It doesn't actually compute the exact eigenvalues and hence treats the third one as if it had a really small imaginary part. For more on how to compute eigenvalues numerically, take an advanced linear algebra or numerical analysis class or google "The QR Algorithm".

If you get tired of writing out all three terms in y, a for loop can help

```
y=c(1)*V(:,1)*exp(L(1,1)*t);
for i=2:3
    y=y+c(i)*V(:,i)*exp(L(i,i)*t);
end
```

This doesn't save any time for a 3×3 example, but try writing out the solution for a 10×10!

Exercises

(1) Demidogs evolve in four stages: larvae, polliwogs, catigorgons, and finally demidogs. Larvae become polliwogs after an average span of 30 days, polliwogs turn into catigorgons after an average span of three days, catigorgons become demidogs after an average span of two days, and demidogs have an average lifespan of one year. Demidogs also spawn an average number of five new larvae every 2 days. Assume that there are currently 20 larvae, 5 polliwogs, 1 catigorgon, and 0 demidogs in our world.

(a) Write a matrix-vector differential equation for the numbers of each stage which will be present in t days. Use Matlab to plot and calculate how many (adult) demidogs there will be after 7, 30, and 365 days. (I recommend three separate plots).

(b) Look at the eigenvalues of your coefficient matrix from part a. How do their values validate or contradict the graphs you saw in part a?

(2) Suppose we have three species competing for survival. Every member of Species 1 eats one Species 2 member per day, every 2 eats one 3 per day, and every 3 eats one 1 per day (this is called Rock-Paper-Scissors model for obvious reasons). Let $y_1(t), y_2(t), y_3(t)$ be the number of each species (in thousands) alive after t days.

(a) What would be the coefficient matrix A in the matrix-vector equation $\mathbf{y}' = A\mathbf{y}$ modeling these species interactions? Use Matlab to compute its eigenvalues and discuss what they suggest about the long term behavior of this system.

(b) Experiment around with different initial conditions to find values which lead to a victory for Species 1. Then find different values which lead to victory for Species 2 and then Species 3. Include plots to demonstrate each scenario. **Warning**: Once one of the species dies, the competition is over. For example, if 1 dies first, then 2 wins because there is no one left to eat them or be eaten by 3.

Section 3.3b - The Battle for Middle Earth Revisited

We have returned to all of the examples from Section 3.1 except our Gondor vs. Mordor combat model. In the Matlab Supplement, we observed that in order to overcome an initial troop deficit of 1/2, an army needed to quadruple its kill rate. In this section, we'll use eigen-shit to verify this observation algebraically. Will this be on the test? No. Is it madness? Maybe. But we will do it nonetheless.

The Mordor-Gondor combat model can be formulated in matrix-vector form as

$$\mathbf{y}' = A\mathbf{y}$$

where

$$A = \begin{bmatrix} 0 & -b \\ -a & 0 \end{bmatrix}, \mathbf{y}(t) = \begin{bmatrix} g(t) \\ m(t) \end{bmatrix}.$$

Suppose we have initial conditions $g(0) = g_0, m(0) = kg_0$ where $k > 1$ is some constant representing the initial size ratio of Mordor to Gondor. The question is: how much bigger does a have to be than b to push Gondor to victory? The eigenvalues of A above are given by

$$\lambda^2 - ab = 0 \Longrightarrow \lambda = \pm\sqrt{ab}$$

and the corresponding eigenvectors are nonzero multiples of

$$\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \mp\sqrt{a} \end{bmatrix}$$

so the general solution is

$$\mathbf{y}(t) = c_1 \begin{bmatrix} \sqrt{b} \\ -\sqrt{a} \end{bmatrix} e^{\sqrt{ab}t} + c_2 \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix} e^{-\sqrt{ab}t}.$$

Plugging in the initial conditions gives the system of equations

$$\begin{aligned} g_0 &= \sqrt{b}c_1 + \sqrt{b}c_2 \\ kg_0 &= -\sqrt{a}c_1 + \sqrt{a}c_2 \end{aligned}$$

which has solution

$$c_2 = \frac{g_0}{2\sqrt{ab}}\left(\sqrt{a} + \sqrt{b}k\right), c_1 = \frac{g_0}{2\sqrt{ab}}\left(\sqrt{a} - \sqrt{b}k\right).$$

(That's why this won't be on the test.) Using these values in the general solution and doing some algebra yields the formulas

$$\begin{aligned} g(t) &= \frac{g_0}{2\sqrt{a}}\left[\left(\sqrt{a} + \sqrt{b}k\right)e^{\sqrt{ab}t} + \left(\sqrt{a} - \sqrt{b}k\right)e^{-\sqrt{ab}t}\right] \\ m(t) &= \frac{g_0}{2\sqrt{b}}\left[-\left(\sqrt{a} + \sqrt{b}k\right)e^{\sqrt{ab}t} + \left(\sqrt{a} - \sqrt{b}k\right)e^{-\sqrt{ab}t}\right] \end{aligned}$$

for the sizes of the armies after t days of combat.

Now in order for $g(t)$ to win, $m(t)$ must reach zero which implies that

$$-\left(\sqrt{a}+\sqrt{b}k\right)e^{\sqrt{ab}t}+\left(\sqrt{a}-\sqrt{b}k\right)e^{-\sqrt{ab}t}=0$$

for some value of t. In other words, there must be a t for which

$$e^{\sqrt{ab}t}=\frac{\sqrt{a}-\sqrt{b}k}{\sqrt{a}+\sqrt{b}k}.$$

Since the left side is positive, this will happen if and only if $\sqrt{a}-\sqrt{b}k>0$ or equivalently

$$a/b>k^2.$$

This is great (if you are a mathematician; bad if you are Gondor). It means that if Mordor has $k=2$ as many troops, then Gondorians will need to kill Mordorians at a rate of $k^2=4$ times as fast.

The real world moral is that it is better to double your army than double your efficiency! The mathematical moral is that Matlab can help avoid some messy algebra...

Section 3.4: Equilibrium Solutions, the Phase-Plane, and Matrix Exponentials

Class objectives:

- Define and compute equilibrium solutions for non-homogeneous systems
- Graph solutions in the phase-plane
- Assess the stability of equilibrium solutions using eigenvalues
- Define and write solutions in terms of Matrix Exponentials.

A SOFOCCNE can be written in matrix-vector form as

$$\mathbf{y}' = A\mathbf{y} + \mathbf{b}(t) \tag{1}$$

where $\mathbf{b}(t)$ is some nonzero, vector-valued function of t. For example, the mixing tanks model

$$\begin{aligned} x_1' &= -0.1x_1 + 0.075x_2 + 1.5 \\ x_2' &= 0.1x_1 - 0.15x_2. \end{aligned}$$

from the beginning of the Section 3.1 notes is non-homogeneous and can be written in the form of Equation (1) with

$$\mathbf{y}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad A = \begin{bmatrix} -0.1 & 0.075 \\ 0.1 & -0.15 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1.5 \\ 0 \end{bmatrix}$$

Moving from homogeneous to non-homogeneous systems is analogous to moving from homogeneous to non-homogeneous first order equations: we find a particular and complementary solution; then combine them using the Principle of Superposition. The Method of Undetermined Coefficients and Variation of Parameters both have analogs for systems of equations[1], the only catch being that the algebra can get tedious. In this section we will just discuss one important case, namely when $\mathbf{b}$ is a constant vector. Our purposes for doing so is first of all, to complete solutions to the "trio" of modeling problems introduced in Section 3.1 by showing how to solve the mixing tanks problem above, but second (and more importantly), to introduce you to the important concepts of equilibrium solutions and phase-plane diagrams.

Equilibrium Solutions

An **equilibrium** solution to the system displayed in Equation (1) is any vector $\mathbf{y}$ which satisfies the equation $\mathbf{y}' = \mathbf{0}$ or equivalently, for which

$$A\mathbf{y} + \mathbf{b} = \mathbf{0} \Longleftrightarrow A\mathbf{y} = -\mathbf{b}.$$

In Matlab, we can solve for $\mathbf{y}$ by either using 'linsolve(A,-b)' or multiplying $-A^{-1}\mathbf{b}$ (assuming A is nonsingular). By hand, it is easier just to write out the system of equations and eliminate

[1] See, for example, Section 5.6 in Edwards and Penney.

variables. For example, in the tank model above we can find an equilibrium solution by setting $x_1' = 0, x_2' = 0$ and solving the resulting system of equations for x_1, x_2 -

$$\begin{aligned} 0 &= -0.1x_1 + 0.075x_2 + 1.5 \\ 0 &= 0.1x_1 - 0.15x_2. \end{aligned}$$

This gives the solution $x_1 = 30, x_2 = 20$ or written as a vector,

$$\mathbf{y}_p = \begin{bmatrix} 30 \\ 20 \end{bmatrix}$$

The reason for using the subscript 'p' in the above equation is that an equilibrium solution is, by definition, a solution to the differential equation and therefore, we know that the general solution to Equation (1) will be

$$\mathbf{y}(t) = \mathbf{y}_c(t) + \mathbf{y}_p$$

where $\mathbf{y}_c(t)$ is the general solution to the complementary equation $\mathbf{y_c}' = A\mathbf{y}_c$, This solution is in turn given by

$$\mathbf{y_c}(t) = c_1\mathbf{v}_1e^{\lambda_1 t} + c_2\mathbf{v}_2e^{\lambda_2 t}$$

where $(\lambda_i, \mathbf{v}_i)$ are eval-evec pairs for A. The general solution to the SOFOCCNE is then

$$\mathbf{y}(t) = \begin{bmatrix} 30 \\ 20 \end{bmatrix} + c_1\mathbf{v}_1e^{\lambda_1 t} + c_2\mathbf{v}_2e^{\lambda_2 t}$$

We could plug in the initial conditions to find c_1, c_2, but you can check here that both eigenvalues λ_1, λ_2 are negative which means that regardless, we will always have

$$\mathbf{y}(t) \to \begin{bmatrix} 30 \\ 20 \end{bmatrix}$$

as $t \to \infty$. This confirms that the amounts of salt in the two tanks will converge to equilibrium values of $30, 20$, respectively, which makes sense since these are the tank volumes multiplied by the incoming concentration.

The Phase-Plane and Stablity

In general, if we can find an equilibrium solution $\mathbf{y}_p$ to a nonhomogeneous system $\mathbf{y}' = A\mathbf{y}+\mathbf{b}$, then the general solution will be

$$\mathbf{y}(t) = \mathbf{y}_p + c_1\mathbf{v}_1e^{\lambda_1 t} + c_2\mathbf{v}_2e^{\lambda_2 t}$$

where $(\lambda_i, \mathbf{v}_i)$ are eval-evec pairs for A. Therefore, determining if solutions converge to equilibrium depends on what happens to the exponents above as $t \to \infty$. For example, if $\lambda_1, \lambda_2 < 0$, then both exponential terms converge to 0 and hence, all solutions, regardless of starting values will converge to equilibrium. In this case, we say that the equilibrium $\mathbf{y}_p$ is *asymptotically stable*[2] If, on the other hand, either λ_1 or $\lambda_2 > 0$, then some solutions which start near the equilibrium will diverge away towards $\pm\infty$ in both components and in this case, we say that the equilibrium solution is *unstable*.

[2].

There are two ways to visualize stable vs. unstable equilibria. The first, is by looking at typical time-series plots of solution components vs. time as we have been doing all semester long. But a second way is to plot solutions in the *phase-plane* where we get rid of t and instead plot our solution vector

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

in the xy plane with the first component y_1 on the x-axis and the second component y_2 on the y-axis. An example comparing these two types of plots for the mixing tanks problem is shown in Figure 1. The plot of a specific solution in the phase-plane is often referred to as a *solution* or *integral* curve for the system of equations.

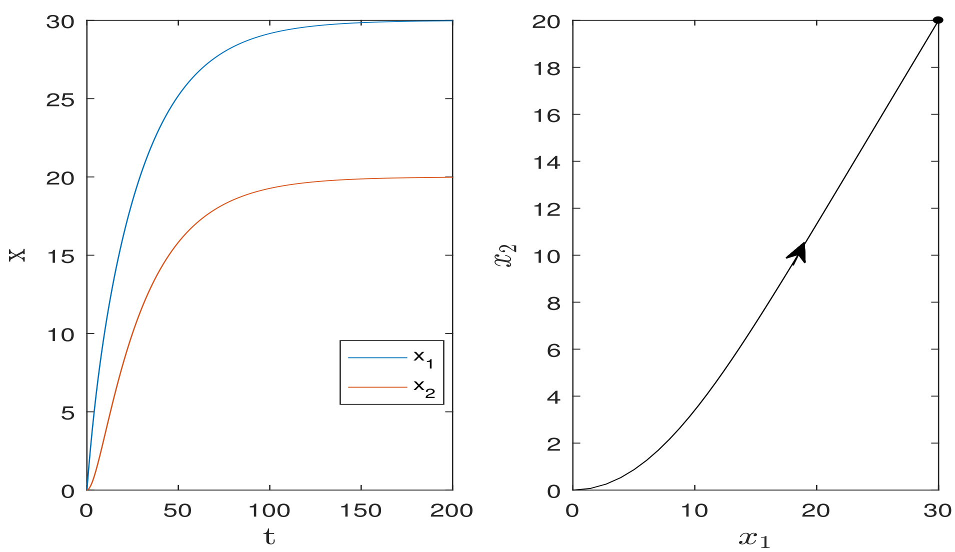
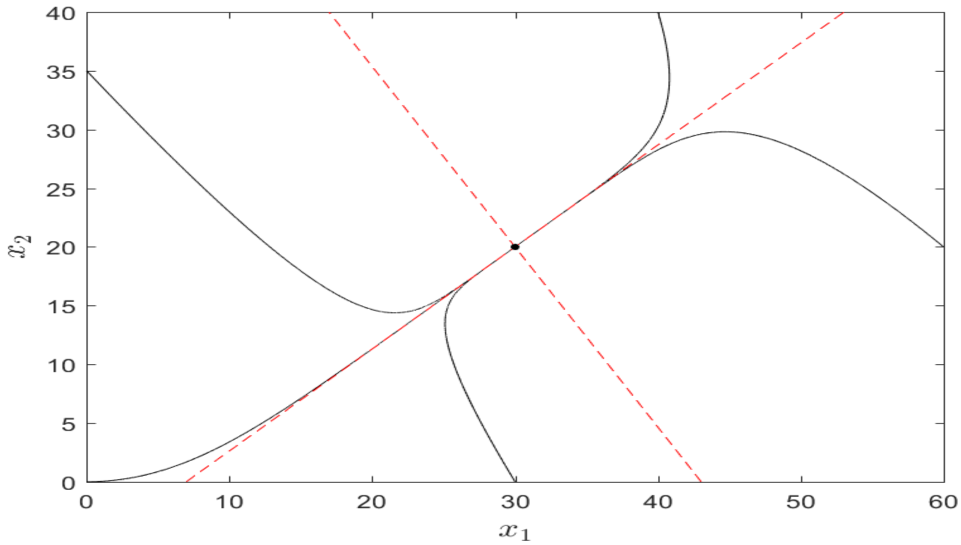


Figure 1: (Left) Time-series plot of solutions to the tank problem with initial conditions $x_1(0) = x_2(0) = 0$. (Right) Phase-plane plot of solution

Notice that in the phase-plane portrait, the solution vector $\mathbf{x}(t)$ converges to equilibrium tangent to a line. What is this line? Well recall from Calc 3 that any vector $\mathbf{v} = (v_1, v_2)$ can be used to define a slope by taking the ratio of its x and y components: $m = v_2/v_1$. So let's plot the two lines whose slopes are determined by the eigenvectors of the coefficient matrix A, but which pass through the equilibrium solution (dashed red lines) along with several solutions $(x_1(t), x_2(t))$ with different initial conditions.



Notice that ALL solutions, regardless of where they start, approach equilibrium along the same eigen-line (I made that term up just now)! This is because, according to Matlab, the eigen-shit for the coefficient matrix is

```
V =

    0.7550   -0.5458
    0.6557    0.8379


L =

   -0.0349         0
         0   -0.2151
```

so the first eigenvalue $\lambda_1 \approx -0.0349$ is larger than the second eigenvalue $\lambda_2 \approx -0.2151$. This means that the term $c_2\mathbf{v}_2 e^{\lambda_2 t}$ decays much faster than the first term $c_1\mathbf{v}_1 e^{\lambda_1 t}$ which means for large t,

$$\mathbf{y}(t) - \mathbf{y}_p \approx c_1\mathbf{v}_1 e^{\lambda_1 t}.$$

The right side of this equation will graph a line with slope-vector $\mathbf{v}_1$ as t increases towards ∞.

The exercises below will ask you to explore phase diagrams for some other systems with real eigenvalues, but phase-plane portraits of complex eigenvalue system get even more interesting. For example, consider the system $x' = -y, y' = x$ which has one equilibrium at $(0,0)$ and eigenvalues $\pm i$. A sample solution is shown in Figure 2 below in both time and phase-plane space. Notice that the constant amplitude oscillations in the time-series plot lead to concentric orbits in the phase plane. Interestingly, one would get the same plot for the system $x' = -4y, y' = 4x$ if we use the same initial conditions, but in reality, the solution curves in this latter example would be spinning around the circle at a faster rate (see exercise below). The case where the eigenvalues have non-zero real part is even more interesting; see Figure 3. The decaying oscillations lead to spirals that move towards the equilibrium of $(0,0)$.

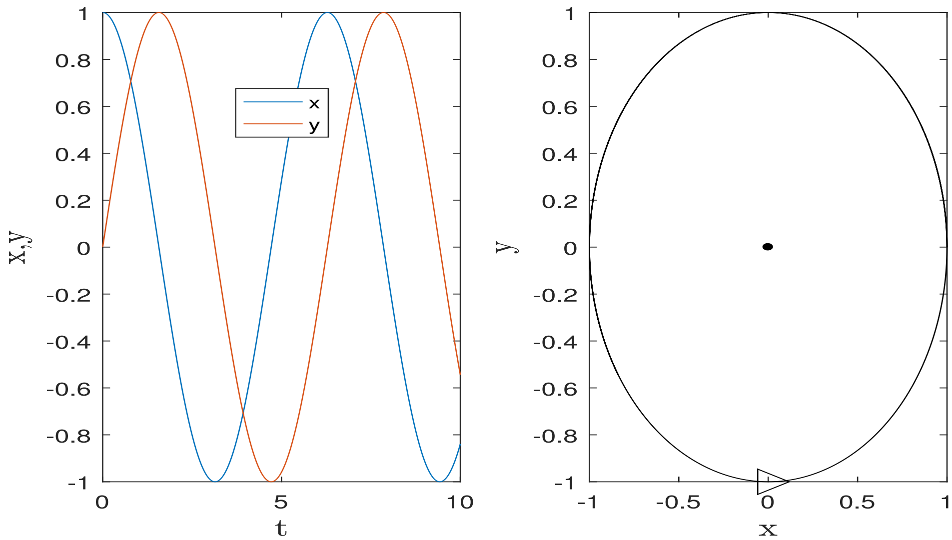


Figure 2: (Left) Time-series plot of solutions to $x' = -y, y' = x$ with initial conditions $x(0) = 1, y(0) = 0$. (Right) Phase-portrait plot of solution

The Matrix Exponential

This course has been dominated by the prevalence of exponential functions. They first popped up as solutions to FOCCHE

$$y' = ay.$$

Then, after introducing the complex exponential, we were able to obtain general solutions to all SOCCHE:

$$my'' + \gamma y' + ky = 0.$$

Finally, after converting SOFOCCHE into matrix vector form

$$\mathbf{y}' = A\mathbf{y}$$

we again obtained exponential solutions based on the e-shit of A. But there is something slightly unsatisfying about having to write the general solution as

$$\mathbf{y}(t) = \sum_{i=1}^{n} c_i \mathbf{v}_i e^{\lambda_i t}$$

in this latter case. In analogy to the solution of $y' = ay$ being $y(t) = y_0 e^{at}$, it would be nice if we could write the solution to $\mathbf{y}' = A\mathbf{y}$ as $\mathbf{y}(t) = e^{At}\mathbf{y}_0$ (note that multiplying in the opposite order doesn't make sense). In this section, we will explore this idea by introducing the matrix exponential and discussing its connections to eigenvalues, eigenvectors, and matrix diagonalization.
Recalling our development of the complex exponential, it seems natural to define the matrix exponential in terms of Taylor series as

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

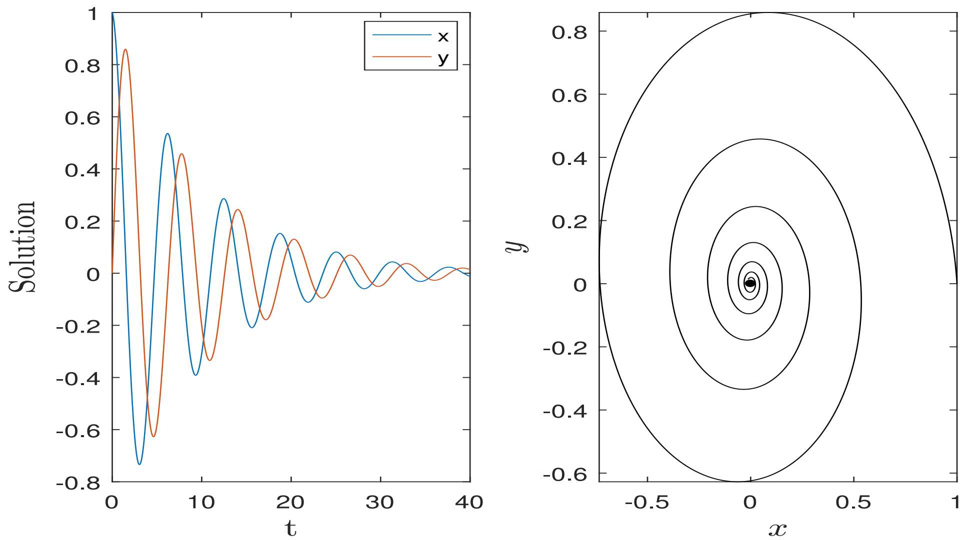


Figure 3: (Left) Time-series plot of solutions to $x' = -0.1x - y, y' = x - 0.1y$ with initial conditions $x(0) = 1, y(0) = 0$. (Right) Phase-portrait plot of solution

Let's see if this makes sense by looking at each term on the righthand side. First, when $n = 0$, we get A^0. Just as $a^0 = 1$ for any real number a, we will define $A^0 = I$ for any matrix A. Second, for $n = 1$ we just get $A^1 = A$. So far so good. Third, for $n = 2$, we get A^2 which we define by matrix multiplication as AA. After that, we define A^n iteratively as $A^{n-1}A$. So the term on the right will be a sum of matrices that all make sense and as long as the sum converges, we seem to be in good shape.

As an example of how to compute e^A, suppose that A is diagonal:

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

for some real numbers a, b. Then

$$A^n = \begin{bmatrix} a^n & 0 \\ 0 & b^n \end{bmatrix}$$

(please confirm this to yourself as well) and hence,

$$e^A = \sum_{n=0}^{\infty} \begin{bmatrix} \frac{a^n}{n!} & 0 \\ 0 & \frac{b^n}{n!} \end{bmatrix} = \begin{bmatrix} e^a & 0 \\ 0 & e^b \end{bmatrix}.$$

This shows that for diagonal matrices, the matrix exponential simply takes the exponential of the diagonal entries.

The fact stated at the end of the last paragraph is not true, however, for more general matrices. For example, try taking the very simple matrix

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Then

$$A^2 = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}, \; A^3 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$$

and in general

$$A^n = \begin{bmatrix} 1 & 2n \\ 0 & 1 \end{bmatrix}.$$

Plugging this into the definition of a matrix exponential,

$$e^A = \sum_{n=0}^{\infty} \begin{bmatrix} \frac{1}{n!} & \frac{2n}{n!} \\ 0 & \frac{1}{n!} \end{bmatrix} = \begin{bmatrix} e^1 & 2e^1 \\ 0 & e^1 \end{bmatrix}.$$

The situation gets even worse for matrices in which BOTH off-diagonal elements are nonzero.

A better way of approaching matrix exponentials is through eigenvalues. We will develop all this in the case where A is 2×2 just to avoid tedious notation, but everything we do carries over to higher dimensional cases as well. We will also assume throughout that A has two distinct eigenvalues λ_1, λ_2 with corresponding evecs $\mathbf{v}_1, \mathbf{v}_2$. Then we know by definition that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for $i = 1, 2$. We can combine these two equations into a single matrix equation by saying

$$\begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 \end{bmatrix}.$$

Now suppose we want to factor both sides of this equation. The left side is by definition

$$AV$$

where V is the matrix whose columns are evecs. On the right, if we define D to be the matrix with the evals on the diagonal (think output of 'eig' in matlab), then we can also factor the right as

$$VD$$

so we obtain the clean equation

$$AV = VD.$$

If V is invertible[3], then we can multiply both sides of the above by V^{-1} to obtain

$$A = VDV^{-1}.$$

This is called the diagonalization or sometimes, spectral decomposition of A.

Why is this product useful? Well, let's go back to our definition of the matrix exponential. We want to compute terms of the form A^n, for example A^2. But if we have obtained a diagonalization of A as VDV^{-1}, then

$$A^2 = AA = VDV^{-1}VDV^{-1} = VD^2V^{-1}$$

since $VV^{-1} = I$. D is diagonal so we know that D^2 just means we square both elements on the diagonal of D. Furthermore, its easy to check that

$$A^n = VD^nV^{-1}$$

[3]This always ends up being the case if the evals are distinct, but you have to take Math 75 to find out more!

for any n so

$$e^A = \sum_{n=0}^{\infty} \frac{VD^nV^{-1}}{n!} = V\left(\sum_{n=0}^{\infty} \frac{D^n}{n!}\right)V^{-1} = Ve^DV^{-1}.$$

Recall that e^D for diagonal D means we just take the exponential of the entries in D.

To apply this, suppose that

$$A = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}.$$

Please check that A has eigenvalues $\lambda_1 = 3, \lambda_2 = -2$ with corresponding eigenvectors $\mathbf{v}_1 = (-2, 1), \mathbf{v}_2 = (1, 2)$. Therefore,

$$e^A = \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} e^3 & 0 \\ 0 & e^{-2} \end{bmatrix} \left(\begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}\right)^{-1}.$$

Using the formula for a matrix-inverse (see LA Supplement):

$$\left(\begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}\right)^{-1} = \frac{1}{-4-1}\begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -2/5 & 1/5 \\ 1/5 & 2/5 \end{bmatrix}$$

so multiplying all three matrices out yields

$$e^A = \begin{bmatrix} -\frac{4}{5}e^2 + \frac{1}{5}e^{-3} & -\frac{2}{5}e^2 + \frac{2}{5}e^{-3} \\ -\frac{2}{5}e^2 + \frac{2}{5}e^{-3} & \frac{1}{5}e^2 + \frac{4}{5}e^{-3} \end{bmatrix}.$$

Connection to Systems of DEs

Yeah, I admit that even I didn't like doing that last calculation. Luckily, Matlab has a nifty 'expm' function which automatically computes the matrix exponential for us in examples so I could have just done

```
>> expm(A)

ans =

    0.5951    2.0683
    2.0683    8.8682
```

The more important takeaway is that we can use the matrix exponential as a nice way of writing solutions to SOFOCCHE. In particular, if $\mathbf{y}' = A\mathbf{y}$, then

$$\mathbf{y}(t) = e^{At}\mathbf{y}_0$$

as promised. And since $e^{At} = Ve^{Dt}V^{-1}$, we also know that

- If the evals of A are negative[4], then $e^{At} \to 0$ as $t \to \infty$.
- If the evals of A are positive, then e^{At} diverges as $t \to \infty$ (meaning all its entries go to $\pm\infty$).

In others words, e^{At} is very much like e^{at} if we think of replacing A by its evals.

[4] Actually, its enough for the real part to be negative.

Exercise

(1) Recall the mixing tanks problem from Exercise 1 in Section 3.1. You previously solved this problem using Euler's Method.

(a) Find the equilibrium solution for the system of equations.

(b) Determine the general solution to the system of equations (using eig in Matlab to do so is acceptable) and explain why the concentrations in both tanks will converge to 1 g/L as $t \to \infty$ for any initial conditions.

(c) Determine the solution corresponding to 0 initial concentrations in both tanks and sketch both a time-series and phase-plane plot of your solution (in Matlab). Which of the two eigen-lines for the system gives the solutions' asymptote as $t \to \infty$?

(2) Consider the system of equations $x' = -4y, y' = 4x$ with initial conditions $x(0) = 1, y(0) = 0$.

(a) You actually solved this system in a previous homework. Find out where and rewrite the solution.

(b) Create both a time-series and phase-plane plot of your solution (in Matlab). How does it compare with Figure 2 above?

(3) Consider the system of linear DEs

$$\begin{aligned} x' &= 4x - 3y \\ y' &= 3x + 4y \end{aligned}$$

(a) You found the general solution to this system in a previous homework as well (although it was in terms of a a, b). Find out where and then use that result to calculate the solution with initial conditions $x(0) = 1, y(0) = 0$.

(b) Create both a time-series and phase-plane plot of your solution (in Matlab). Explain the shape of the resulting graphs by appealing to the eigenvalues of the coefficient matrix.

(4) Suppose that $\mathbf{y}$ is the solution to the 2×2 nonhomogeneous system of equations $\mathbf{y}' = A\mathbf{y} + \mathbf{b}$ for some constant vector $\mathbf{b}$. Sketch a graph of some sample solution curves in the phase-plane for each of the following examples.

(a) A has one positive and one negative real eigenvalue.

(b) A has two complex eigenvalues with positive real part.

(c) A has one negative and one zero eigenvalue.

(5) Suppose that the system $\mathbf{y}' = A\mathbf{y} + \mathbf{b}$ has equilibrium solution $\mathbf{y}_p = (1, 0)$ and that A has eigenvalues $\lambda_1 = 1, \lambda_3 = 3$ with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

(a) Sketch (by hand) a phase-portrait showing the eigenlines and some possible solution curves.

(b) Suppose your solution curve passes through the point $(2, 0)$. What will be the limiting behavior of the x and y components in $\mathbf{y}(t)$?

(6) Show that if $\det(A) \neq 0$, then the only equilibrium solution to a homogeneous system $\mathbf{y}' = A\mathbf{y}$ is $\mathbf{y} = (0, 0)$. What could happen if $\det(A) = 0$?

(7) Suppose that the coefficient matrix A for a linear system of equations $\mathbf{y}' = A\mathbf{y}$ with solution vector $\mathbf{y}(t) = (x(t), y(t))$ has two distinct eigenvalues λ_1, λ_2 with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2$. For each scenario listed in (a) and (b) below, write down an example of specific values for $\lambda_1, \lambda_2, \mathbf{v}_1, \mathbf{v}_2$ which would yield the desired behaviors

(a) The system has solution curves which converges to 0 tangent to the line $y = x$ as $t \to \infty$.

(b) All solution curves exhibit constant amplitude oscillations with period 6π.

(8) Suppose that

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Compute e^A in two ways

(a) Directly from the definition.

(b) Using ‘expm’ in Matlab

Section 3.5: Systems of Second Order Equations

Section objectives:

- Convert systems of second order equations into systems of first order equations.
- Solve the resulting systems of equations using Matlab and interpret the output

We have gone through the process of solving SOFOCCHE[1] using eigenvalues and eigenvectors and have scratched the surface of dealing with non-homogeneous terms. In this section we will see how a simple change of variables trick allows us to convert systems of second order equations into systems of first order equations. The advantage of this trick is that we could technically do away with all of Chapter 2 and just consider it all an application of Chapter 3 which creates a nice unifying aspect to all we have done thus far. The disadvantage is that it actually makes calculations more cumbersome because you double the number of equations you have to deal with. Therefore, we will limit our attention here to one or two key examples and then develop some better methods in the following section. We begin with the familiar example of a spring-mass system.

Converting SOCCHE to SOFOCCHE

Consider a typical SOCCHE of the form

$$my'' + \gamma y' + ky = 0.$$

We have already pointed out some similarities between finding solutions to such equations and our current pursuit of solutions to SOFOCCHE (we even have a "characteristic equation" in both cases). Now we will see why they are so similar.

What we are going to do is define two new variable $y_1 = y$ and $y_2 = y'$. The reason why this is a reasonable thing to try is because a second order equation has two degrees of freedom for defining initial conditions, one for the displacement and one for the velocity, just like a system of two first order equations has two degrees of freedom for defining $y_1(0)$ and $y_2(0)$. So really, a SOCCHE can be thought of as two FOCCHE: one for position and one for velocity. To make this more concrete, note that

$$y_1' = y' = y_2.$$

Furthermore, using the definition of y_2 and our SOCCHE:

$$y_2' = y'' = -\frac{\gamma}{m}y' - \frac{k}{m}y = -\frac{\gamma}{m}y_2 - \frac{k}{m}y_1.$$

Combining the last two equations yields the system

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -\frac{k}{m}y_1 - \frac{\gamma}{m}y_2 \end{aligned} \tag{1}$$

[1] At least as long as we can find "enough" eigenvectors

with the initial conditions $y_1(0) = y_0$ and $y_2(0) = v_0$. This is a SOFOCCHE with coefficient matrix

$$A = \begin{bmatrix} 0 & 1 \\ -k/m & -\gamma/m \end{bmatrix}$$

and its characteristic equation is

$$\det(A - \lambda I) = \lambda^2 + (\gamma/m)\lambda + (k/m) = 0.$$

Multiply this by m to reveal an old friend: the characteristic equation from Chapter 2. All this shows that the two characteristic equations we have thus far defined truly are one in the same equation.

A Two-Mass System

Now we will see how to extend the arguments above to the two-mass system shown in the figure below: We will refer to the car on the right as Car 1 and the car on the left as Car 2

Figure 1: Example of a two-mass, three spring system.

with $x_i(t)$ denoting the displacement from equilibrium (with the positive direction pointing right) of Car i in meters after t seconds. We will also label the springs from right to left as Spring 1, Spring 2, and Spring 3. Using Newton's Second Law, we have

$$mx_1'' = F_1$$
$$mx_2'' = F_2$$

where F_i is the sum of the forces acting upon Car i. As before, the forces acting upon each car can be split up into two classes:

(I) Spring Forces: Hooke's Law states that the force exerted by a spring is proportional to its displacement: $F = -kx$ where k is a constant whose value depends on the elasticity of the spring (see also Section 2.1). Now if x_1, x_2 are both positive, then

- Spring 1 is stretched (from equilibrium) by x_1
- Spring 2 is stretched by $x_1 - x_2$
- Spring 3 is stretched by x_2

and the forces exerted by Springs 1 and 2 on Car 1 are both in the negative direction:

$$F_{S_{11}} = -k_1 x_1)$$
$$F_{S_{21}} = -k_2(x_1 - x_2)$$

while the forces exerted on Car 2 by Springs 2 and 3 are positive and negative respectively:

$$F_{S_{22}} = k_2(x_1 - x_2)$$
$$F_{S_{32}} = -k_3 x_2$$

(II) Frictional Forces: contribute forces on Car i proportional to x_i' say with constant of proportionality $-\gamma_i$.

Combining this information yields the system of two second order equations:

$$\begin{aligned} m_1 x_1'' &= -\gamma_1 x_1' - (k_1+k_2)x_1 + k_2 x_2 \\ m_2 x_2'' &= -\gamma_2 x_2' + k_2 x_1 - (k_2+k_3)x_2 \end{aligned} \tag{2}$$

To convert this system to a SOFOCCHE, we define the four variables:

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_1' \\ y_3 &= x_2 \\ y_4 &= x_2'. \end{aligned}$$

Then using a combination of substitution and rearrangement of the equation in (2), we obtain a system of four first order equations:

$$\begin{aligned} y_1' &= x_1' = y_2 \\ y_2' &= x_1'' = -(k_1+k_2)/m_1 x_1 - \gamma_1/m_1 x_1' + k_2/m_1 x_2 \\ &\qquad = -(k_2+k_2)/m_1 y_1 - \gamma_1/m_1 y_2 + k_2/m_1 y_3 \\ y_3' &= x_2' = y_4 \\ y_4' &= x_2'' = k_2/m_2 x_1 - (k_2+k_3)/m_2 x_2 - \gamma_2/m_2 x_2' \\ &\qquad = k_2/m_2 y_1 - (k_2+k_3)/m_2 y_3 - \gamma_2/m_2 y_4 \end{aligned}$$

We can write this new system in matrix-vector formulation as $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$ where the coefficient matrix A is given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -(k_1+k_2)/m_1 & -\gamma_1/m_1 & k_2/m_1 & 0 \\ 0 & 0 & 0 & 1 \\ k_2/m_2 & 0 & -(k_2+k_3)/m_2 & -(\gamma_2/m_2) \end{bmatrix}.$$

An example of the solution when $x_1(0) = 1, x_2(0) = -1,\ x_1'(0) = x_2'(0) = 0$ is shown in Figure 2. Notice that in this particular example, we have both $x_1(t) = \cos(t\sqrt{3})$ and $x_2(t) = -\cos(t\sqrt{3})$. The next section will show how to obtain this solution explicitly with the aid of Matlab.

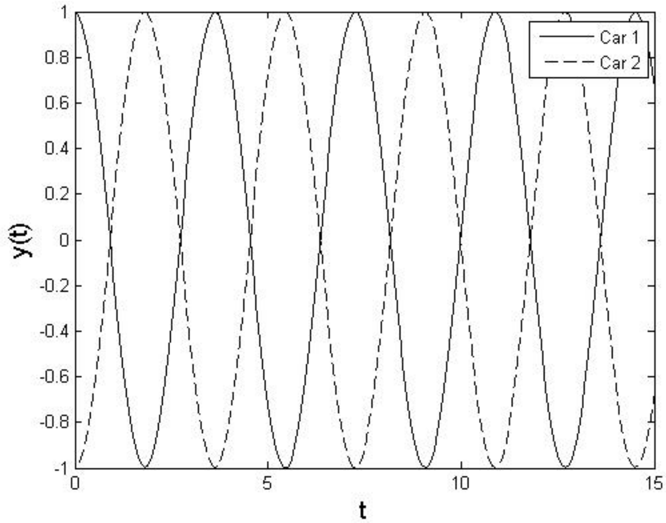


Figure 2: Sample displacements in a frictionless two-mass mechanical oscillator with opposite initial displacements, zero initial velocity, and masses/spring constants of 1.

Example of solving a two car spring mass system with Matlab

First, we define all parameters and enter them into the coefficient matrix:

```
g1=0;
g2=0;
k1=1;
k2=1;
k3=1;
m1=1;
m2=1;
A=[0 1 0 0;-(k1+k2)/m1 -g1/m1 k2/m1 0;0 0 0 1;k2/m2 0 -(k2+k3)/m2 -g2/m2]
```

```
A =

     0     1     0     0
    -2     0     1     0
     0     0     0     1
     1     0    -2     0
```

Next, we define a vector

$$\mathbf{y}_0 = \begin{bmatrix} x_1(0) \\ x_1'(0) \\ x_2(0) \\ x_2'(0) \end{bmatrix}$$

which stores the desired initial positions and velocities.

```
y0=[1;0;-1;0]

y0 =

     1
     0
    -1
     0
```

To obtain the general solution, we need the eigenvalues and eigenvectors of A

```
[V,L]=eig(A)


V =

   0.0000 + 0.3536i   0.0000 - 0.3536i   0.5000 + 0.0000i   0.5000 + 0.0000i
  -0.6124 + 0.0000i  -0.6124 + 0.0000i   0.0000 + 0.5000i   0.0000 - 0.5000i
   0.0000 - 0.3536i   0.0000 + 0.3536i   0.5000 + 0.0000i   0.5000 + 0.0000i
   0.6124 + 0.0000i   0.6124 + 0.0000i   0.0000 + 0.5000i   0.0000 - 0.5000i


L =

   0.0000 + 1.7321i   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 - 1.7321i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 1.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 - 1.0000i
```

Recall from the Matlab Supplement that the eigenvalues $\lambda_1, ...\lambda_4$ are stored on the diagonals of the matrix L and corresponding eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_4$ are stored in the columns of the matrix V. The general solution is then

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 + c_3 e^{\lambda_3 t}\mathbf{v}_3 + c_4 e^{\lambda_4 t}\mathbf{v}_4$$

To solve the given initial value problem, we first note that our initial conditions tell us

$$\mathbf{y}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + c_4\mathbf{v}_4.$$

This is actually a system of four equations for the four unknowns $c_1, ..., c_4$:

$$\begin{aligned} 1 &= c_1 v_{11} + c_2 v_{12} + c_3 v_{13} + c_4 v_{14} \\ 0 &= c_1 v_{21} + c_2 v_{22} + c_3 v_{23} + c_4 v_{24} \\ -1 &= c_1 v_{31} + c_2 v_{32} + c_3 v_{33} + c_4 v_{34} \\ 0 &= c_1 v_{41} + c_2 v_{42} + c_3 v_{43} + c_4 v_{44} \end{aligned}$$

where v_{ij} is the number in row i and column j on the matrix V above. This is not something you want to solve by hand, especially with all the imaginary numbers. But an alternative way of writing this system is as

$$\mathbf{y_0} = V\mathbf{c}$$

where

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

We have already seen how to solve such systems using "linsolve"

```
c=linsolve(V,y0)
```

```
c =

  -0.0000 + 1.4142i
  -0.0000 - 1.4142i
   0.0000 + 0.0000i

   0.0000 - 0.0000i
```

Luckily, two entries (c_3, c_4) are zero! Returning to our solution, we now know that the position of the first car is

$$x_1(t) = c_1 v_{11} e^{\lambda_1 t} + c_2 v_{12} e^{\lambda_2 t}$$

We can simplify the coefficients $c_1 v_{11}$ and $c_2 v_{12}$ using Matlab:

```
[c(1)*V(1,1),c(2)*V(1,2)]
```

```
ans =

   0.5000 - 0.0000i   0.5000 + 0.0000i
```

Wow, that's nice! After the smoke clears,

$$x_1(t) = \frac{1}{2} e^{\lambda_1 t} + \frac{1}{2} e^{\lambda_2 t}.$$

Furthermore, the eigenvalues $\lambda_1 = i\sqrt{3}$ and $\lambda_2 = -i\sqrt{3}$ are complex conjugates we can use Euler's trig identities to simplify this to

$$x_1(t) = \cos(t\sqrt{3}).$$

We invite you to verify that we also get

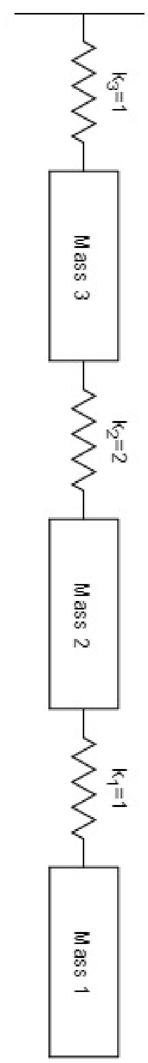
$$x_2(t) = -\cos(t\sqrt{3})$$

completing Figure 2.

Note that the systems actually has two pairs of complex conjugate eigenvalues: $\pm i$ and $\pm\sqrt{3}i$. With pure imaginary eigenvalues, we know that solutions $\mathbf{y}(t)$ will be all combinations of sines and cosines with periods 2π or $2 * \pi/\sqrt{3}$. This means that there are two *natural frequencies* of oscillation for our system and hence, we can expect two *resonant frequencies*. The next section will further investigate this hypothesis.

Exercises

(1) To scare off future soldiers, a demigorgon attaches three dead carcasses to the ceiling above the portal entryway separated by springs (Masses 1,2,3 in the figure on the right). It pulls each mass down by one meter and then pushes the lower mass with a upward velocity of 1 m/s before releasing. Assume that air resistance is negligible (no damping in the Upside Down) and all masses are 70 kg for the sake of simplicity.



(a) Write three second order equations for the displacements of the three masses after t seconds. Please define all variables and state your orientation!

(b) Convert your system in part (a) to a system of first order equations in matrix vector form. Alternate displacements and velocities in your variable definitions (y_1 is the displacement of first mass, y_2 is the velocity of first mass, etc...)

(2) Let x, y be solutions to the system of two second order equations

$$\begin{aligned} x'' &= -3x + y \\ y'' &= x - 3y \end{aligned}$$

along with the initial conditions $x(0) = 1, y(0) = 0,\ x'(0) = y'(0) = 0$.

(a) Convert this to a system of four first order equations in the unknowns $y_1 = x, y_2 = x', y_3 = y, y_4 = y'$.

(b) Determine the eigenvalues of the corresponding matrix of coefficients (using Matlab). What do the eigenvalues tell you about the behavior of this system?

(c) Determine a formula for and sketch a plot of the solutions $x(t), y(t)$ vs. t.

(3) Suppose that the spring-mass system in Figure 1 is operating in an ideal friction free environment (no damping), both cars have a mass of 1 kg, the springs all have identical spring constants $k_1 = k_2 = k_3 = 2$, and that Car 1 and 2 are released from equilibrium with initial velocities of 1 m/s and -1 m/s, respectively. Let $x(t), y(t)$ denote the displacement of the cars after t seconds.

(a) Write down a system of four first order equations, including initial values, which describes how the four variables x, x', y, y' evolve over time.

(b) The (approximate) eigenvectors and eigenvalues of the coefficient matrix for the system are given below along with the constants c_1, c_2, c_3, c_4 in the general solution which match the given initial conditions. Use this information to find a real valued formula for $y(t)$ and calculate its period and amplitude.

```
V =

   0.00 + 0.25i   0.00 - 0.25i   0.00 - 0.41i   0.00 + 0.41i
  -0.65 + 0.00i  -0.66 + 0.00i   0.58 + 0.00i   0.58 + 0.00i
   0.00 - 0.25i   0.00 + 0.25i   0.00 - 0.41i   0.00 + 0.41i
   0.65 + 0.00i   0.66 + 0.00i   0.58 + 0.00i   0.58 + 0.00i


L =

   0.00 + 2.45i   0.00 + 0.00i   0.00 + 0.00i   0.00 + 0.00i
   0.00 + 0.00i   0.00 - 2.45i   0.00 + 0.00i   0.00 + 0.00i
   0.00 + 0.00i   0.00 + 0.00i   0.00 + 1.41i   0.00 + 0.00i
   0.00 + 0.00i   0.00 + 0.00i   0.00 + 0.00i   0.00 - 1.41i


c =

  -0.75
  -0.75
   0.00
   0.00
```

(4) Compare the graphs of solutions to the undamped, two-mass, three spring system from the start of this section under the following three initial displacement scenarios assuming zero initial velocity.

(a) Both initial displacements are 1.

(b) One initial displacement is 1 and the other is -1.

(c) One initial displacement is 1 and the other is 0.

For each scenario, compute the coefficients $c_1, ..., c_4$ in the general solution and discuss how their values in conjunction with the eigenvalues of the coefficient matrix explain the graphs.

(5) Consider the spring-mass system shown in Figure 1 (the one from the class demo). Choose one question of interest which examines how the motion of the cars depends on one or more system parameters (eg. initial conditions, spring constants, masses, damping coefficients) and write a brief report which details your investigations, including some graphs and commentary.

Section 3.6: Eigenvalues and Resonant Frequencies

Section objectives:

- Determine resonant frequencies for spring-mass systems

If we add an external forcing $g(t)$ to Car 1 in Figure 1 from Section 3.4, we obtain a new system of two equations:

$$\begin{aligned} m_1 x_1'' &= -\gamma_1 x_1' - (k_1 + k_2)x_1 + k_2 x_2 + g(t) \\ m_2 x_2'' &= -\gamma_2 x_2' + k_2 x_1 - (k_2 + k_3)x_2 \end{aligned} \tag{1}$$

In class, we used a periodic forcing function of the form $g(t) = a\sin(\omega t)$ and observed that there were two resonant frequencies $\omega_T < \omega_O$ for the system, one of which resulted in tandem motion (ω_T) and one which resulted in opposing motion (ω_O). The goal of this section is to show why an undamped two-mass system will always yield two resonant frequencies by introducing the idea of a stiffness matrix. Although second order systems can always be handled by conversion to a system of SOFOCCHE as in the last section and we recommend you just stick with that approach if you need to find a specific solution (as long as you can use technology), that approach does have the problem of doubling the dimension and can get tedious in 3- or more car systems such as the one shown in Figure 1 below. Stiffness matrices don't offer significantly more in terms of computational efficiency, but they can give a quick and dirty method for computing resonant frequencies.

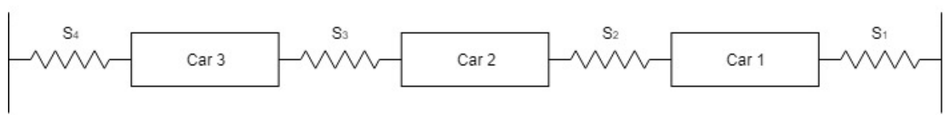


Figure 1: Example of a three-mass, four-spring system

Stiffness Matrices

We will use the example in Figure 1 to introduce the concept of a stiffness matrix. Let $x_i(t)$ denote the displacement of Car i after t seconds for $i = 1, 2, 3$ and let k_i be the spring constant for spring S_i, $i = 1, 2, 3, 4$. Newton's Second Law leads to a system of three equations for the x_i'' as follows:

$$\begin{aligned} m_1 x_1'' &= -k_1 x_1 - k_2(x_1 - x_2) = -(k_1 + k_2)x_1 + k_2 x_2 \\ m_2 x_2'' &= k_2(x_1 - x_2) - k_3(x_2 - x_3) = k_2 x_1 - (k_2 + k_3)x_2 + k_3 x_3 \\ m_3 x_3'' &= k_3(x_2 - x_3) - k_4 x_3 = k_3 x_2 - (k_3 + k_4)x_3. \end{aligned}$$

In matrix-vector form, we can rewrite this system as

$$M\mathbf{x}'' = K\mathbf{x} \tag{2}$$

where

$$M = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \; K = \begin{bmatrix} -(k_1+k_2) & k_2 & 0 \\ k_2 & -(k_2+k_3) & k_3 \\ 0 & k_3 & -(k_3+k_4) \end{bmatrix}, \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The matrix M is called the mass matrix, K is called the stiffness matrix, and $\mathbf{x}$ is called the displacement vector. Note that in the two-mass, three spring example from Section 3.5, we would get the stiffness matrix

$$K = \begin{bmatrix} -(k_1+k_2) & k_2 \\ k_2 & -(k_2+k_3) \end{bmatrix}$$

We could also rewrite the system in (2) as

$$\mathbf{x}'' = B\mathbf{x}$$

where $B = M^{-1}K$ is the matrix in which we divide row i of K by m_i. Now we know from last section that if we convert to a system of six first order equations $\mathbf{y}' = A\mathbf{y}$, then the general solution will be a combination of the vectors

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

where α is an eigenvalue of A, but can we find the α's directly from B instead? This would cut out the middle step of having to write out the 6×6 system of equations. To see how α is related to B, let's plug $\mathbf{x}(t) = \mathbf{v}e^{\alpha t}$ into the equation $\mathbf{x}'' = B\mathbf{x}$ to get

$$\alpha^2 \mathbf{v}e^{\alpha t} = B\mathbf{v}e^{\alpha t}$$

or after cancelling the exponentials:

$$B\mathbf{v} = \alpha^2 \mathbf{v}.$$

Look familiar? It is the same as the definition of eigenvalues and eigenvectors except with an α^2 in there. This means that α^2 is an eigenvalue of B and we have the following procedure:

- We divide each row of the stiffness matrix K by the mass of the corresponding car to form B
- We compute the eigenvalue/eigenvector pairs $(\lambda_i, \mathbf{v}_i)$ of B.
- We take $\pm\sqrt{\lambda_i}$ for each eigenvalue to form the general solution:

  $$\mathbf{x}(t) = \sum_{i=1}^{k} \mathbf{v}_i (a_i e^{\sqrt{\lambda_i} t} + b_i e^{-\sqrt{\lambda_i} t})$$

  where k is the number of masses.

Note that if there are k masses, there are actually $2k$ constants a_i, b_i in the above expression corresponding to the k initial positions along with the k initial velocities.

Example with the Two-mass system

Let's go back to the two mass system from the end of the Section 3.4 notes where we had $m_1 = m_2 = 1$ and $k_1 = k_2 = k_3 = 1$. Since the masses are all 1, we have

$$B = K = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

for this system. The eigenvalues of this matrix are $\lambda_1 = -1, \lambda_2 = -3$ with corresponding representative eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Taking square root of the eigenvalues yields the exponents in the solution vector: $\alpha = \pm i, \pm\sqrt{3}i$. Note that these were the eigenvalues we saw in the Matlab output from Section 3.4. We can now form the general solution

$$\mathbf{x}(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left(a_1 e^{it} + b_1 e^{-it}\right) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \left(a_2 e^{i\sqrt{3}t} + b_2 e^{-i\sqrt{3}t}\right) \tag{3}$$

and then plug in appropriate initial conditions for the initial displacements and velocities. For example, in the case where we release the cars in opposite directions, we would have the initial displacement vector

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Plugging this into Equation (3) gives us two equations for the four unknowns a_1, a_2, b_1, b_4:

$$\begin{aligned} 1 &= a_1 + b_1 + a_2 + b_2 \\ -1 &= a_1 + b_1 - a_2 - b_2 \end{aligned}$$

We get another two equations if we use the initial velocity vector

$$\mathbf{x}'(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Differentiating the general solution and plugging this in yields

$$\begin{aligned} 0 &= ia_1 - ib_1 + i\sqrt{3}a_2 - i\sqrt{3}b_2 \\ 0 &= ia_1 - ib_1 - i\sqrt{3}a_2 + i\sqrt{3}b_2 \end{aligned}$$

or after dividing both sides by i:

$$\begin{aligned} 0 &= a_1 - b_1 + \sqrt{3}a_2 - \sqrt{3}b_2 \\ 0 &= a_1 - b_1 - \sqrt{3}a_2 + \sqrt{3}b_2. \end{aligned}$$

To solve these four equations, you can either use linsolve in Matlab or in this example, you can add the first two equations together and the second two equations together to eliminate a_2, b_2 and get two equations for a_1, b_1:

$$\begin{aligned} 0 &= 2a_1 + 2b_1 \\ 0 &= 2a_1 - 2b_1. \end{aligned}$$

The only solution here is $a_1 = b_1 = 0$. Now we go back and subtract equations instead to get

$$\begin{aligned} 2 &= 2a_2 + 2b_2 \\ 0 &= 2\sqrt{3}a_2 - 2\sqrt{3}b_2. \end{aligned}$$

which has solution $a_2 = b_2 = 1/2$. Nice - this means that in our general solution, we get

$$\mathbf{x}(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \left(\frac{1}{2}(e^{i\sqrt{3}t} + e^{-i\sqrt{3}t}) \right) = \begin{bmatrix} \cos\sqrt{3}t \\ -\cos\sqrt{3}t \end{bmatrix},$$

confirming our early observation that opposing initial conditions lead to opposing motion! The exercises below will ask you to confirm that if we change the initial displacement of Car 2 to +1, then we get solutions $x_1(t) = x_2(t) = \cos(t)$ which models tandem motion.

Resonant Frequencies

In the previous example, we saw two pure imaginary eigenvalues with two different angular frequencies. These eigenvalues correspond to the two natural frequencies of our system and building off our work in Section 2.4, we can also expect that these are the two resonant frequencies for the system. This connection is further strengthened by the example from the last section: starting the cars off in opposite directions led to opposing oscillations with the higher natural angular frequency $\sqrt{3}$ while starting them off in the same direction led to tandem oscillations with the lower frequency 1.

To confirm all this works, we would like to go through an example of finding the general solution to an undamped two-mass system with periodic forcing term $\mathbf{g}(t) = \mathbf{a}\exp(i\omega t)$. For simplicity, we will assume that all the masses are 1 although the general case can be handled by first dividing all spring and forcing constants by the masses. Our second order system then takes on the form

$$\mathbf{x}'' = K\mathbf{x} + \mathbf{g}. \tag{4}$$

We choose as our trial solution

$$\mathbf{x}_p(t) = \mathbf{A}e^{i\omega t}$$

where $\mathbf{y}$ is an unknown vector of coefficients. Plugging this into Equation (4) for $\mathbf{x}$ gives

$$-\omega^2\mathbf{A}e^{i\omega t} = K\mathbf{y}e^{i\omega t} + \mathbf{a}e^{i\omega t}.$$

We can eliminate the exponentials terms and rearrange to obtain

$$(K + \omega^2 I)\mathbf{A} = -\mathbf{a}.$$

This is a system of equations we can solve for $\mathbf{A}$ IF the matrix $K + \omega^2 I$ is invertible. But wait a minute, in Section 3.2 we saw that this matrix is NOT invertible exactly when $\det(K+\omega^2 I) = 0$ which is the equation we would use to find the eigenvalues of K (just replace $+\omega^2$ with $-\lambda$). Therefore, the values of forcing frequencies ω which lead to singularities here are exactly those values for which $-\omega^2$ is an eigenvalue of the stiffness matrix K. This is also how we computed the natural frequencies of the system!

We can use this observation to quickly identify resonant frequencies of a system and determine how they depend on system parameters. For example, suppose that the masses are all 1, the spring constants for the outer springs (1 and 3) are the same value k, but we double the spring constant of the inner spring (2) to $2k$. Then the stiffness matrix is

$$K = \begin{bmatrix} -3k & 2k \\ 2k & -3k \end{bmatrix}$$

which has characteristic equation

$$\lambda^2 + 6k\lambda + 5k^2 = (\lambda + 5k)(\lambda + k) = 0.$$

The eigenvalues are $-k, -5k$ which means that resonant frequencies are $\sqrt{k}, \sqrt{5k}$. Would you have guessed that the opposing resonant frequency would be $\sqrt{5}$ times as large as the tandem one in this example?

Final Remarks and Analogies for Linear Systems

A secondary goal of this chapter (next to actually solving systems of equations) has been to demonstrate how linear algebra helps us solve multidimensional problems by making analogies with one dimensional problems. We would like to close the chapter with several other examples. First, it is actually possible to a homogeneous two-mass system

$$\begin{aligned} m_1 x_1'' &= -\gamma_1 x_1' - (k_1 + k_2)x_1 + k_2 x_2 \\ m_2 x_2'' &= -\gamma_2 x_2' + k_2 x_1 - (k_2 + k_3)x_2 \end{aligned} \tag{5}$$

in an analogous way to our standard SOCCHE

$$my'' + \gamma y' + ky = 0$$

with the solution y replaced by the solution vector $\mathbf{y} = (x_1, x_2)$ and all the constants m, γ, k replaced with matrices M, Γ, K. To do so, we first move all terms to the left side of the equation in (5)and rewrite the system in vector form as

$$\begin{bmatrix} m_1 x_1'' \\ m_2 x_2'' \end{bmatrix} + \begin{bmatrix} \gamma_1 x_1' \\ \gamma_2 x_2' \end{bmatrix} + \begin{bmatrix} (k_1 + k_2)x_1 - k_2 x_2 \\ -k_2 x_1 + (k_2 + k_3)x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{6}$$

Then we define the three matrices

$$M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \Gamma = \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix}, K = \begin{bmatrix} (k_1 + k_2) & -k_2 \\ -k_2 & (k_2 + k_3) \end{bmatrix}$$

and voila, equation (6) simplifies to

$$M\mathbf{y}'' + \Gamma\mathbf{y}' + K\mathbf{y} = \mathbf{0}. \tag{7}$$

Second, it is also possible to obtain a matrix analog to our old characteristic equation

$$m\lambda^2 + \gamma\lambda + k = 0$$

Recalling that we got this by plugging in $e^{\lambda t}$, let's plug $\mathbf{y}(t) = \mathbf{v}e^{\lambda t}$ into Equation (7) and see which $\lambda, \mathbf{v}$ pairs yield a solution. Doing so and cancelling the $e^{\lambda t}$ terms gives the relationship

$$(\lambda^2 M + \lambda\Gamma + K)\mathbf{v} = \mathbf{0}.$$

Any $\lambda, \mathbf{v}$ which satisfies this equation leads to a solution. Recalling from Section 3.2 that any matrix-vector equation of the form $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions if and only if $\det(A) = 0$, we get a new characteristic equation

$$\det(\lambda^2 M + \lambda\Gamma + K) = 0 \tag{8}$$

for determining the exponents λ corresponding to solutions.

Before moving onto our final example, let's stop to appreciate the coolness of this. We have now shown that all FOCCHE/SOFOCCHE and SOCCHE/SOSOCCHE have exponential solutions, potentially with complex exponents, and all four cases have characteristic equations which determine the exponents of solutions. The relationships are summarized in the table below. And once you know the exponents of the solutions, you have the keys to unlocking the asymptotic behavior.

Object	Single Equation	System of equations
First Order DE	$y' = ay$	$\mathbf{y}' = A\mathbf{y}$
First Order Char equation	$a - \lambda = 0$	$\det(A - \lambda I) = 0$
Second Order DE	$my'' + \gamma y' + ky = 0$	$M\mathbf{y}'' + \Gamma\mathbf{y}' + K\mathbf{y} = \mathbf{0}$
Second Order Char equation	$m\lambda^2 + \gamma\lambda + k = 0$	$\det(M\lambda^2 + \Gamma\lambda + K) = 0.$

Finally, at the beginning of this chapter we hinted that the equation $\mathbf{y}' = A\mathbf{y}$ should have a solution of the form $e^{At}\mathbf{y}_0$ by analogy with $y' = ay$ having solution $e^{at}y_0$. This is in fact true, but requires us to first define the matrix exponential function. The actual definition is typically done in terms of Taylor series (similar to what we did to define e^{it}), but we can also take another route in terms of eigenvalue-eigenvector problems. Recall that the general solution to a SOFOCCHE is

$$\mathbf{y}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t}$$

where the $\lambda_i, \mathbf{v}_i$ form eigenvalue-eigenvector pairs for A. Using the definitions of matrix multiplication, we can rewrite this expression as

$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_1 t}\end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Recalling that

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = V^{-1}\mathbf{y}_0,$$

where $V = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix}$ changes this to

$$\mathbf{y}(t) = V \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_1 t}\end{bmatrix} V^{-1}\mathbf{y}_0.$$

Now it seems natural to represent the matrix sandwiched in between V and V^{-1} as

$$e^{Dt}$$

where D is the diagonal matrix of eigenvalues so adding this new definition in gives

$$\mathbf{y}(t) = Ve^{Dt}V^{-1}\mathbf{y}_0.$$

Finally, defining

$$e^{At} = Ve^{Dt}V^{-1}$$

the above simplifies to

$$\mathbf{y}(t) = e^{At}\mathbf{y}_0$$

as desired.

Practically speaking, this representation isn't so useful, but theoretically, it is beautiful and helps cast further light on how eigenvalues relate matrices to scalars. For example, we know that $e^{at} \to 0$ as $t \to \infty$ if $a < 0$, but blows up if $a > 0$. We can say the same thing for e^{At}: it coverges to the zero matrix if its eigenvalues are negative and blows up if they are positive. Of course, this is only the tip of the iceberg. But you will have to take Math 145 (Applied Linear Algebra) and Math 157 (Partial Differential Equations) to learn more.

Exercises

(1) Determine the solutions $x_1(t), x_2(t)$ for the displacements of Cars 1 and 2 in the undamped two-mass, three-spring system with $m_1 = m_2 = k_1 = k_2 = k_3 = 1$ when

  (a) the two cars are displaced by 1 m in the same direction and released.

  (b) Only Car 1 is displaced by 1 m and released.

  Plot both solutions in Matlab and compare with our in class experiment.

(2) Calculate the resonant frequencies for undamped two-mass, three spring systems with the following parameter values and say something about each (like "wow, the opposing frequency is $\sqrt{5}$ times as large as the tandem one when we double the tension of the inner spring relative to the outer spring!)

  (a) $m_1 = m_2 = m$; $k_1 = k_2 = k_3 = k$.

  (b) $m_1 = m_2 = 1$, $k_2 = k$, and $k_1 = k_3 = 2k$.

  (c) $m_1 = m_2 = 1$, $k_1 = k_2 = k$, and $k_3 = 2k$.

  (d) $m_1 = m, m_2 = 2m$ and all the spring constants $= 1$.

(3) (This exercise requires the use of 3×3 determinants which are not covered in this text, but which you probably learned in Calc 3). Determine the resonant frequencies of the three-car, four spring system in Figure 1 in the case where all the masses and spring constants are 1.

Part IV
Concluding Remarks

Epilogue: A nonlinear world

Most of this course has focused on modeling with *linear* differential equations. Linear differential equations (or systems of linear equations) are equations which depend on the unknown function(s) and its (their) derivatives in a linear way. A generic first order linear equation can be written in the form

$$y' + p(t)y = g(t)$$

for some functions p, g while a linear second order equation has a similar form and a system of linear first order equations is a collection of coupled first order linear equations. Hands down the most important idea for solving linear systems is:

> **The Principle of Superposition**: If y_1 is a solution to a linear system with forcing g_1 and y_2 is a solution to the same linear system with forcing g_2 then $y_1 + y_2$ is a solution to the same system with forcing $g_1 + g_2$.

This principle comes along with two useful side effects:

- If y_1, y_2 are both solutions to a homogeneous system, then so is any linear combination of y_1, y_2.
- If $y_p(t)$ is a solution to a nonhomogeneous system and $y_c(t)$ is a solution to the complementary homogeneous system, then $y(t) = y_c(t) + y_p(t)$ is also a solution to the nonhomogeneous problem.

The first allows us to form *general solutions* to homogeneous problems from a small number of particular solutions and the second allows us to extend these to nonhomogeneous problems.

Unfortunately, nonlinear equations do not follow the principle of superposition and hence, there are not as many generic, clean methods for finding explicit solutions. But the last 50-60 years have led to some important breakthroughs in this field, particularly in the development of dynamical systems and chaos theory[1]. The purpose of this final section is to give you a brief introduction to the study of nonlinear systems including:

- Setting up nonlinear models to describe models for competition.
- Stability of equilibrium solutions and qualitative analysis of solutions.

We'll focus on two examples: the logistic equation for population growth and modeling competition between species.

[1] For a popular introduction to his subject, James' Gleick's "Chaos" is a fun read; for more mathematical details, Devaney's "An Introduction to Chaotic Dynamical Systems" is a good start while Hirsch and Smale's "Differential Equations, Dynamical Systems, and Linear Algebra" is a now classic text on the subject which also recovers what we have done in here.

Logistic growth

Back in Section 1, we discussed the Malthusian Model for population growth in which it is assumed that the population grows at a rate proportional to its current size. This assumption is somewhat unrealistic since it implies that the per capita growth rate is constant: individuals give birth and die at the same rates regardless of the population size. In the 1830's, Verhulst[2] proposed a modified population growth model in which he assumed that the per capita death rates of individuals in the population are actually ***density dependent***; that is, they increase as the size of the population increases. To see what happens to a population under this assumption, let $y(t)$ denote the number of individuals alive after t years. Suppose that individuals, on average, give birth once every a years and have an average lifespan of $1/(d_1 + d_2 y(t))$ (that is, the average lifespan decreases as the number of individuals increases). Then the per capita birth rate of individuals is $b = 1/a$ per year and the per capita death rate is $d_1 + d_2 y(t)$. Therefore, a differential equation modeling the evolution of $y(t)$ over time is

$$\begin{aligned} y' &= \text{“Rate in” - “Rate out”} \\ &= by - (d_1 + d_2 y)y \\ &= (b - d_1)y - d_2 y^2 \\ &= ry(1 - y/K) \end{aligned} \tag{1}$$

where $r = b - d_1$ is called the ***intrinsic growth rate*** of the population and $K = r/d_2$ is called the ***carrying capacity*** (we shall investigate why later!) The equation above is called the logistic growth equation and it is a separable, first order equation and has been frequently used to model the growth of populations over time[3].

Instead of solving this equation directly (which can be done using separation of variables and partial fractions), let's instead see if we can describe the long term behavior of solutions without too much algebra. Recall from basic calculus that the sign of the derivative of a function (positive, negative or zero) tells us whether the function itself is increasing, decreasing, or constant. The table below shows the sign of y' for the logistic equation in different regions of y values along with the corresponding behavior of the solution y:

y	Sign of y'	Behavior of y
$(-\infty, 0)$	$-$	Decreasing
0	0	Equilibrium
$(0, K)$	$+$	Increasing
K	0	Equilibrium
(K, ∞)	$-$	Decreasing

This table allows us to guess the asymptotic behavior of solutions based on the initial starting values of y. For example, if $0 < y_0 < K$, then $y(t)$ will be increasing as long as $y(t) < K$, but since $y'(K) = 0$, it can never pass through K so $y(t) \nearrow K$ as $t \to \infty$. Similarly, if

[2] Verhulst, P.F. (1838) Notice sur la loi que la population suit dans son accroissement. *Corr. Math. et. Phys.* 10, 113-120.

[3] For example, the growth of protozoa populations in vitro (Gause, G.F (1964) *The struggle for existence.* Dover) and national populations (Pearl and Reed (1920) *Proceedings of the National Academy*, pg. 275.)

$y_0 > K$, then $y(t)$ will decrease towards K as $t \to \infty$. (It is for this reason that K is called the carrying capacity of the system).

Another way of thinking this is via a **phase line diagram** like the one on the right in Figure 1. Here we have basically taken the two-dimensional plot of some solution curves from Figure 1a and smashed it down to one dimension representing y with t implicitly shown through the direction of the arrows. Notice that solutions which start near the equilibrium solution of 0 tend to move away from 0 while solutions which start near K tend to move towards K. For these reasons, we call 0 a **source** and K a **sink** (these names make even more sense if you think about three-dimensional analogs).

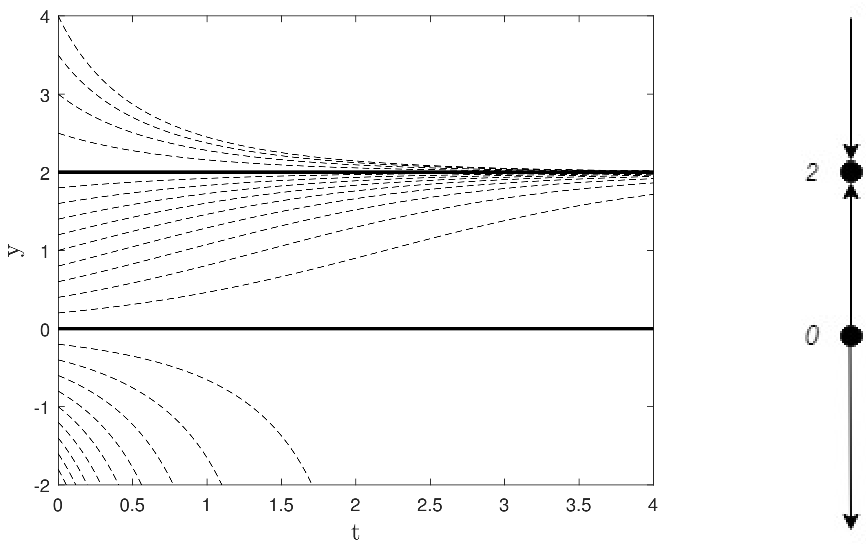


Figure 1: (Left) Some solutions curves of the Logistic Equation with $K = 2$. (Right) Phase Line Reduction

Having introduced these basic ideas in one-dimension, let's build up to a more complicated example.

Modeling Competition between species

In the late 1910's following World War I, an Italian biologist by the name of Umberto D' Ancona was studying selachian (i.e. shark) levels in the Adriatic sea and noticed an unexplainable sharp rise in population levels during this time. He took this data to his cousin, a mathematician by the name of Vito Volterra, who developed a mathematical model to explain this phenomenon (a similar model was independently developed by another mathematician A.J. Lotka[4] to describe herbivore-plant relationships). For a more exciting example, everyone knows that vampires live off humans and humans benefit nothing from the existence of vampires (except possibly some mildly entertaining movies). Let's set up a simple system of differential equations to describe the competition between humans and vampires

[4]Lotka, A.J. (1920) Analytical Note on Certain Rhythmic Relations in Organic Systems, Proc. Natl. Acad. Sci. U.S., 6, 410-415.

in the United States and estimate the ultimate outcome of this competition. We let $h(t)$ and $v(t)$ denote the sizes of the human and vampire populations, respectively, at time t. We will measure t in units of average human lifespans (so $t = 1$ corresponds to about 75 years). We will measure h and v in millions. Such models are called ***predator-prey*** models (in this example, human being the prey and vampires the predator).

To set up a system of equations governing the interaction between humans and vampires, we begin by stating our assumptions. The following list indicates one possible set of such assumptions:

We assume that in the absence of vampires, humans have a constant per capita birth rate of b_h per lifespan. Note that we are thinking of b_h as the "average" birth rate per human so according to the CIA, $b_h \approx 1.03$ (the total fertility rate is 2.06). We assume that the per capita death rate of humans stays constant at 1 (since time is measured in human lifespans!)

We assume that in the absence of humans, each vampire dies due to starvation after, on average, $1/d_v$ years (so the per capita death rate is d_v).

The per capita rate at which humans encounter vampires is proportional to the number of vampires, i.e. the per capita rate is $mv(t)$ where $m > 0$. A human/vampire encounter will result in a human to vampire conversion with probability p_v, the human killing the vampire (go Buffy!) with probability $p_h < p_v$, or nothing with probability $1 - p_v - p_h$.

The differential equations for h and v can then be set up using the balancing principle:

$$h' = b_h h - h - mp_v vh = ah - bhv$$
$$v' = mp_v hv - d_v v - mp_h hv = chv - dv$$

where $a = b_h - 1 \approx 0.03$, $b = mp_v$, $c = m(p_v - p_h)$, and $d = d_v$. This is again a system of two equations in two unknowns, but it is now a nonlinear system since both equations involve product $h \cdot v$. Also note that both equations are generalizations of the logistic equation from the previous section.

If we set $h' = 0, v' = 0$, we find that there are two equilibrium solutions to the system:

$$\begin{bmatrix} h \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix} = \begin{bmatrix} d/c \\ a/b \end{bmatrix}$$

The value of the first equilibrium does not depend on any parameters and just represents extinction of both species. The second equilibrium represents human vampire coexistence. Note that its value decreases as b (the human death rate due to predation) and c (the human-vampire conversion rate) increase and that its value increases as d (the vampire death rate) and a (the human birth rate) increase. In this sense, overly aggressive predators are bad for everyone!

Calculating equilibrium solutions is easy, but now we have the additional problem of determining towards which equilibrium, if any, the populations of humans and vampires will converge. One things that we can say is that the vampires can grow ($v' > 0$) if and only if

$$h > d/c$$

so if the humans reach too small a size, the vampires will also start to die off from lack of food. On the other hand, if $< a/b$, then $h' > 0$ so that the humans can grow as long as the size of the vampires is small. This suggests the possibility of "cyclic orbits": humans increase in size, providing more food for vampires to grow, who in turn kill off more humans, which means less food, which means a decrease in vampire populations, which means that humans can start thriving again and so on. Very different from what we observed with the logistic equation!

Stability of Equilibrium Solutions

To gain some intuition about possible limiting behaviors for nonlinear systems of equations, let's recall a few facts about the stability of equilibrium solution for linear equations. In Chapter 3, we saw that the general solution to a SOFOCCNE $\mathbf{y}' = A\mathbf{y} + \mathbf{b}$ has the form

$$\mathbf{y}(t) = \mathbf{y}_p + c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} \tag{2}$$

where $(\lambda_1, \mathbf{v}_1), (\lambda_2, \mathbf{v}_2)$ are eigenvalue-eigenvector pairs of A and $\mathbf{y}_p$ is an equilibrium solution[5]. Equation (2) tells us that the stability of an equilibrium depends on the eigenvalues λ_1, λ_2 of A. Two negative real eigenvalues imply that all solutions converge to equilibrium, for example, while two positive eigenvalues imply that all solutions would diverge away from the equilibrium. For complex eigenvalues, we saw that solutions either circle an equilibrium (corresponding to pure imaginary eigenvalues) or they spiral into/out from the equilibrium depending on whether the real part of the eigenvalues is positive/negative.

With this review in mind, let's get back humans and vampires. Letting $\mathbf{x} = \begin{bmatrix} h \\ v \end{bmatrix}$, We can rewrite the equations governing this system in vector form as

$$\mathbf{x}' = \begin{bmatrix} ah - bhv \\ chv - dv \end{bmatrix}.$$

For nonlinear systems such as this one, there is no coefficient matrix so to speak. Instead, we need to linearize the system near an equilibrium to analyze stability. You may recall from Calc 3 that the Jacobian matrix of a function $\mathbf{f}(x, y) = (f_1(x, y), f_2(x, y))$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ is the matrix of partial derivatives

$$J(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y}. \end{bmatrix}$$

To apply this in our human-vampire example, define

$$\mathbf{f}(h, v) = (ah - bhv, chv - dv)$$

as the right hand side of the equation. Then the Jacobian Matrix is

$$J(h, v) = \begin{bmatrix} a - bv & -bh \\ cv & ch - d. \end{bmatrix}$$

[5]All this provided that we can find two LI eigenvectors $\mathbf{v}_1, \mathbf{v}_2$.

If we plug in our coexistent equilibrium $(d/c, a/b)$, then we get the matrix

$$J = \begin{bmatrix} 0 & -bd/c \\ ca/b & 0 \end{bmatrix}$$

which has characteristic equation

$$\lambda^2 + ad = 0$$

and complex eigenvalues $\pm i\sqrt{ad}$. This implies that the coexistent equilibrium is a center just like in the previous example! This is further illustrated in Figure 2 and you can play around with further examples in the Matlab hv gui.

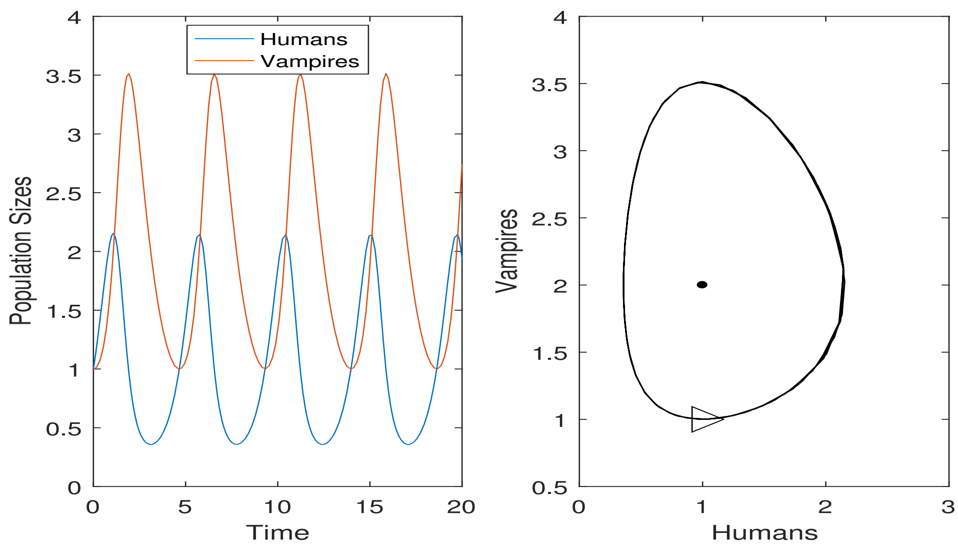


Figure 2: (Left) Time-series plot of human-vampire populations with $a = 2, b = c = d = 1$ and initial conditions $x(0) = 1, y(0) = 1$. (Right) Phase-portrait plot of the same solution

Concluding Remarks

Nonlinear modeling gets even more interesting in three-dimensions where solutions can converge to strange sets such as the Lorentz Attractor shown in Figure 3. Nonlinear systems have also given rise to the field of chaos theory which has played a pivotal roles in popular movies such as The Butterfly Effect, Sliding Doors, and Jurassic Park. While a linear world is orderly and predictable, a nonlinear world is chaotic and surprising: a small difference in initial conditions can lead to divergent long term behavior. We have only glimpsed the surface of this exciting world, but hope that you will continue on to higher level math courses to learn more!

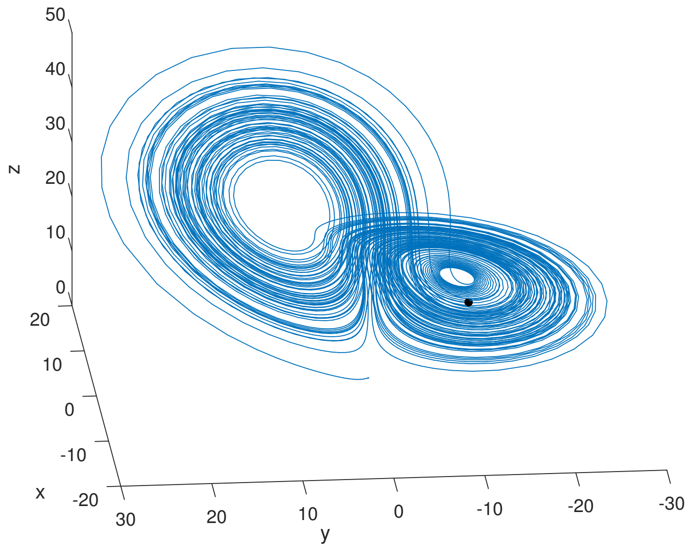


Figure 3: Sample solution to the Lorenz system $x' = s(y-x), y' = x(r-z)-y, z' = xy-bz$ with parameters $r = 28, s = 10, b = 8/3$

Exercises

(1) Use phase-line diagrams to determine the asymptotic behavior, as $t \to \infty$, of solutions to each of the following DE's for all possible values of $y(0)$.

(a) $y' = ay$, $a > 0$.

(b) $y' = y(y-1)(y-3)$

(c) $y' = (y+1)^3$

(d) $y' = -(y+1)^3$

(e) $y' = (y+1)^2$

(2) Suppose that the proportion $p(t)$ of all Pacific students infected by an influenza outbreak is evolving according to the logistic equation

$$\frac{dy}{dt} = \frac{y}{5}(1-2y) \tag{3}$$

and that initially, one student is infected[6]. What proportion of students will eventually become infected?

(3) Motivated by an article written by the biologist Stephen Jay Gould [7] which points out that most species of snails consist of animals whose shells curl almost exclusively in one direction, Clifford Taubes[8] suggests the following differential equation for the proportion, $p(t)$, of all snails in a particularly species which curl to the left:

$$\frac{dp}{dt} = ap(1-p)(p-1/2)$$

where a is a species dependent parameter. Determine the asymptotic proportion of left handed snails, as $t \to \infty$, if (a) $1/4$ of all snails are initially left-curling and (b) $3/4$ of all snails are initially left-curling.

(4) A rogue group of Guerilla fighters of size $x(t)$ is at war with a large, conventional marching army of size $y(t)$. The conventional army incurs losses at a rate proportional to the size of the Guerrilla army, but the Guerilla army incurs losses at a rate proportional to the product of the sizes of the two armies.

(a) Write down a system of differential equations for x, y.

(b) Determine all equilibrium solutions to the system in (a).

(c) What are the eigenvalues of the Jacobian matrix at an equilibrium solution?

(d) Use Matlab (Euler's, ode solver, other) to plot some sample solutions in time series and phase-plane format. Discuss the stability of equilibrium solutions and try to find some scenarios which will lead to Guerrilla victories!

---

[6] The use of the logistic equation to model the spread of contagious diseases was first proposed by Daniel Bernoulli in 1760. More recently, the logistic equation was used to study the spread of H1N1 in small towns (`http://www.intmath.com/blog/h1n1-and-the-logistic-equation/3498`)

[7] Gould, S.J. (1995) Left snails and right minds. *Natural History*, 10-18.

[8] Taubes, C.H. (2001) *Modeling differential equations in biology.* Prentice Hall, New Jersey.

(5) Returning to Vito Volterra's original predator (selachian) and prey (fish) model, suppose that prey gives birth at constant rate b_f to new prey, die to natural causes at rate d_f, and are killed by sharks at a rate proportional to the current number of sharks. Assume that the birth rate of sharks is limited by the number of prey present so that the per capita birth rate of sharks is proportional to the current number of fish. In the absence of prey, sharks die at rate d_s due to starvation. Suppose, in addition, that both fish and sharks are fished at a rate proportional to their current sizes.

(a) Write down a system of equations to describe the evolution of the predator-prey populations.

(b) Find the coexisting equilibrium and show that a decrease in fishing increases the equilibrium size of the shark population.

(c) Calculate the Jacobian matrix at the coexistent equilibrium and find its eigenvalues. Is the equilibrium a sink, source, or neither?

(d) Use Matlab (Euler's, ode solver, other) to plot some sample solutions in time series and phase-plane format. Do the results confirm your work in (c)? How do the results depend on system parameters (especially the fishing rate)?

(6) In epidemiology, the Susceptible-Infected-Recovered (SIR) model is often used to describe the spread of infectious diseases, such as seasonal influenza, which infer immunity upon recovered individuals. In this model, we let $S(t), I(t)$, and $R(t)$ denote the proportion of susceptible, infected, and recovered individuals in a community of fixed size N, t days after the start of an outbreak. We assume that susceptible individuals encounter infected individuals at a rate proportional to the current number of infecteds and such encounters result in the spread of the infection with probability p. We also assume that infecteds recover after an average of d days.

(a) Ignoring births and deaths due to other causes, set up a system of first order equations to describe how the proportions of infected and susceptible individuals evolve over time.

(b) What is the critical proportion of susceptibles required for the disease to spread?

(c) What are the equilibrium solutions and corresponding eigenvalues of the Jacobian matrix in this example?

(d) Use Matlab (Euler's, ode solver, other) to plot some sample solutions in time series and phase-plane format. Experiment with different values for β, r and discuss the results.

Principles of Modeling with Differential Equations

Review and Additional Practice

In this class, we have encountered many examples of real (and not so real) world models which involve differential equations. Here are some of the principles we have applied in setting up equations in such scenarios:

Balancing Principle: The rate of change of a quantity is the "rate in" minus the "rate out".

Newton's Second Law: The sum of the forces acting upon an object is equal to the product of the object's mass and acceleration.

Newton's Law of Cooling: The rate of change of the temperature of an object is proportional to the temperature difference between the object and its surroundings.

Hooke's Law: The force exerted by a spring is proportional to the spring's current displacement from rest.

Kirchoff's Voltage Law : The sum of all voltages around a closed loop is 0.

We have also practiced translating verbal statements such as *the rate at which x decreases is proportional to y* into differential equations such as

$$\frac{dx}{dt} = -ky$$

where $k > 0$ and talked about per capita birth/death rates in population models.

In the modeling scenarios described below, set up a differential equation or system of equations for the unknown quantity or quantities of interest including the initial conditions when specified. Also list all methods which we could use to solve (exactly or approximately) the given system of equations (you do not actually need to solve them though!) and anything else you know about it (linear/nonlinear, homogeneous/nonhomogeneous, constant coefficient, first/second order, matrix of coefficients if applicable).

1. A researcher believes that a certain type of tumor cell has an average lifespan of d days and while alive, divides an average of b times per day to create two identical cells. Initially there are 100 cells in the tumor. Quantity: $N(t)$ = number of tumor cells after t days.

2. A 110° light bulb is placed in a room whose temperature is periodically varying over a 24 hour period between a high of 25° at time 0 and a low of 15° halfway through the cycle. Quantity: $y(t)$ = temperature of the bulb after t hours.

3. Mosquitoes in Stockton have an average per capita birth rate of a per day, but a constant number k die every day due to local pest control measures. Quantity: $m(t)$ = Number of mosquitos (in thousands) after t days.

4. Micah's leaky wading pool initially contains 10 gallons of water with a concentration of 1 ounce of dirt per gallon. Water leaks from the pool at a rate of r_2 gallons per minute. Pure water is pumped into the pool at a rate of r_1 gallons per minutes. Quantity: $x(t)$ = amount of dirt in the tub after t minutes.

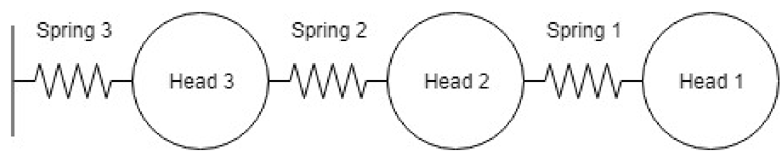
5. A closed circuit consists of a resistor with a resistance of $R = 10^4$ Ohms, a capacitor with a capacitance of $C = 10^{-3}$ Farads, and a driving force $g(t) = 10\sin(\pi t)$ volts. The initial voltage drop across the capacitor is 0. Quantity: $v(t)$ = the voltage across the capacitor after t seconds.

6. A body of mass m kg is dropped from a tall building of height H. Suppose that the only forces acting upon the body are gravity which supplies a constant force of $9.81\, m/s^2$ and air resistance which slows the body down at a rate proportional to the square of its velocity. Quantity: $y(t)$ = Distance fallen after t seconds.

7. The 1.8 meter tall Green Arrow shoots an arrow of mass 250 g straight up into the air with an initial velocity of 5 m/s. Same assumptions as previous question. Quantities: $y_1(t)$ = the arrow's height above the ground and $y_2(t)$ = the arrow's upward velocity after t seconds.

8. One end of a spring is attached to a zombie head of mass m kg while the other end is attached to the ceiling. The head is pulled y_0 meters down from its resting position and released. Assume that air resistance is proportional to velocity and that the ceiling is high enough so that the zombie head never hits the ground. Quantity: $y(t)$ = displacement of the top of the zombie head from equilibrium in the downward direction in meters after t seconds.

9. Amos is towing the heads of three rebellious crew members which are attached with springs to the back of the spaceship Rocinante as in the figure below. Assume that the Rocinante is travelling at a constant velocity and that there are no damping forces acting upon the heads. Quantities: The displacements of the three heads after t seconds.



10. Suppose we have a system consisting of two interconnected tanks both of which hold 20 gal of pure water. Water containing a concentration of 1 oz/gal of salt is pumped into Tank 1 at a rate of 1 gal/min. The well stirred mixture flows from Tank 1 to Tank 2 at a rate of 2 gal/min. The mixture flows from Tank 2 back to Tank 1 at a rate of 1 gal/min and is pumped out of Tank 2 at a rate of 1 gal/min. Quantities: $x_1(t), x_2(t)$ = the amount of salt (oz) in Tanks 1 and 2, respectively, after t minutes.

11. The amount of quantity x increases at a rate proportional to the amount of y, but decreases at a rate proportional to the amount of z, the amount of y increases at a rate proportional to z, but decreases at a rate proportional to x, and the amount of z increases at a rate proportional to x, but decreases at a rate proportional to y. Quantities: Amount of each quantity after t units of time.

12. Zombies are killed at a rate proportional to the number of humans because they are dumb, but humans can hide and are killed at a rate proportional to the product of the number of humans and the number of zombies. However, when a human dies, they become a zombie about half the time. Humans also give birth to new humans at a

constant per capita rate, but only die from zombie attacks. Quantities: $h(t), z(t) =$ numbers of humans and zombies after t days.

13. A biologist classifies a certain species of fish according to one of three life phases: juvenile, reproductive, or elderly. Juveniles become reproductive after an average span of four weeks, reproductive fish become elderly after an average of eight weeks, and elderly fish die after an average of two weeks. In addition, reproductive fish give birth to juveniles at an average per capita rate of ten per week. Today, there are 500 juvenile fish, 300 reproductive fish, and 200 elderly fish. Quantities: $j(t), r(t), e(t) =$ numbers of juvenile, reproductive, and elderly fish which will be present after t weeks.