

# Week 9 Seminar: Deterministic Models

MATH DIVULGED

July 10, 2020

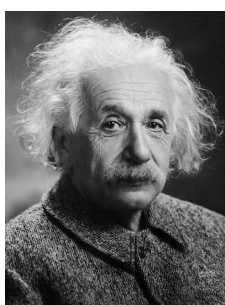
## §1 Introduction

The goal in this handout and lecture is to introduce students to how mathematics is applied in the modern world to real problems. I hope you will have a better idea of what pursuing a career in mathematics could be like, and what mathematics can be beyond just the competition level.

This will be a very long handout with a lot of information - feel free to scroll around, discover what topics are being discussed, and read what you are interested in. In fact, this handout barely scratches the surface to all the various deterministic models. There is no pressure for you to want to explore everything; instead, focus on what intrigues you the most.

### §1.1 Models

What is a *model*? To put it simply, a model is a representation of reality. As humans, we strive to understand the universe around us, and throughout history, we have derived tools such as mathematics, science, and other forms of thinking to further our understanding. However, it is easy for students and even experienced scientists to mistake science as an exact art.



*"As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality." -Albert Einstein*

We should be clear that mathematics is only a tool invented by humans to interpret and predict reality, and is not exactly how reality always works. Understanding this, you must realize that models are the basis for scientific understanding. People create models to **approximate** the world around us through observations and mathematical influence.

Models can come in all shapes and sizes. Simple models can be easy to create, require little data, but are less accurate and detailed. Advanced models will be tediously difficult to get right, require more parameters and data, but are more accurate and oftentimes, rewarding as a result. **But**, that does not mean that simple models are useless. Simple models can show a broad interpretation of what is happening in a scenario, which can provide crucial intuition and understanding.

For example, we can take the classroom example of a ball dropping. When you first learn physics, you learn equations that describe the parabolic motion of objects due to gravity, and your teachers will always remind you that these equations apply when you neglect air resistance. Air resistance is a very complicated subject to learn, and to predict. We can spend hours creating models to analyze air flow around a moving ball, but, if we just want to understand the basic movement of the ball, a simple model in physics without air resistance is sufficient. It will tell us approximately how the ball will travel with a bit of deviation from reality, but it is enough. We won't always need a near-perfect model to understand the world around us.

## §1.2 Uses

Models usually have a purpose. People want to do things with models.

A relevant example to 2020 is the Coronavirus pandemic. How do governments decide on policy and economic decisions? Mathematicians work with biologists and other professionals to create detailed models in an attempt to predict what factors could influence the spread of the disease throughout a community. These models and predictions that are given to policymakers help influence officials to enact laws, help, and general advice. These numbers can't be perfectly accurate, but they are attempts by scientists to provide foresight, and they are certainly more accurate than blind guesses.

Managing the wildlife and ecology of a country is also often crucial to governments. No one wants to be responsible for the extinction of a valued species, or the destruction of beautiful ecosystems. But the actions of hunters, the introduction of new species, or even accidental chemicals and pesticides can all significantly impact biomes and the organisms within.

How do we approach a problem with modeling? The general gist of modeling is to observe behavioral trends of some quantity, and use known functions to relate those into a mathematical formulation. I'll provide an example:

### Example 1.1

A population of a single species occupies a closed area with a limited quantity of natural resources. Model the size of the population with time.

This seems like a very daunting task. However, we can make some clever observations about the scenario. The rate of change of the population should be somewhat proportional to the size of the population. The larger the population, the more possible reproduction. So, we can conclude that

$$\text{Rate} \sim \text{Size}.$$

Second, a smaller population should grow near proportional to the size, but a larger population should be inhibited by the limited resources. This species can't multiply forever, because they share the same pool of resources. So, as a population gets larger, there should be a factor multiplied onto the size that should reduce the rate of the population's growth. A function which can do this is the linear equation (Constant - Size). A larger size means that this value is smaller, reducing the rate. Thus,

$$\text{Rate} \sim \text{Size} \cdot (\text{Constant} - \text{Size}).$$

This is a great mathematical formulation of the population's size - it does make several assumptions and shortcuts - but it is a good approximation to provide a basic understanding of populations.

The only problem is, how can we solve this relation in terms of mathematics?

## §2 Techniques

The term *deterministic* refers to the concept of being "determinable". In models, this implies a model will always give one distinct answer, a result that can be computed. The opposite, *stochastic* models, will be discussed in the next seminar handout.

Deterministic models can use a variety of techniques, but a popular theme is the use of *differential equations* (DE), or in other words, equations involving derivatives. Why do derivatives appear so often? Derivatives are analogous to rates, and observing details about rates can oftentimes be easier than observing details about the actual quantity. For instance, with example 1.1, we observed the rate of the population's change. Thus, knowing how to handle derivatives in modeling is a crucial aspect to creating models successfully.

Upon first thought, it seems that integration should be able to solve DEs. However, not only is integration often incapable of dealing with DEs effectively, integrals are often impossible to solve by hand as well. Modeling often presents challenging DEs which yield no simple solutions. So, mathematicians employ the most powerful weapons of the twenty-first century: the computer. In this lesson, I will only demonstrate the basics for a computational approach to DEs.

### §2.1 Derivative Definition

For more on calculus, visit our handout from the Calculus seminar. However, I will briefly introduce the concept of a derivative.

A slope between two points can be defined as the change in  $y$  over the change in  $x$  on a graph. A *derivative* is a slope between two points on a graph; however, those two points are unimaginably close to each other. It is for this reason, that a derivative is known as an "instantaneous slope": it is the exact slope of the curve at a specific exact point. Mathematically, we can write:

**Definition 2.1.** A **Derivative**  $f'(x)$  or  $\frac{dy}{dx}$  of a function  $y = f(x)$  satisfies

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The value  $\Delta x$  is the infinitesimal increase from  $x$  to its next point for the slope calculation,  $x + \Delta x$ .

There are plenty of neat tricks for calculating derivatives that were discussed in the calculus seminar handout, but we will not really need them here.

## §2.2 Discretization

In mathematics, the term continuous pertains to no breaks or stops within a domain. On the other hand, discrete refers to broken intervals or specific points. As mathematicians, we often deal with continuous functions and problems. Reality is, for the most part, continuous. However, computers will commonly struggle, as their main strength is in computing numbers, not equations. Thus, our task is to convert the definition of a derivative into a numerical process that computers can employ. There will be no real discussion of computer science here, only math.

In the definition of a derivative, we assumed that  $\Delta x$  approached 0 in the limit. This gives us the exact derivative. However, if we instead let  $\Delta x$  be an incredibly small value close to 0 but not 0, we can numerically approximate a derivative at a certain point. From now on, we will treat  $\Delta x$  as a very small number - this will often be chosen by the designer of a model.

Using the definition of the derivative, we can come up with the approximation:

$$f'(x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

which can rearrange into

### Theorem 2.2

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x.$$

Using this technique, we can solve for discrete values of a function  $f(x)$  with intervals of  $\Delta x$ . In actual computation, we can pick varying values of  $\Delta x$  to accommodate for the function's properties, but this is an advanced topic.

The error in Theorem 2.2 is proportional to  $\Delta x^2$  - search up Taylor expansions if you want to learn why. This error may seem dramatic; however, the error on this is sufficiently small if we choose a sufficiently small  $\Delta x$ .

### Example 2.3

Given that  $f'(x) = 2x$  for all  $x$ , and that  $f(0) = 1$ , use Theorem 2.2 to estimate  $f(n)$ .

*Solution.* We can pick a very small value of  $\Delta x$  where  $n = m\Delta x$  for an integer  $m$ . Thus, using Theorem 2.2, we have that

$$f(\Delta x) = f(0) + 2 \cdot 0 \cdot \Delta x = f(0).$$

$$f(2\Delta x) = f(\Delta x) + 2\Delta x^2 = f(0) + 2\Delta x^2.$$

$$f(3\Delta x) = f(2\Delta x) + 2 \cdot 2\Delta x^2 = 6\Delta x^2.$$

...

$$f(n) = f(m\Delta x) = f(0) + m(m-1)\Delta x^2.$$

□

We were able to solve this by hand in terms of the variables, but a computer would also be able to quickly plug in numbers such as  $\Delta x$ ,  $m$  and  $f'(x) = 2x$  to quickly compute any value as well. For more difficult derivatives, we would struggle more to solve for the final function, but a computer can still breeze through calculations, as long as it is computing in a smart way.

### §2.3 Linearization

For Example 2.3, the recursive method we used to compute  $f(n)$  could be coded into a computer; however, there is a more general method commonly used, which I will now present.

*Solution.* By Theorem 2.2, we can write  $f(k\Delta x) = f((k-1)\Delta x) + 2(k-1)\Delta x^2$  for some integer  $k$ . We have already discretized a domain of numbers using the interval  $\Delta x$  from 0 to  $n$ , so we know we are searching for the values of  $f(k\Delta x)$  at each  $x = k\Delta x$ . Let  $a_k = f(k\Delta x)$  for each  $k$  from 0 to  $m$ . Then, we have

$$a_k = a_{k-1} + 2(k-1)\Delta x^2 \implies a_k - a_{k-1} = 2(k-1)\Delta x^2.$$

Since we know the numerical values for  $\Delta x$  and  $k$ , the right hand side is just a numerical value while the left hand side contains our unknown variables. So, our  $m+1$  equations for each  $k$  form a system of linear equations.

If you are familiar with matrices, a system of linear equations can be written into matrix form

$$B \cdot A = C$$

where  $A$  is our vertical vector of  $a_k$  for each  $k$ ,  $C$  is a vertical matrix of  $2(k-1)\Delta x^2$ , and  $B$  is a matrix of  $-1, 0, 1$  for the coefficients of  $a_k$ . A computer can easily calculate  $C$  and  $B$ , and thus, using algorithms for matrix inversion, the computer can calculate all the values of  $a_k = f(k\Delta x)$  at once:

$$A = B^{-1} \cdot C.$$

□

Although as humans, this is not a closed-form solution for us, it is a method to employ for computers, who can perform these calculations incredibly quickly.

Another point to discuss is how solutions for differential equations are generated. Having an equation is required, but there also needs to be known **boundary values**, or known points of specific variables. For instance, given a slope for a linear equation, it is impossible to pinpoint exactly what the points on the line are without being given at least 1 point on the line. The same is true for differential equations - we need a starting point. In the example above, the starting point is the known  $f(0) = a_0 = 1$ .

Furthermore, unlike Example 2.3, differential equations will often have derivatives equal to an expression in terms of the variables there. This complicates our equations after applying Theorem 2.2, as they are not necessarily linear. Thus, using approximations and known states of variables, we can create linear equations which often involve many, many variables. These can be solved simultaneously using this method of matrices. For more information, you can search up the "Finite Difference Method".

If you didn't understand this method fully, perhaps because of a lack of experience with matrices, don't worry! At your age and level, you don't need to be concerned over the exact details. Instead, I hope you will take away the general concept of numerically solving DEs, and the ideas and examples of modeling I have yet to show.

## §2.4 Other Methods

There are a wide variety of different computational approaches designed by mathematicians to solve DEs, and Ordinary DEs - a class of equations which have special properties. These algorithms and solvers are widely employed by mathematicians in projects, and are often more convenient than hand-designing a method. If you are interested in these methods, feel free to search up Runge-Kutta as a great starting point.

However, discretizing can sometimes only be half the battle. If we ever need to find a point in between two discretized points on our domain, one possible technique to use is interpolation. There are many computer algorithms and methods for interpolating a value in a domain given a set of data.

Given the wide variety of techniques to solve DEs, we will depart from discussing the techniques, and spend the rest of the handout discussing common and interesting applications of DEs to different fields in reality.

## §3 Population Dynamics

Ecology is one aspect where modeling is incredibly applicable due to the volatility of ecosystems to change, whether the change is an invasive species, climate change, deforestation, or even just hunting. Studying how populations in an area grow and shrink becomes possible simply from observing logical traits of species in regards to their growth and death rates.

### §3.1 Single Species Dynamics

We will return to the example posed in the introduction section.

#### Example 3.1

A population of a single species occupies a closed area with a limited quantity of natural resources. Model the size of the population with time.

Previously, we had formulated the proportionality

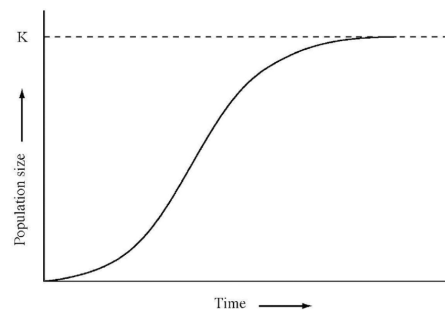
$$\text{Rate} \sim \text{Size} \cdot (\text{Constant} - \text{Size}).$$

We can represent these quantities with variables, using  $P$  for the size of the population and  $t$  for time:

$$\frac{dP}{dt} = aP \left( 1 - \frac{P}{K} \right),$$

where  $a, K$  are two constants dependent on data. I should note that this equation can also be interpreted, rather as a product of two facts, as the difference of two quantities:  $aP$  and  $aP^2/K$ . These can be considered as the growth and death rates of the population, try to convince yourself that as an exercise.

This differential equation for  $P$  in terms of  $t$  is certainly solvable with ordinary calculus techniques and a bit of manipulation; however, a computer is capable, using the variety of methods available for solving DEs, to numerically solve for  $P$  at discrete points once given a boundary value: preferably,  $P(0)$ . The closer the points are to each other, the better result we get.



The resulting graph for  $P$  is known as a **Logistic Growth** curve. As you can see, the population will slow down growth and approach a specific number - this is known as the **carrying capacity** and is equivalent to the value of  $K$ . This can be verified by checking numerical results, or by simply plugging in  $K$  into the derivative equation and finding  $dP/dt = 0$ .

This case was relatively simply, but rarely in nature will there be a scenario in which only 1 species interacts with itself. Usually, for animals, there are predators, food sources, etc. Thus, for a more comprehensive model, we will want to aim for a model revolving around multiple species.

### §3.2 Predator-Prey Dynamics

The simplest representation of a food chain involves just one species which acts as prey, and one species which acts as a predator. The concepts and observations to derive a model are barely different from a single species scenario; the rate of change of one species becomes dependent not only on itself, but also on the other species.

Take a predator and prey relationship between two species,  $X$  and  $Y$ , respectively. Let the population of  $X$  be  $x$  and the population of  $Y$  be  $y$ . Since we know that the rate is proportional to how many individuals the species has, we should analyze the rate per population, or  $x^{-1}dx/dt$ .

For a single individual, there is a constant reproductive rate for a single individual - this individual won't reproduce significantly faster just because there is a larger population.

Similar to single species, species X has intra-competition (competition within the species) for the same resources, so this individual we are analyzing should have a lowered chance of surviving (a higher death rate) when the size of X is larger. Lastly, the more predators there are - species Y - the more likely the individual is to die. Thus, as a mathematical equation,

$$\begin{aligned}\frac{1}{x} \frac{dx}{dt} &= a_0 - a_1x - a_2y \\ \implies \frac{dx}{dt} &= x(a_0 - a_1x - a_2y),\end{aligned}$$

where  $a_0, a_1, a_2 > 0$  are constants. Remember, these constants are either dependent on data, or are picked to demonstrate scenarios.

Similarly, for predators the equation ends up as

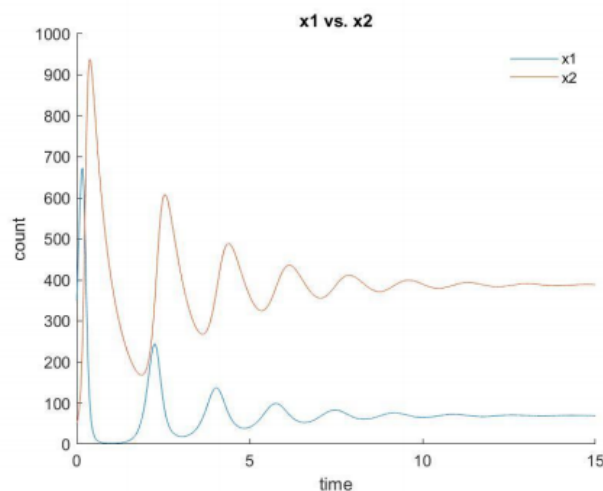
$$\frac{dy}{dt} = y(b_0 + b_1x - b_2y)$$

for constants  $b_0, b_1, b_2 > 0$ . The only difference with the prey equation is that the coefficient on  $x$  is positive in this case - more prey should improve the survivability of an individual in Y.

A simpler model neglecting intraspecies competition is widely known as the **Lotka-Volterra** equations. However, by including this degree of competition, we add complexity to our model, generating more accurate models at the expense of requiring more parameters.

Thus, our model arrives at system of differential equations. Fortunately, this can be approached exactly the same as discussed in section 2, with either a linearized matrix containing discretized points for both  $x$  and  $y$ , or using built-in ODE solvers and algorithms which can handle vectorized differential equations. Note that by setting the vector  $(x, y)$  as a vector  $\vec{w}$ , the system of differential equations can be expressed as a single vector differential equation using some linear algebra, which is perfectly solvable via computer.

However, the points of interest don't arise solely from solving a given model, but rather in analyzing the various results that occur from various starting values  $x(0), y(0)$ .





In this example, we can clearly see that the populations of the species oscillate back and forth, but seem to approach a stable level after a period of time. This should make sense - a change in an ecosystem (the introduction of the predator and prey in our case) should cause a chaotic start which slowly tends to a stable existence.

There is a lot of mathematics behind calculating and predicting fixed points and equilibriums in ecosystems not just of 2 species, but of many. Much of the details are a bit advanced for the style of this handout. If you are interested in deriving results on your own, take the model in the Multispecies Dynamics section and analyze the partial derivatives - this should give you a system of linear equations which can be solved using matrices.

However, oscillations don't always tend quickly to "equilibrium". In a few rare cases, where the coefficients are just right, we can actually observe oscillations in reality. One famous example is the **Lynx-Snowshoe Hare Cycle** where a species of wild cat and rabbit oscillate every couple years. The reality of it is much more complex than a simple 2-species predator-prey relationship, but the key point is that such a simple model was able to reveal the possibility of oscillation and the resulting behavior, providing key intuition on the plausibility of certain theories and hypotheses.

### §3.3 Multispecies Dynamics

An ecosystem in reality is often teeming with life of all sizes and types - the number of relevant species is very large, and they can all play a crucial role in maintaining an equilibrium within a specific region.

Regardless, some species will have larger impacts on an ecosystem than others. If a particular species is impacted and causes a dramatic change on the state of an ecosystem, the species is referred to as a **keystone species**. There can be many keystone species in an area, but to model more realistic ecosystems, we must add a layer of complexity to our models.

We can extend the model from the predator-prey scenario by generalizing the concept used. Like our analysis there, we will analyze the rate of a population's growth per size, or the rate of a single individual. It is easy to quantify growth as the difference between factors that encourage survivability versus factors that promote death.

Assume that there are  $n$  species with sizes  $x_1, x_2, \dots, x_n$ . Each species should have different rate of death; some animals will survive longer without sustenance and some will die faster. We will analyze the general species  $x_i$ . There exists a decay or death rate  $D_i > 0$ . If the ecosystem has only that one individual in it,  $D_i$  is the rate at which it dies.

Now, taking other species into account, there is a "relationship" constant  $r(i, j)$  between species  $x_i$  and  $x_j$  which is determined by the interaction of the two species. If  $i$  feeds on  $j$ , then  $r(i, j) > 0$  while  $r(j, i) < 0$ , and vice versa. If the two competitions don't prey on each other, then the relationship should be 0.

Taking decay and interspecies relations into account, the model can be summed up as the system of differential equations of the form

**Theorem 3.2 (A Multispecies Model)**

$$\frac{dx_i}{dt} = x_i \left( -D_i + \sum_{k=1}^n r(i, k)x_k \right).$$

There can be obvious changes and modifications which can be made, and there are certainly models which use a completely separate way of thought. However, we designed this model logically, and thus, the results should be logical as well. As general practice, one of the species, say  $x_1$  without loss of generality, should be set as the "sun". The sun has an unlimited quantity of energy which can be absorbed by plants, which act as predators of the sun. To implement this function, per iteration of the equation solver, the population size  $x_1$  should be fixed at a constant value, thereby powering the remainder of the food chain.

Like the predator-prey scenario, many oscillations should occur, but in most scenarios, species will eventually tend to a stable size until a change erupts. You can imagine a  $n$ -dimension region with  $n$  different hyperplanes (derived from partial derivatives) intersecting. These fixed regions can be pre-computed with methods briefly mentioned earlier.

Without a large amount of data and field information, it is hard to determine accurate numbers of  $D_i$  and  $r(i, j)$  for a specific, real ecosystem. However, that does not mean you cannot implement this model and experiment with keystone species.

For instance, Yellowstone National Park faced a crisis in the past century - the over-hunting of wolves ended up causing serious damage to the ecosystem, as the predators turned out to be a crucial keystone species. To see an example of this in action, you could select a number of important species in Yellowstone and create a food chain with the model. Then, at a certain point in time, manipulate the rates and constants so that wolves drop quickly in number and observe what happens to the ecosystem. Some questions which could be partially answered are: How will slowly reintroducing wolves impact the ecosystem? What other possible solutions could work? What animals and plants are also keystone species that should not be touched?

Other possible uses of a multispecies model would be analyzing the introduction of invasive species. Due to globalization of the world, animals are being spread around to different ecosystems, often creating damaging changes to the existing food chains. How can we mitigate damage from a specific invasive species? What type of animals could pose dangers if introduced to the ecosystem? What other animals or plants can we introduce to combat existing invasive species?

## §4 Epidemics

As mentioned in the introduction to this handout, modeling epidemics is a serious task which could help prevent critical damage to entire regions if done effectively. Models are used to help predict the severity of diseases, where the disease is heading, and how intense current infections are. This information can be important context for policymakers and

governments to take action and limit the spread and threat of a disease.

Although it is undeniable that the US could have taken quicker and more efficient action earlier with the Coronavirus Pandemic, mathematicians all over the country in universities and organizations like the CDC have made a multitude of different models to help our country. We should take time to thank not only our nation's first responders, medical staff, and researchers, but also those modeling teams who have helped supply our country with their best guesses for what to do to prepare for, prevent, and minimize harm.

In this section, I will cover the most basic concepts in modeling diseases, discuss the impacts of vaccines and measures such as quarantine, and also mention a few possible extensions to this basic model to improve its application to reality.

### §4.1 SIR States

One popular technique used in modeling is the idea of states. A "state" is a particular category of properties that we can use to classify a specific object. For instance, water has three main states in terms of phases: ice, liquid water, and steam. A traffic light has three states in terms of color: red, yellow, and green. Dogs can be classified in two states: dogs under 10 years of age, and dogs over 10 years of age.

The real magic behind using states in models is through fluidity. If an object can transition between its possible states, we can use mathematics to quantify the number of objects in each state at a point in time, or the probabilities of being in specific states (this will be covered in the stochastic seminar).

For general diseases, people can be classified in four states: susceptible, infected, recovered, and dead. Someone who is susceptible can still contract the disease, someone who is infected is currently experiencing the disease, someone who is recovered has already had the disease but now has built up immunity to it, and dead implies those who unfortunately died to the disease.

One of the most popular, simple models for diseases only looks at the first three states: susceptible (S), infected (I), and recovered (R), thereby earning the name "SIR". In this model, we will use  $S$ ,  $I$ , and  $R$  to denote the number of people in the population who are in each respective state.

The common SIR model makes several assumptions regarding the scenario. First, the population remains constant - no one comes in and no one leaves. Mathematically,  $S + I + R$  is constant and is equal to the population size. Second, the population must be sufficiently large - the more people, the more accurate average rates and fractions will be, and the more we can assume everyone to have ordinary, random behavior. There are more rigorous assumptions to make, but for the purpose of understanding, these are sufficient.

### §4.2 SIR Equations

Imagine being part of the infected state. As an individual, you are waiting on time - it takes a certain amount of time per person to recover from a disease. In reality, this time varies depending on health, medication, and many other bodily factors. However, on

average, with a large number of people, we can assume each person recovers at the same rate and expect fairly accurate predictions of recovery rates.

Thus, as an individual, the rate at which you "fractionally" transition to the recovered state is constant. We can make claims regarding fractions of people since we are dealing with a big population - these fractional constant rates multiplied by  $I$  will give an approximation to how many people are actually transitioning per unit time. Thus, the rate at which  $R$  grows (people recovering) is proportional to  $I$  times a constant, or in terms of derivatives,

$$\frac{dR}{dt} = \gamma I$$

for a positive constant  $\gamma$ . It is worth checking that the derivative is positive, since we are increasing the size of  $R$  with recoveries.

Each person in the susceptible state share the same rate at becoming infected, on average. However, the rate is no longer constant like it is with recovery. As one individual, you have a higher chance of being infected with the disease if there are more people in the population who are in the infected state. Thus, the rate for an individual to transition to the infected state is proportional to  $I$ . Multiplying these individual rates by the overall susceptible population  $S$  implies that the rate at which  $S$  decreases is proportional to  $S \cdot I$ . In terms of derivatives,

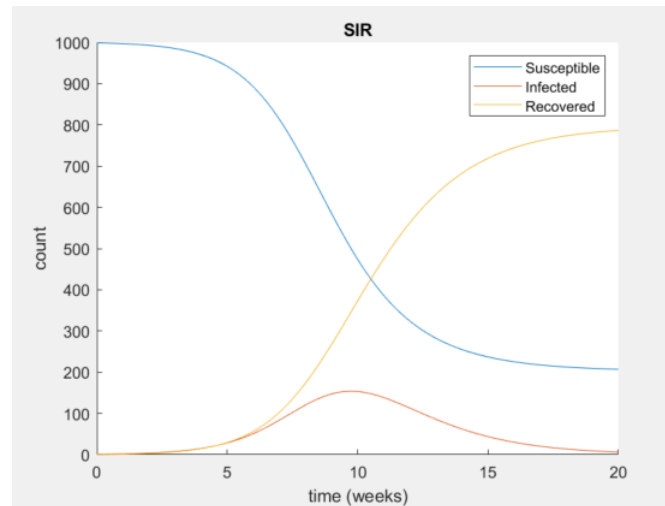
$$\frac{dS}{dt} = -\beta SI$$

for positive constant  $\beta$ . The derivative is negative here, because the susceptible population drops when people are infected.

Taking into account the fact that  $S + I + R$  is a constant, the derivatives of  $S, I, R$  must sum to zero (otherwise, the population is changing!). So,

$$\frac{dI}{dt} = -\frac{dS}{dt} - \frac{dR}{dt} = \beta SI - \gamma I.$$

In most epidemics, the population is a majority in the susceptible state, and there are only several individuals (a miniscule fraction of the population) who are in the infected state, and no one in the recovered state. From these deductions, which arose from observations and assumptions, we have three variables, three starting points, and three differential equations. As mentioned in the previous section, systems of differential equations are just as easy to solve for a computer as a single equation due to linear algebra. That's not to say it is impossible to deduce key properties of the solutions by hand, but there is not too much to gain from that.



From the example graph, we can see three curves for each of the three states. The infected curve is probably the most interesting out of the three: it resembles a little bump very similar to a bell curve - you could verify this mathematically. However, the shape itself reveals key information about when the epidemic is at its worst - the time when  $I$  is the largest is known as the peak of the epidemic.

It is also important to understand that  $I$  is not an indicator of how many people have been infected in the population since the beginning of the epidemic; it only serves as an indicator of how many people at a time  $t$  are in the infected state. To compute how many people have been infected in total at a time  $t$ , we can take a look at

$$I(t) - I(0) + R(t) - R(0) = S(0) - S(t).$$

This quantity tells us how many people are currently infected, and how many have already recovered. We can find how many people were infected over the course of the entire epidemic by evaluating  $S(0) - S(\infty)$ , or any significantly large number.

There is also importance regarding the two constants in our equations,  $\beta$  and  $\gamma$ . It should be pretty obvious that a larger  $\beta$  implies that the disease is more contagious, since it increases the rate of contraction. A larger  $\gamma$  implies a less serious disease, since people are recovering at faster rates. Any disease where the recovery is faster than infection is doomed to die out, while the opposite signals an outbreak.

In fact, an outbreak occurs when  $dI/dt > 0$  for an initial period of time where  $I$  is small while  $S$  is large. Using our equation,

$$\beta SI - \gamma I > 0 \implies \frac{\beta S}{\gamma} > 1.$$

This is known as the **reproductive number**  $R_0$ . The higher a disease's reproductive number, the higher the infected peak rises, and the more threatening it is to society. There is also motivation in this expression to analyze the SIR model in terms of proportions - that is, to have  $S, I, R$  to represent the fractions of the population that are in those respective states. As such,  $S(0) \approx 1$  and so the reproductive number becomes  $R_0 = \beta/\gamma$ .

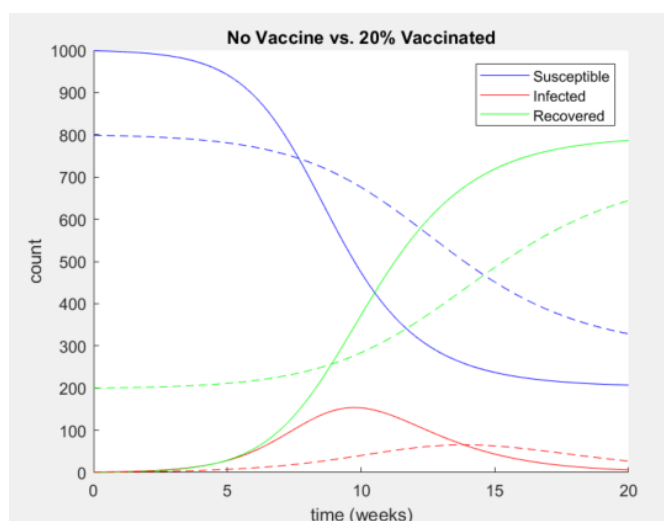
For context: Ebola is under 2; the common cold is between 2 and 3; COVID-19 has an estimated number of 3.28; measles, a pure, air-transmission disease, is above 12.

### §4.3 Vaccination and Prevention

Since the creation of the vaccine, humanity isn't completely helpless to powerful epidemics. Aside from the silly anti-vaxxer movement, many people rely on vaccinations to reduce their risk of contracting a certain disease. Biologically, vaccines operate by inserting a weakened variation of the disease into one's body, allowing the body to learn how fight it off and produce necessary antibodies to protect itself from the disease in the future.

Mathematically, there are two classifications for vaccines: **Leaky** and **All-or-Nothing**. Leaky vaccines are vaccines which "partially work" and only reduce the chance of individuals contracting the disease, which is a decrease in  $\beta$ . All-or-nothing vaccines are vaccines which provide certain individuals complete immunity to the disease while providing no benefit for everyone else - this is a decrease in  $S(0)$  and an increase in  $R(0)$ . In reality, all vaccines are not one or the other; they are in between, meaning some people can get near full immunity while others get reduced chances, and even some may experience no benefit at all. These results all depend on the vaccine, and the person who is being administered the vaccine.

Using the mathematical interpretations of the two vaccine types, we can incorporate these and compare the results of vaccination to the results of no vaccination. Below, we will analyze the all-or-nothing vaccine, where 20% of the population has been successfully immunized.

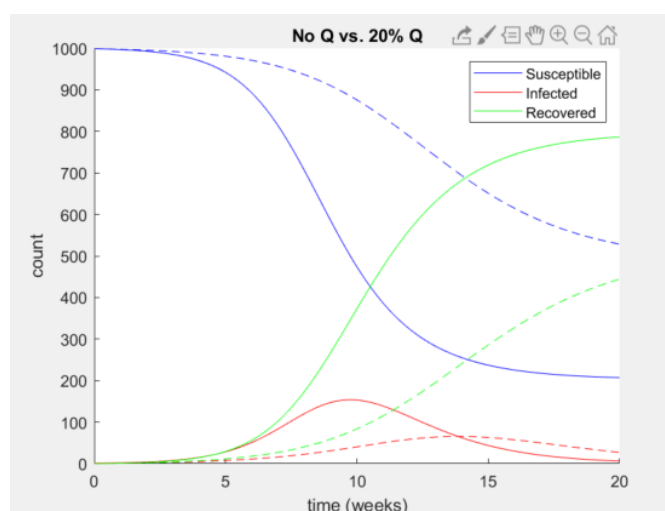


The dotted line represents the vaccinated scenario, and we can observe the clear differences vaccination produces. First, the peak of the infected curve is lower. Second, the infected peak is shifted rightwards, implying a longer epidemic. Third, checking  $S(0) - S(\infty)$  correctly tells us that although less people recovered overall, there were less total infections. Is this good or bad?

The third observation is undoubtedly an advantage to vaccination. However, the first two may be met with skepticism: do we want a shorter, more intense epidemic (the "get-it-over-with" mindset), or a prolonged, less intense epidemic? For a serious disease which hurts the economy, some people may instinctively opt for the first option. In the short run, getting the economy back up faster is great. However, should we risk people's health for a better economy? Furthermore, high infected peaks can overstress hospitals

and medical workers, as these institutions can only manage a limited capacity of patients. A longer epidemic can provide hospitals more time and space to care for those who have been infected at one point in time. A majority of individuals in this scenario will choose the moral, latter option. Since many vaccines are produced by private companies, and are only regulated by the government, they are typically always publicly available. The government typically does not have the power to withhold a safe, successful vaccine.

However, there are other preventive measure similar to vaccination that the government can enforce. In our pandemic today, quarantine has become a top method endorsed, or un-endorsed, by governments across the world. By isolating oneself from society, an individual can essentially "remove" themselves from the susceptible state, reducing  $S$  very similarly to an all-or-nothing vaccine. We can solve this scenario and provide a comparison.



The dotted line indicates the scenario where 20% of the population has quarantined themselves. Computationally, the  $S(0)$  value was decreased, but added back in when producing the graph to provide a more intuitive quantity for the susceptible state (those in quarantine are technically still susceptible).

Like the vaccination, the infected peak is lower but shifted to the right, while the total number of infected  $S(0) - S(\infty)$  has clearly lowered for the quarantine scenario. It is easy to observe that a mere 20% quarantine ratio can have such an astounding impact on a disease's impact. It is worth mentioning that the disease used in these examples has a reproductive number of 2 - it is relatively contagious and comparable to stronger strains of the seasonal flu. For more contagious diseases, preventive measure become even more crucial.

One popular biology concept is **herd immunity**, which is the notion that by decreasing the susceptible population significantly, the infected count will be so low that we protect some of those who are still susceptible - typically those with problems and are incapable of vaccinating or quarantine measures. If there were no preventive measures, these individuals would have a much higher chance of having been infected. Herd immunity is easily observable in the above graphs.

Lastly, it is crucial to mention that the SIR model is in no way, a perfect model for epidemics. It lacks sufficient states, and makes assumptions about the population that

ignore population density, clusters, communities, etc. However, specifics aside, it does encompass a large fraction of the general behavior of diseases, and thus provides good insight on how an epidemic can progress. It's a great starting point, and also a great tool for teaching epidemiology. Different techniques can be added to create better disease models, and some of these will be mentioned in the stochastic lecture.

## §5 Conclusion

Differential equations are certainly not limited to just the couple examples that we've discussed in this handout. Furthermore, deterministic models are not only limited to just pure differential equation problems. There is an unimaginable variety of strategies that can be employed to produce deterministic results.

Some other examples for you to think about:

Traffic can be modeled by implementing cars as objects with a set of information. Speed, acceleration, and position can all be related via differential equations - for instance, a car will accelerate if the car in front of them is farther away while hitting the brakes when the car in front is too close. By expressing everything with DEs, mathematicians can model traffic to locate congestion and plan roads, intersections, lanes, and traffic lights accordingly.

Queues, or lines, in restaurants, stores, and amusement parks can heavily impact the number of customers waiting. Rates and distributions can be used to model queue lengths and revenue. How many cashiers should a fast food restaurant have? Too many, and there will be too many times during the day where there are too few customers in line - this is overpaying the workers and can cause a loss of income. How much should an amusement park value a "fast pass" at? How many visitors would they want occupying the fast lane? How can parks optimize the number of people in rides?

And then there's physics and engineering. Nearly all physical concepts have relations to differential equations, and there are countless methods devised to tackle different scenarios.

Modeling isn't just a small branch in mathematics with mild relevance today; it's a full blown field that stretches across the entire spectrum of human development. Since the birth of mathematics, math has been a tool to aid human kind in exploring, understanding, and developing the world. Gauss, the prince of mathematics, even helped astronomers model the orbit of Ceres, a dwarf planet, when it had supposedly gone missing. Good luck on all your competitions and learning, and I hope you will stick with math when you grow up as well.