Math 360-1: Applied Analysis
Northwestern University, Lecture Notes

Written by Santiago Cañez

These are notes which provide a basic summary of each lecture for Math 360-1, the first quarter of "MENU: Applied Analysis", taught by the author at Northwestern University. The book used as a reference is the 2nd edition of *Differential Equations: A Modeling Perspective* by Borrelli and Coleman. Watch out for typos! Comments and suggestions are welcome.

Contents

Lecture 1: Modeling with ODEs

A *differential equation* is an equation which relates an unknown function to one or more of its derivatives. For example, $y' = y$ is a differential equation characterizing functions $y(t)$ which equal their own derivative. In this case, all solutions (i.e. functions which satisfy the given equation) are of the form $y(t) = ce^t$. More generally, we can consider equations of the form

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

where f is some function which encodes a relation between the function y and variable t. Equations involving single-variable derivatives like this are called *ordinary differential equations* (ODEs for short), while equations involving partial derivatives of multivariable functions are called *partial differential equations* (PDEs); this particular quarter focuses on the first type, with the second type to be studied in the second quarter. Equations involving only first-derivatives are called *first-order*, those involving second derivatives are called *second-order*, and so on. Certain differential equations are *linear* whereas others are *nonlinear*, and some are *homogeneous* (also called *non-driven*) while others are *inhomogeneous* (or *driven*); these are all terms we will define as we go.

Differential equations are the cornerstone of many modern applications of mathematics, since it is often the case that when describing some phenomena one can derive a relation between a function of interest and its derivatives, and it is then a differential equation which models this behavior. A goal of this course, apart from learning how to find solutions of differential equations and interpret/analyze those solutions, is to understand *why* a certain differential equation is the correct one needed to model a given scenario, and what the form of the equation tells us about the behavior of the system in question, even without having explicit solutions available. For instance, the first-order nonlinear differential equation

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

models a function $y(t)$ which describes some type of *population*—why? Why does it make sense to model population growth by this particular equation? Why not something else? The second-order linear ODE

$$y'' = -y$$

models the motion of a spring—why? Looking ahead to next quarter, the partial differential equation

$$u_t = u_{xx} + u_{yy}$$

models the behavior of a temperature function $u(x, y, t)$ across a 2-dimensional region—why? We'll come to understand precisely why this is, and how we can derive/analyze the solutions.

(We used the terms "linear" and "nonlinear" above, which for now we take to refer to simply the type of expressions involving y appear: linear equations are those in which y occurs to only a first power, whereas nonlinear equations have higher power of y or more elaborate expressions like $\sin y$. We'll give precise definitions of these terms later on, using the type of terminology developed in a linear algebra course.)

Compartment models. To start with, a simple class of differential equations we can study are those coming from *compartment models*. In such a model, the net rate of change of a function or quantity of interest $y(t)$ at a particular "time" t is equal to the rate of stuff coming "in" minus the rate of stuff going "out":

$$y' = (\text{rate in}) - (\text{rate out}).$$

For instance, we consider the population of fish in some lake, with $y(t)$ denoting the population at time t. We assume that the birth rate of fish is proportional to the population size, as is the death rate, so that both of these rates are of the form

$$\text{birth rate} = ay \qquad \text{death rate} = by.$$

The birth rate is the rate "in" and the death rate is the rate "out", so we get the equation

$$y' = ay - by$$

as one which models the population of interest. Suppose in addition that fish are harvested at a constant rate H, which should also be incorporated into the rate "out":

$$y' = ay - (by + H).$$

Let us write this more compactly as

$$y' = ky - H \text{ where } k = a - b.$$

The ultimate goal, if possible, is to find the function $y(t)$ which satisfies this equation. Finding such solutions explicitly will not always be possible for arbitrary differential equations, but in this case we can proceed as follows. For now, do not focus too much on *why* we are doing the manipulations we'll do; this is something we'll clarify later when we start talking about solving differential equations in a more formal manner.

The given ODE can be rewritten as

$$y' - ky = -H.$$

Multiplying both sides through by the function e^{-kt} gives

$$y'e^{-kt} - ke^{-kt}y = -e^{-kt}H.$$

Some intuition for why we did this comes from considering the related ODE $y' = ky$ with no "harvesting" term: in this case the solution is $y(t) = e^{kt}$ and so we might expect that even with a harvesting term, the function we are after should still be related to an exponential, so it makes sense (maybe?) to incorporate an exponential into our given equation. An important observation here is that, since e^{-kt} is never zero, we are not losing any information when manipulating our original equation into this new one:

$$y \text{ satisfies } y' - ky = -H \text{ if and only if } y \text{ satisfies } y'e^{-kt} - ke^{-kt}y = -e^{-kt}H.$$

This is crucial if we want to argue in the end that we will have found *all* possible functions satisfying the ODE on the left.

And now the magic: the left hand side we end up with can be written as the derivative of one single expression:

$$y'e^{-kt} - ke^{-kt}y = (e^{-kt}y)'.$$

Indeed, that this happens is the entire reason why we multiplied through by e^{-kt} in the first place! (We'll look at the more general version of this technique, known as the technique of *integrating factors*, soon enough. For now, you can just accept this as "magic".) Thus, our manipulated ODE can be written as

$$(e^{-kt}y)' = -e^{-kt}H,$$

which says that the function $e^{-kt}y$ should be an antiderivative of $-e^{-kt}H$. But any antiderivative of $-e^{-kt}H$ is of the form $\frac{1}{k}e^{-kt}H + C$ for some constant C, so we conclude that

$$e^{-kt}y = \frac{1}{k}e^{-kt}H + C \text{ for some constant } C.$$

Thus, after multiplying through by e^{kt} (which is never zero!), we find that

$$y(t) = Ce^{kt} + \frac{H}{k}.$$

Hence, all solutions of $y' = ky - H$ are of the form $y(t) = Ce^{kt} + \frac{H}{k}$ for a constant C. We call this the *general solution* of the given ODE.

Qualitative behavior. So, in the case of $y' = ky - H$, it is possible to describe all solutions explicitly. But even without this explicit solution, we can still derive a good amount of information about the behavior of solutions from the given ODE alone. This idea will be even more important in situations where it is *not* possible to describe solutions explicitly.

The constant function $y(t) = \frac{H}{k}$ is in fact a solution to the given ODE since $y' = 0$ and

$$ky - H = k\left(\frac{H}{k}\right) - H = 0$$

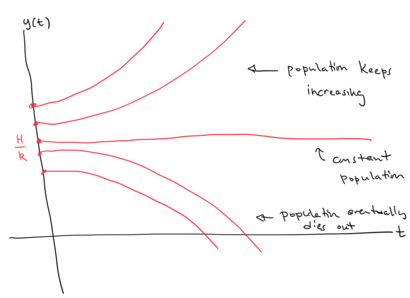
as well. (This corresponds to taking $C = 0$ in the general solution derived above.) This is the only constant solution. Now, suppose a solution $y(t)$ has "initial value" $y(0)$ which is larger than $\frac{H}{k}$. A general fact about ODEs we will look at later is that, under some mild assumptions, solutions are unique once we specify some "initial" value $y(t_0) = y_0$. (To be clear, "initial" does not necessarily mean the value at $t = 0$, but rather just any one value we are initially specifying.) In this case, this implies that no other solution apart from the constant solution $y = \frac{H}{k}$ can ever attain the value $\frac{H}{k}$, so that the *graph* of any other solution can never intersect the graph of this constant. In particular, for a solution $y(t)$ with $y(0)$ initially larger than $\frac{H}{k}$, *all* values of $y(t)$ will also be larger than $\frac{H}{k}$, and similarly if $y(0) < \frac{H}{k}$, then all values of $y(t)$ will be smaller than $\frac{H}{k}$.

So, consider a solution with $y(0) > \frac{H}{k}$, so that $y(t) > \frac{H}{k}$ for all t. (These solutions correspond to taking $C > 0$ in the general solution derived above.) Then

$$y' = ky - H > k\left(\frac{H}{k}\right) - H = 0,$$

so $y' > 0$ for such solutions, which means that these solutions are *increasing*. Note this also makes sense from the general solution: $y(t) = Ce^{kt} + \frac{H}{k}$ increases (exponentially to ∞ in fact) when $C > 0$, but the point is that we don't need the explicit solution in order to determine this qualitative behavior.

Similarly, a solution with $y(0) < \frac{H}{k}$ will have $y' < 0$ at all points, so such solutions *decrease*. Thus, we get the following rough picture of the behavior of our solutions:



In terms of our fish model, this says that for an initial fish population of $\frac{H}{k}$, the population will remain constant with the harvesting and death rates together exactly balancing out the birth rate; for an initial population larger than $\frac{H}{k}$, the fish population will be blow up and increase exponentially (so the birth rate over powers the death and harvesting rates); while for an initial population smaller than $\frac{H}{k}$ the population will decrease and eventually die out, with the death and harvesting rates overpowering the birth rate).

Lecture 2: More on Modeling

Warm-Up. Suppose we model a fish population as before, with birth and deaths rates proportional to the population size, but with a periodic harvesting rate given by the function $H(t) = 1 + \cos t$. So the harvesting is not constant, but varies and repeats every 2π units, in whatever units we are measuring time t. For a simplification, let us assume that the constants of proportionality describing the birth and death rates differ by 1. We claim that there is then a unique initial population $y(0)$ for which the population $y(t)$ is itself periodic as well.

We can model this scenario via the differential equation

$$y' = ay - (by + H(t)) = y - (1 + \cos t)$$

where ay is the birth rate, by the death rate, and our simplifying assumption means that $a - b = 1$. To solve this explicitly we use the same idea as before: find a way to manipulate this equation into one of the form

$$(\text{some function of } y \text{ and } t)' = \text{some function of } t,$$

after which we will be able to solve for y. This will always be possible (at least to the extent that we can find various required antiderivatives) for a first-order linear ODE, but we'll discuss this next time in detail. Again, for now, it is sort of "magic" that this works.

Rewrite the given ODE as

$$y' - y = -(1 + \cos t)$$

and multiply through by the nonzero expression e^{-t}:

$$e^{-t}y' - e^{-t}y = -e^{-t} - e^{-t}\cos t,$$

which does not affect the possible solutions. The left-hand side is now the derivative of $e^{-t}y$:

$$(e^{-t}y)' = -e^{-t} - e^{-t}\cos t,$$

so $e^{-t}y$ is an antiderivative of $-e^{-t} - e^{-t}\cos t$. Any such antiderivative (after using the technique of integration by parts on the second term, which would be worth brushing up on) looks like

$$e^{-t} - \frac{1}{2}e^{-t}(\sin t - \cos t) + C$$

for a constant C, so we get

$$e^{-t}y = e^{-t} - \frac{1}{2}e^{-t}(\sin t - \cos t) + C \text{ for some constant } C.$$

Thus, after multiplying through by e^{t}, we find that all solutions of the given ODE are of the form

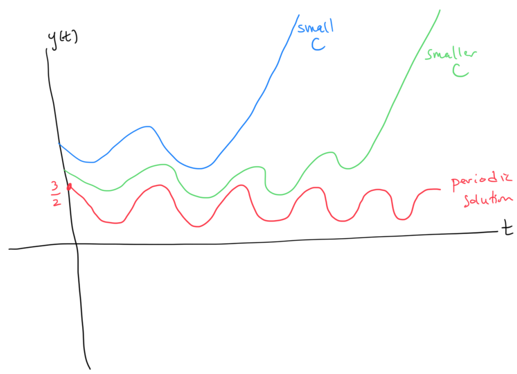
$$y(t) = Ce^{t} + 1 - \frac{1}{2}(\sin t - \cos t) \text{ for } C \text{ a constant.}$$

In order for this solution to exhibit periodic behavior, the exponential term Ce^t should not be present, so that $y(t) = 1 - \frac{1}{2}(\sin t - \cos t)$ will indeed be a periodic solution. Thus we need $C = 0$ and the initial population in this case is thus

$$y(0) = 1 - \frac{1}{2}(\sin 0 - \cos 0) = \frac{3}{2}.$$

Thus, for an initial population of $\frac{3}{2}$ (in whatever units we are using to measure population), the population will behave periodically.

For other initial populations, when $C \neq 0$, the population overall will not be periodic and will *eventually* behave exponentially. Now, for small values of C, for a while the periodic part $\frac{1}{2}(\sin t - \cos t)$ of the solution will overpower the exponential part Ce^t, so that the solution will "appear" periodic for some time, but eventually the exponential part overpowers the periodic part:



The smaller C is, the longer it takes for the population to no longer appear to be periodic, but it will get there for sure.

Quick observation. Let us highlight one more aspect of the ODE above, which was in fact also apparent in the model with constant harvesting rate. Recall the general solution we found above:

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

The periodic solution when $C = 0$ is one *particular* solution of the ODE $y' = y - (1 + \cos t)$. Now, the part of the general solution which can vary depending on C, Ce^t, is not a solution of this ODE, but rather it is a solution of the analogous *homogeneous* or *non-driven* ODE $y' = y$ instead! If we interpret the harvesting function $H(t) = 1 + \cos t$ as some "external" input which "drives" the behavior of the system, then the equation $y' = y$ which omits this describes the behavior without any such external driving force, which is why this is "non-driven". (What the book calls *driven* and *non-driven* are also commonly called *inhomogeneous* and *homogenous* respectively, but, as the book says, these latter terms can also have other meanings, which is the book uses driven and non-driven instead. I'll likely switch back and forth between the two, since I'm more used to inhomogeneous and homogeneous, but it should be clear form context what we mean. We'll clarify the difference between driven and non-driven equations more as we go on.)

So, the upshot is that the general solution to $y' = y - (1 + \cos t)$ can in fact be written as the sum of one particular solution of this equation and *arbitrary* solutions of the corresponding non-driven equation $y' = y$:

$$y(t) = \underbrace{Ce^t}_{\text{general non-driven solution}} + \underbrace{1 - \frac{1}{2}(\sin t - \cos t)}_{\text{particular solution}}.$$

This will be true for all *linear* equations, which is the main reason why linearity is so useful. We'll clarify this next time.

Non-linear model. The types of population models we have looked at so far, those of the form

$$y' = \underbrace{ay}_{\text{birth}} - (\underbrace{by}_{\text{death}} + \underbrace{H(t)}_{\text{harvesting}})$$

is probably not so realistic since, as we've seen, the solutions to such equations can exhibit exponential growth: in the "real world", a fish population in a lake should not be able to grow and grow without restriction since, intuitively, the lake environment itself should only be able to support so many fish. If nothing else, the number of fish which can actually be in the lake should be constrained by their size and the size of the lake, since you can't have more than one fish in the same physical location!

So, we need a better model. One way to get this is to note that the death rate should perhaps not actually be proportional to the population alone, but maybe should be proportional to the population squared. In other words, perhaps it should be true that as the number of fish increases, the death rate should also increase at a faster rate itself, so that more and more fish die off as the population begins to get too large. Thus, we might model the population using something like

$$y' = ay - (by^2 + H)$$

instead, where the death rate is by^2 instead of by. (In theory, we could incorporate an even greater rate of death by using larger powers of y, but it turns out that in practice taking a quadratic death rate like this is usually good enough.) This form of ODE $y' = ay - by^2 - H$ is called a *logistic equation*, and provides a much better model of population than the one we used previously.

Example 1. Consider the case where $a = b = 1$ and $H = 0$. Then our logistic equation is

$$y' = y - y^2.$$

As with the previous population models we had, it turns out that this equation can be rewritten, for $y \neq 0$ and $y \neq 1$, as:

$$\left(\ln\left|\frac{y}{y-1}\right|\right)' = 1.$$

We'll see how this can be derived later, but the point is that it will be possible here to obtain an explicit solution: the function $\ln|\frac{y}{y-1}|$ must be of the form $t + C$ since it is an antiderivative of 1, and from this an exact value of y can be obtained.

Example 2. But, we don't need explicit solutions in order to determine qualitative behavior. Consider the logistic equation where $a = 3, b = 1, H = 2$:

$$y' = 3y - y^2 - 2.$$

This has two constant solutions: $y = 1$ and $y = 2$, which are found by factoring the right side:

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

Thus, for initial populations of 1 or 2, the population will remain constant.

Now, as has been alluded to and we will clarify later, solutions are unique once we specify a particular value. So, no other solution can ever have population 1 or 2 at any time, so the graphs of other solutions do not intersect the horizontal lines $y = 1$ and $y = 2$. For an initial population $y(0) > 2$, the derivative y' looks like

$$y' = -(\text{positive})(\text{positive}) < 0,$$

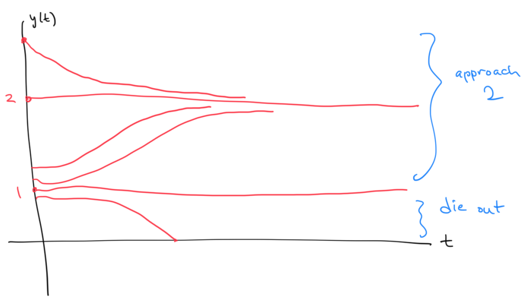
so such solutions decrease towards $y = 2$; for $1 < y(0) < 2$, we have

$$y' = -(\text{positive})(\text{negative}) > 0,$$

so such solutions increase towards $y = 2$; and for $y(0) < 1$, we get

$$y' = -(\text{negative})(\text{negative}) < 0,$$

so such solutions decrease. Hence our populations behave as follows:



The point is that for initial populations larger than 1 (in whatever units), the population overall tends to approach 2 as times goes on—either decreasing or increasing towards 2 depending on whether the initial population is larger than or smaller than 2—while for initial populations smaller than 1 the population will eventually die out. This behavior of *not* being able to grow without restriction is what makes this a more realistic model for population growth; $y = 2$ is the population which the lake environment can actually support.

Lecture 3: 1st Order Linear ODEs

Warm-Up. Uranium and thorium atoms experience *radioactive decay* at a rate proportional to the mass of each present. Moreover, uranium atoms decay specifically into thorium atoms. We want a system of two ODEs which models the mass of samples of uranium and thorium.

Denote the mass of uranium at time t by $x(t)$ and the mass of thorium at time t by $y(t)$. For each we use a compartment model:

$$\text{rate of change} = (\text{rate in}) - (\text{rate out}).$$

First, there is no new uranium being produced, while uranium is decaying (i.e. being lost) at a rate $k_1 x(t)$ for some positive constant k_1. Thus we have

$$x'(t) = 0 - k_1 x(t) = -k_1 x(t).$$

Now, we *do* have new thorium being produced, since the decaying uranium turns into new thorium. Thus the "rate in" for the thorium mass is precisely the rate $k_1 x(t)$ of uranium being lost. But, thorium is lost due to decay at a rate $k_2 y(t)$ for some positive k_2. Thus we get:

$$y'(t) = k_1 x(t) - k_2 y(t).$$

Again, the point is that y' depends on *both* x and y, since "rate in" depends on x and "rate out" on y. Thus, we get the following system of ODEs which model our scenario:

$$x'(t) = -k_1 x(t)$$
$$y'(t) = k_1 x(t) - k_2 y(t).$$

This is what the book calls a *linear cascade*, since the "rate out" of one compartment "cascades" into the "rate in" into the next.

In this case, the solutions can be found by first solving for $x(t)$ based on the first equation, so $x(t) = C_1 e^{-k_1 t}$ for C_1 an arbitrary constant, and then plugging this into the second equation:

$$y'(t) + k_2 y(t) = k_1 C_1 e^{-k_1 t},$$

which can then be solved via the up-until-now "ad hoc" method we've used previously, but which we will now clarify in more detail. Solving more general systems of linear equations is something we'll look at later on, and will use a fair amount of linear algebra.

1st order linear ODEs. We now consider more carefully first-order ODEs of the form

$$y' + p(t)y = q(t).$$

Something which at first appears to be in a more general form such as

$$a(t)y' + b(t)y = c(t)$$

can be brought into the previous form after dividing through by $a(t)$; of course, this only works when $a(t) \neq 0$, which is a restriction we'll address later.

The first thing to address is why we call this a *linear* ODE. The quick-and-dirty answer is that we only have y occurring to the first power, but it is worth making this language more explicit. Be aware that this is not done in the book in a clear way if at all, so we are definitely pushing things a bit further. The key point is that the ODE

$$y' + p(t)y = q(t) \text{ can be written as } L(y) = q(t)$$

where L is the "operation" which takes a function y as input and outputs the function $y' + p(t)q$.

Linear differential operators. The operation L defined above can be viewed as a mapping *from* a space of functions *to* a space of functions:

$$L : \{\text{functions}\} \to \{\text{functions}\}$$

For now we won't be more precise about what set of functions we are looking at, but at the very least they should be functions for which it is actually possible to take a derivative. (If you have a taken an abstract linear algebra course before, we should really be looking at the *vector space* of infinitely-differentiable functions, but we'll avoid use this full-blown "vector space" terminology here since abstract linear algebra is not a prequisite.) So, L sends the function y to the function $L(y) = y' + p(t)q$. Moreover, we say that L is a *differential operator* since it is defined via taking derivatives of the input.

The key property we care about here is that L is *linear*, which means that it satisfies the following two conditions:

- $L(y_1 + y_2) = L(y_1) + L(y_2)$ for any functions y_1 and y_2; and
- $L(cy) = cL(y)$ for any function y and scalar c.

(In the language of an abstract linear algebra course, this says that L is a linear map, or a linear transformation.) Let us verify these conditions for $L(y) = y' + p(t)y$:

$$\begin{aligned}
L(y_1 + y_2) &= (y_1 + y_2)' + p(t)(y_1 + y_2) \\
&= y_1' + y_2' + p(t)y_1 + p(t)y_2 \\
&= (y_1' + p(t)y_1) + (y_2' + p(t)y_2) \\
&= L(y_1) + L(y_2) \\
L(cy) &= (cy)' + p(t)(cy) \\
&= cy' + cp(t)y \\
&= c(y' + p(t)y) \\
&= cL(y).
\end{aligned}$$

The differential operator $D(y) = y' + y^2$, by contrast, is non-linear. Indeed, for functions y_1 and y_2 we have:

$$D(y_1 + y_2) = (y_1 + y_2)' + (y_1 + y_2)^2 = y_1' + y_2' + y_1^2 + 2y_1y_2 + y_2^2 = D(y_1) + D(y_2) + 2y_1y_2$$

so the extra $2y_1y_2$ term shows that D does not satisfy the first required linearity condition. (The second requirement also fails: $D(cy) = (cy)' + (cy)^2 = cy' + c^2y^2 \neq cy' + cy^2 = cD(y)$.) This is caused by the presense of the "non-linear" y^2 term.

Why do we care? The question remains as to why we care about the linearity condition, or in other words what is special about linear ODEs as opposed to non-linear ODEs? The key comes in the realization that solutions of a linear ODE:

$$L(y) = q(t)$$

where L is a linear differential operator, can be described in a particularly nice way. Suppose that y_p is one *particular* solution of $L(y_p) = q(t)$, and let y be any other solution, so that $L(y) = q(t)$ as well. Think of y in the following way:

$$y = y_p + (y - y_p).$$

Observe that $y - y_p$ does not satisfy the given ODE, but it *does* satisfy the analogous non-driven (or homogeneous) equation $L(y_h) = 0$:

$$L(y - y_p) = L(y) - L(y_p) = q(t) - q(t) = 0.$$

(The first step is where linearity is used.) This shows that we can write the solution y of the driven equation $L(y) = q(t)$ as the sum of the particular solution y_p and a homogeneous solution $y_h = y - y_p$ of $L(y_h) = 0$:

$$y = \underbrace{y_p}_{\text{particular}} + (\underbrace{y - y_p}_{\text{non-driven}}).$$

Thus, as long as we are able to find *one* particular solution of $L(y) = q(t)$ and *all* solutions of $L(y) = 0$, we can form *all* possible solutions of $L(y) = q(t)$. That is, arbitrary solutions of $L(y) = q(t)$ can be obtained by "translating" solutions of $L(y) = 0$ by adding y_p to them.

It is this fact which makes linear ODEs simpler to deal with that non-linear ODEs. We saw this already in the linear fish population models, were the general solutions were indeed expressed in this way. (To use some language from linear algebra for those of you who are familiar with it: the solutions of $L(y) = 0$ form a linear subspace of our space of functions, and solutions of $L(y) = q(t)$ form an *affine* subspace of this space of functions, which is obtained by translating the linear subspace.)

Integrating factors. Let us return to our first-order linear ODE:

$$y' + p(t)y = q(t),$$

where we now have a better sense of what "linear" actually means and why it matters. We seek to describe all possible functions $y(t)$ satisfying this equation. We will do so by manipulating this equation in an appropriate way, in a manner analogous to previous examples we looked at.

The claim is that there is a function $M(t)$ we can find with the property that the expression obtained by multiplying the left side of the ODE above by $M(t)$ can be written as the derivative of a single function. More precisely, we want a function $M(t)$ such that

$$y'M(t) + p(t)yM(t) = (M(t)y)'.$$

Such a function is called an *integrating factor* for this ODE, since it will turn our ODE into one which can be solved by computing an antiderivative. The equation we want to hold is:

$$y'M(t) + p(t)yM(t) = (M(t)y)' = M'(t)y + M(t)y',$$

and after comparing both sides we get that $M'(t) = p(t)M(t)$ should hold true. But now we know what $M(t)$ should be: $M(t) = e^{P(t)}$ where $P(t)$ is an antiderivative of $p(t)$, meaning a function satisfying $P'(t) = p(t)$. Indeed, for this function we do have

$$M'(t) = (e^{P(t)})' = e^{P(t)}P'(t) = e^{P(t)}p(t) = p(t)M(t)$$

as desired. Of course there are many possible candidates for $P(t)$, but all we need is one.

Multiplying through by this integrating factor, we see that our original ODE becomes:

$$y'e^{P(t)} + p(t)ye^{P(t)} = e^{P(t)}q(t), \text{ or equivalently } (e^{P(t)}y)' = e^{P(t)}q(t).$$

Thus $e^{P(t)}y$ must be an antiderivative of $e^{P(t)}q(t)$:

$$e^{P(t)}y = \int e^{P(t)}q(t)\,dt + C \text{ where } C \text{ is a constant.}$$

(I'm abusing notation somewhat here in using indefinite integral notation to denote a *particular* antiderivative of $e^{P(t)}q(t)$ as opposed to the general antiderivative, which is why the extra $+C$ term is written explicitly.) Hence, after dividing through by $e^{P(t)}$, we get that all solutions of our original first-order linear ODE $y' + p(t)y = q(t)$ are of the form:

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int e^{P(t)}q(t)\,dt$$

where C is a constant and $P'(t) = p(t)$. Specifying an initial value $y(t_0) = y_0$ will single out a specific constant C. (Going forward, an ODE with an initial value specified will be called an *initial value problem*, or IVP for short.)

Note that this general solution is in the form derived above when discussing why linearity was important: the expression $Ce^{-P(t)}$ with C varying gives all possible solutions of the non-driven equation $y' + p(t)y = 0$ and the $e^{-P(t)} \int e^{P(t)} q(t)\, dt$ term describes one particular solution of the driven equation $y' + p(t)y = q(t)$:

$$y(t) = \underbrace{Ce^{-P(t)}}_{\text{non-driven}} + \underbrace{e^{-P(t)} \int e^{P(t)} q(t)\, dt}_{\text{particular}}.$$

In a sense, the non-driven portion describes the behavior of the solution due to initial conditions alone, while the driven (or "particular") part describes the behavior due to external data.

Back to fish models. The solution to the fish model ODE we looked at previously:

$$y' = ky - H, \text{ or } y' - ky = -H$$

where H is constant can indeed be derived from the general solution above. Here, $p(t) = -k$ and $q(t) = -H$. Thus $P(t)$, an antiderivative of $p(t)$, can be taken to be $P(t) = -kt$. Then

$$e^{P(t)} q(t) = -He^{-kt},$$

and an antiderivative of this can be taken to be

$$\int e^{P(t)} q(t)\, dt = \int -He^{-kt}\, dt = \frac{H}{k} e^{-kt}.$$

Hence the general solution derived above becomes:

$$\begin{aligned} y(t) &= Ce^{-P(t)} + e^{-P(t)} \int e^{P(t)} q(t)\, dt \\ &= Ce^{kt} + e^{kt} \left(\frac{H}{k} e^{-kt} \right) \\ &= Ce^{kt} + \frac{H}{k}, \end{aligned}$$

which is precisely the solution we derived back on the first day. In the example $y' - y = -H(t)$ with periodic harvesting rate $H(t) = 1 + \cos t$, the solution we derived previously comes from the choice of

$$\int e^{P(t)} q(t)\, dt = \int -e^{-t}(1 + \cos t)\, dt = e^{-t} - \frac{1}{2} e^{-t}(\sin t - \cos t)$$

as an antiderivative, which can be found through integration by parts.

Lecture 4: More on 1st Order ODEs

Warm-Up. Suppose sugar is being dissolved in a gallon of water, with the mixture leaking out through a hole at the bottom at the same time. We will take $p(t) = \frac{1}{t}$ gal/min as the rate at which the mixture leaks out at time t, and $q(t) = e^t$ g/min as the rate (g is grams) at the rate at which sugar is added. We will also assume that clean water is added as well so that the volume of the mixture remains 1 gallon throughout. We want to determine the mass $y(t)$ of sugar in the mixture at time $t \geq 1$.

The rate at which this masses changes is given by:

$$y' = e^t - \frac{1}{t}y(t), \text{ or equivalently } y' + \frac{1}{t}y = e^t,$$

where e^t is the rate in and $\frac{1}{t}y(t)$ (the entire mixture leaks out at a rate of $\frac{1}{t}$, and of this $\frac{y(t)}{1}$ is the fraction of which is sugar, so $\frac{1}{t}y(t)$ is the rate at which sugar leaks out) is the rate out. To solve this first-order linear ODE, we take $P(t) = \ln t$ as an antiderivative of $p(t) = \frac{1}{t}$, so that

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int e^{P(t)} e^t \, dt = \frac{C}{t} + \frac{1}{t} \int te^t \, dt = \frac{C}{t} + \frac{1}{t}(t-1)e^t = \frac{C}{t} + \frac{te^t - e^t}{t}$$

is the general solution.

Qualitative behavior. In the general solution for $y' + \frac{1}{t}y = e^t$ derived above, we can see that $y(t) \to \infty$ as $t \to \infty$, so all solutions are unbounded and escape to infinity. In particular, the solution with $C = 0$:

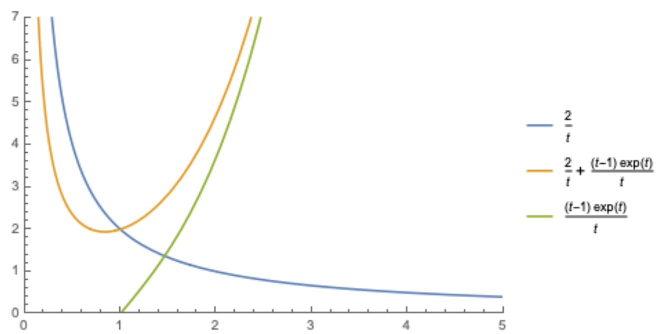
$$y_d = \frac{te^t - t}{t}$$

does this. As mentioned last time, we can interpret this particular solution as describing the effect the external "driving" term e^t has on a solution. The $\frac{C}{t}$ part of the general solution is in fact the general solution of the corresponding non-driven equation:

$$y' + \frac{1}{t}y = 0.$$

The point is that this part $y_h = \frac{C}{t}$ (h stands for "homogeneous") in a sense describes the behavior of the solution due to initial data.

Let us plot all three of these functions: y_h, y_d, and the solution y obtained when $C = 2$:



Initially, the behavior of the solution y in yellow is controlled by the behavior of the non-driven solution y_h in blue, meaning that at this point y is responding mainly to initial data. But eventually the external driven behavior overpowers, and y will respond more to this external data as modeled by y_d in green. In this case the initial response eventually wears off, since y_h approaches 0, but the external response y_d—and hence y—blow up.

Interpretation via responses. This distinction between "response to initial data" vs "response to external data" applies to any first-order linear ODE. Indeed, consider the IVP

$$y' + p(t)y = q(t), \; y(t_0) = y_0.$$

The general solution without regard to the initial condition is

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int e^{P(t)} q(t)\, dt$$

where $P'(t) = p(t)$. We can pick a particular antiderivative of $e^{P(t)}q(t)$ via the following expression:

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int_a^t e^{P(s)} q(s)\, ds$$

where a is some constant and the upper bound t is now variable. Indeed, one part of the Fundamental Theorem of Calculus says precisely that differentiating this integral on the right results in the integrand evaluated at the upper bound:

$$\frac{d}{dt} \int_a^t e^{P(s)} q(s)\, ds = e^{P(t)} q(t),$$

so that this integral expression *is* an antiderivative of $e^{P(t)}q(t)$. (Of course, in nice cases this integral can be computed explicitly so that the final result is not written in terms of an integral, but this is not always achievable.)

The point now is that we can, from this, write down an expression for the specific solution satisfying $y(t_0) = y_0$. First, we make it so that at $t = t_0$ the integral disappears and only the unknown C is left to determine: this can be done by setting t_0 to be the lower bound of the integral:

$$y(t) = Ce^{-P(t)} + e^{-P(t)} \int_{t_0}^t e^{P(s)} q(s)\, ds.$$

Then $y(t_0) = Ce^{-P(t)}$ alone since integrating from t_0 to t_0 results in zero. Hence in order to satisfy $y(t_0) = y_0$, we need

$$Ce^{-P(t_0)} = y_0, \text{ so } C = y_0 e^{-P(t_0)}.$$

Thus the specific solution we want is:

$$y(t) = \underbrace{y_0 e^{-P(t_0)} e^{-P(t)}}_{\text{initial value is } y_0} + \underbrace{e^{-P(t)} \int_a^t e^{P(s)} q(s)\, ds}_{\text{initial value is } 0}.$$

Since the driven part has initial value as 0, we interpret this as saying that this part does not response to the initial input and instead is the part which responds to the external (driven) input, whereas the non-driven (homogeneous) part responds only to the initial input:

$$y(t) = \underbrace{y_0 e^{-P(t_0)} e^{-P(t)}}_{\text{response to initial data}} + \underbrace{e^{-P(t)} \int_a^t e^{P(s)} q(s)\, ds}_{\text{response to external data}}.$$

Thus, solutions of first-order linear ODEs can also be broken down in this way, making clear how solutions respond to the various factors involved.

Particular solutions. As we've seen, the general solution to a first-order linear ODE can be obtained from adding to the general solution of the corresponding non-driven ODE *one* particular solution of the ODE at hand. (This will also be true of higher-order linear ODEs, as well as systems of linear ODEs, as we will see.) One way to obtain such a particular solution is as we did above, by computing an antiderivative of the form

$$\int e^{P(t)} q(t)\, dt.$$

But, this is not always feasible to do easily, and even then it might give the "wrong" particular solution, meaning not the one which exemplifies some behavior we might be looking for.

For instance, consider the ODE

$$y' + 2y = \cos t.$$

Due to the periodic nature of the cosine term, we might guess that perhaps solutions should behave "periodically" in some sense, and maybe even that there is a solution which is periodic on the nose. The point is that using an integrating factor in this case will lead to computing a particular solution of the form

$$e^{-2t} \int e^{2t} \cos t\, dt$$

and if we are not careful about the choices we make, we might not get a periodic solution. Actually, in this case you do, but of course the method of integration by parts is needed.

But if all we are after is *a* solution, perhaps we can find one through more ad-hoc means, say by making a good guess as to what it might look like. In the case at hand, the fact that y' will depend on $\cos t$ suggests that y itself should possible depend on cosine, *and* other functions which result in cosine after differentiating or anti-differentiating. So, let us guess that we can find a solution of the form

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

for unknown constants A and B. The goal is then to find values of A, B which guarantee that this guess is indeed an actual solution. To be a solution, it would have to be true that:

$$y_p' + 2y_p = \cos t, \text{ so } (-A \sin t + B \cos t) + 2(A \cos t + B \sin t) = \cos t.$$

After comparing like terms on both sides, and in particular the coefficients of $\sin t$ and $\cos t$, this results in the requirements that

$$-A + 2B = 0 \text{ and } B + 2A = 1.$$

This pair of equations has solution $A = \frac{2}{5}, B = \frac{1}{5}$, so we find that

$$y_p(t) = \frac{2}{5} \cos t + \frac{1}{5} \sin t$$

is a solution of $y' + 2y = \cos t$. This solution is indeed periodic, and all other solutions are of the form

$$y(t) = Ce^{-2t} + \frac{2}{5} \cos t + \frac{1}{5} \sin t$$

where $y_h = Ce^{-2t}$ describes the general solution of the non-driven ODE $y' + 2y = 0$. In this case, this method of finding a particular solution is likely to be faster than carrying out the required integration above, although it depends on making an initial good guess as to what form the solution will take. Thus we can finally see that all solutions are asymptotically periodic as $t \to \infty$, and specifically approach the particular solution $y_p = \frac{2}{5} \cos t + \frac{1}{5} \sin t$.

Lecture 5: Existence and Uniqueness

Warm-Up. Consider the ODE $y' + \frac{1}{t}y = \cos t$. We find a particular solution of the form

$$y_p(t) = A\sin t + B\cos t + \frac{C\sin t}{t} + \frac{D\cos t}{t}$$

and then describe the asymptotic behavior of all solutions. First, why do we guess for a solution of this form? Simply because we look for functions which tend to give $\cos t$ (the driving term) and $\frac{1}{t}$ (the coefficient of y) when differentiating enough times. So, we should look for a solution with $\cos t$ and $\sin t$, but then also products of these with $\frac{1}{t}$, and so on. It turns out that this is good enough in this example, but if not we could also tack on terms with quadratic denominator, or cubic, etc until we find one that works. (So, very ad-hoc.)

Now, we compute y_p':

$$\begin{aligned} y_p' &= A\cos t - B\sin t + \frac{tC\cos t - C\sin t}{t^2} + \frac{-tD\sin t - D\cos t}{t^2} \\ &= A\cos t - B\sin t + C\frac{\cos t}{t} - C\frac{\sin t}{t^2} - D\frac{\sin t}{t} - D\frac{\cos t}{t^2}. \end{aligned}$$

Thus in order for y_p to be a solution of the given equation, we must have:

$$y_p' + \frac{1}{t}y_p = A\cos t - B\sin t + (C+B)\frac{\cos t}{t} + (A-D)\frac{\sin t}{t} = \cos t.$$

Comparing expressions on both sides, we find that this equality is satisfied when

$$A = 1,\ B = 0,\ C + B = 0,\ A - D = 0$$

so $A = 1 = D$ and $B = 0 = C$. Thus

$$y_p(t) = \sin t + \frac{\cos t}{t}$$

is a particular solution of the given ODE.

To find all solutions then, we simply add to this all solutions of the non-driven equation $y' + \frac{1}{t}y = 0$, which are all of the form $y_h = Ce^{-\ln t} = \frac{C}{t}$. Thus all solutions of the given driven ODE are of the form

$$y = y_h + y_p = \frac{C}{t} + \sin t + \frac{\cos t}{t}.$$

As $t \to \infty$, the first and third terms die-off, so asymptotically all solutions approach $\sin t$.

Solution intervals. Above we saw that all solutions of $y' + \frac{1}{t}y = 0$ are given by $y = \frac{C}{t}$. Now, consider the ODE

$$ty' + y = 0.$$

Dividing by t gives the previous equation, but the point is that technically the solutions of the two are somewhat different. In particular, given the initial condition $y(0) = 0$, the solution to $ty' + y = 0$ is $y = 0$, whereas $y' + \frac{1}{t}y = 0$ is not a well-defined equation at $t = 0$ and the solutions $y = \frac{C}{t}$ we derived above are not defined here. So, dividing by t misses the solution which is defined for $t = 0$. In this case, the solution $y = \frac{C}{t}$ we do get is only defined on the union $(-\infty, 0) \cup (0, \infty)$. Given an initial condition like $y(1) = 3$, where the solution is $y = \frac{3}{t}$, the largest interval around 1 on which the solution exists is only $(0, \infty)$.

In a similar way, consider the IVP

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

The solution to this (which can be derived using the method of "separation", which we'll discuss soon enough) is

$$y = \frac{2}{t(t-2)} = -\frac{1}{t} + \frac{1}{t-2}.$$

In this case, the largest interval around the initial input $t = 1$ on which the solution is defined is $(0, 2)$, since it is undefined at $t = 0$ and $t = 2$.

The upshot is that, in general, a solution to a given IVP might only exist on some interval around the initial input.

Existence and uniqueness. So we now come to the fact that in most (nice) cases, solutions of an IVP are guaranteed to exist and to be unique. This is a crucial property required to derive many of the qualitative facts we'll derive; for instance it guarantees that no two solution curves can intersect, which we have already used a few times.

For first-order linear IVP, this is easy since we have an explicit form of all solutions: all solutions of $y' + p(t)y = q(t)$ look like

$$y = Ce^{-P(t)} + e^{-P(t)} \int_{t_0}^{t} e^{P(s)} q(s)\, ds$$

and the only one which satisfies the initial condition $y(t_0) = y_0$ is the one with $C = y_0 e^{-P(t_0)}$. So, solutions to first-order linear IVPs always exist and are unique, at least under the assumption that $p(t)$ and $q(t)$ are continuous on some interval containing t_0, which is needed in part in order to ensure that the the integral above in the general solution actually exists.

But for more general non-linear IVPs, where explicit solutions are not readily available, how can we guarantee existence and uniqueness of solutions?

Fixed-point formulation. The key to understanding the existence and uniqueness of solutions is the observation that a given initial value problem can be reformulated as what's called a *fixed-point* problem. Consider a first-order IVP of the form

$$y' = f(y,t),\ y(t_0) = y_0$$

where f is a continuous function of two variables. We can anti-differentiate both sides to get that y should then satisfy the following an *integral equation* of the form:

$$y(t) = C + \int_a^t f(y(s), s)\, ds \text{ for some constants } C \text{ and } a.$$

The point is that, by the Fundamental Theorem of Calculus, the integral on the right is indeed an antiderivative of $f(y(t), t)$:

$$\frac{d}{dt} \int_a^t f(y(s), s)\, dx = f(y(t), t).$$

The assumption that $f(y,t)$ is continuous is needed in order to ensure that this integral exists and is differentiable. The values of C and a can be found in order to satisfy the required initial condition

$y(t_0) = y_0$ as follows: take $a = t_0$ so that the value of the integral on the right at $t = t_0$ is exactly 0:

$$y(t_0) = C + \int_{t_0}^{t_0} f(y(s), s)\, ds = C + 0$$

so that $C = y_0$ as a result. Thus, this shows that $y(t)$ satisfies $y'(t) = f(y(t), t), y(t_0) = y_0$ if and only if it satisfies

$$y(t) = y_0 + \int_{t_0}^{t} f(y(s), s)\, ds.$$

So, we have rephrased the original initial value problem as an integral equation instead. Now, define a mapping T which takes as input a function y and outputs the function Ty defined by the right-side above: the value of Ty at t is

$$(Ty)(t) = y_0 + \int_{t_0}^{t} f(y(s), s)\, ds.$$

The requirement that y satisfy $y(t) = y_0 + \int_{t_0}^{t} f(y(s), s)\, ds$ is then precisely the requirement that the result of applying the mapping T to y results in y itself: $Ty = y$. This condition says that y is *fixed-point* of T, and so the conclusion is that a solution of the given initial value problem is the same as a fixed-point of T.

Existence and uniqueness again. Thus, saying that an initial value problem has a unique solution (there are two parts to this: the existence of a solution, and its uniqueness), is equivalent to the statement that the mapping T defined above has a unique fixed point. Did this reformulation gain us anything? Yes, because there are various conditions on $f(y, t)$ which will guarantee that T does have a unique fixed point. Here is one, which is the one given in the book:

> Suppose f and $\frac{\partial f}{\partial y}$ are both continuous on some rectangle containing (t_0, y_0). Then there exists an interval $[t_0 - \delta, t_0 + \delta]$ centered at t_0 on which the initial value problem $y' = f(y, t), y(t_0) = y_0$ has a unique solution.

To be clear, the interval $[t_0 - \delta, t_0 + \delta]$ is one on which the unique solution $y(t)$ is defined.

Example. Consider the following IVP:

$$y' = -\sin y + t \cos t, \;\; y(1) = 2.$$

In this case, $f(y, t) = -\sin y + t \cos t$ is continuous everywhere on $\mathbb{R}^2$, and so certainly some some rectangle containing $(1, 2)$, and so is

$$\frac{\partial f}{\partial y} = -\cos y.$$

Thus the existence and uniqueness theorem guarantees that the IVP above not only has a solution which exists on some interval around 1, but that it only has one. In this case, it is not possible to write down this solution explicitly, but nonetheless it exists.

Towards a proof. For those of you who had some real analysis (or quite a bit actually, to the level of having studied *metric spaces*), here we say a bit about the proof of the existence of a unique fixed point of the mapping:

$$T : y(t) \mapsto y_0 + \int_{t_0}^{t} f(y(s), s)\, ds.$$

The existence of such a fixed point comes from the fact that we can make T into what's called a *contraction* by choosing a small enough interval around t_0, which is the interval alluded to in the statement of the existence and uniqueness theorem. To say that T is a contraction is to say that it "shrinks distances" in the sense that the following inequality holds for some $M < 1$:

$$|(Tg)(t) - (Th)(t)| \leq M|g(s) - h(s)| \text{ for all } t, s \text{ in the domain of } g \text{ and } h.$$

The *Contraction Mapping Principle* (which goes by various other names as well) says that any contraction on a *complete metric space* has a unique fixed point. The complete metric space to which this theorem is applied in our case is the space $C[t_0 - \delta, t_0 + \delta]$ of continuous functions on the to-be-determined interval $[t_0 - \delta, t_0 + \delta]$. This contraction property is thus, ultimately, what underlies our existence and uniqueness theorem.

The difference $(Tf)(t) - (Tg)(t)$ (assuming $t > t_0$) is given by:

$$\left(y_0 + \int_{t_0}^{t} f(g(s), s)\, ds\right) - \left(y_0 + \int_{t_0}^{t} f(h(s), s)\, ds\right) = \int_{t_0}^{t} [f(g(s), s) - f(h(s), s)]\, ds,$$

so that

$$|(Tf)(t) - (Tg)(t)| = \left|\int_{t_0}^{t} [f(g(s), s) - f(h(s), s)]\, ds\right| \leq \int_{t_0}^{t} |f(g(s), s) - f(h(s), s)|\, ds.$$

A version of the *Mean Value Theorem* then gives a bound of the form:

$$|f(g(s), s) - f(h(s), s)| \leq \left(\sup\left|\frac{\partial f}{\partial y}\right|\right) |g(s) - h(s)|$$

where the supremum is taken over points in the to-be-determined closed interval $[t_0 - \delta, t_0 + \delta]$. (The assumption that $\frac{\partial f}{\partial y}$ be continuous is needed to guarantee that this supremum exists as a finite number.) Setting M to be the supremum of $\frac{\partial f}{\partial y}$, we then get:

$$|(Tf)(t) - (Tg)(t)| = \left|\int_{t_0}^{t} [f(g(s), s) - f(h(s), s)]\, ds\right| \leq \int_{t_0}^{t} |f(g(s), s) - f(h(s), s)|\, ds \leq \int_{t_0}^{t} M|g(s) - h(s)|\, ds.$$

If we bound $|g(s) - h(s)|$ by its supremum—which we'll call $\sup|g - h|$—we finally get:

$$|(Tf)(t) - (Tg)(t)| \leq \int_{t_0}^{t} M \sup|g - h|\, ds = M \sup|g - h|(t - t_0).$$

If instead $t < t_0$, we get a similar bound only using $t_0 - t$ at the end, so either way we can phrase the resulting inequality as:

$$|(Tf)(t) - (Tg)(t)| \leq M \sup|g - h||t - t_0|.$$

The point is that by choosing $|t - t_0|$ to be small enough—which comes from the choice of δ in $[t_0 - \delta, t_0 + \delta]$—we can guarantee that T is what's called a *contraction*, which means that it "shrinks" distances in an appropriate way. In particular, for $\delta < \frac{1}{2M}$, we have

$$|(Tf)(t) - (Tg)(t)| \leq M \sup|g - h||t - t_0| \leq M \sup|g - h|\delta < \frac{1}{2} \sup|g - h|,$$

which gives the required "contraction" property. (Any constant $0 < c < 1$ in place of $\frac{1}{2}$ will also work.) Then we can appeal to a fact from analysis that any contraction on a *complete metric* *space* (of which the space of continuous functions on a closed intervals is an example) has a unique fixed point, and this gives our existence and uniqueness theorem. This argument actually works in a slightly more general setting where $\frac{\partial f}{\partial y}$ might not be continuous (what you need is for $f(y,t)$ to satisfy what's called a *Lipshcitz* condition in the first variable, which continuity of the partial derivative definitely implies), and the general version of this existence and uniqueness theorem is called the *Picard-Lindelöf Theorem.*

Lecture 6: Slope Fields

Warm-Up. Consider the IVP $ty' = 2y, y(0) = 0$. Certainly $y = 0$ is a solution, but in fact there are infinitely many solutions, so why does the Existence and Uniqueness Theorem not apply here? The answer is that if we rewrite the given ODE in normal form:

$$y' = \frac{2}{t}y,$$

neither the function $f(y,t) = \frac{2}{t}y$ nor $\frac{\partial f}{\partial y} = \frac{1}{t}$ are continuous on some region containing the initial point $(0,0)$, so there is no guarantee of a unique solution.

How do we know this IVP has infinitely many solutions? We can solve the rewritten first-order equation $y' - \frac{2}{t}y = 0$ using an integrating factor, by picking $P(t) = -2\ln t$ as an antiderivative of $p(t) = -\frac{2}{t}$. We get

$$y(t) = Ce^{2\ln t} = Ct^2$$

for C constant as an infinite family of solutions, and they all satisfy $y(0) = 0$. (Technically, due to the discontinuity at $t = 0$ in $p(t) = -\frac{2}{t}$, we get solutions which are at first only defined for $t \neq 0$, but then we notice that the solutions we do get happen to be defined at $t = 0$ as well.)

But we can get even more solutions than this. For instance,

$$y(t) = \begin{cases} 0 & t \leq 0 \\ t^2 & t > 0 \end{cases}$$

is also a solution (you can check that this function is indeed differentiable at 0, so that it is differentiable everywhere), and

$$y(t) = \begin{cases} t^2 & t \leq 0 \\ 0 & t > 0 \end{cases}$$

is another. By replacing t^2 with more arbitrary Ct^2 terms, we get even more piecewise-defined solutions. Or, we can take things like

$$y(t) = \begin{cases} Ct^2 & t \leq 0 \\ Dt^2 & t > 0 \end{cases}$$

more generally. (These can be thought of as obtained by solving $y' = \frac{2}{t}y$ for $t > 0$, solving it for $t < 0$, and then "splicing" these solutions together at $t = 0$.)

Other uses of uniqueness. Uniqueness of solutions to IVPs can be used to describe the form solutions take as follows. Consider a first-order non-driven linear ODE:

$$y' + p(t)y = 0.$$

Suppose we did not already know how to solve this explicitly, so that we did not know all solutions were of the form $y = Ce^{-P(t)}$ where $P'(t) = p(t)$. We can derive this fact using uniqueness alone.

First, we can verify directly that for $P(t)$ an antiderivative of $p(t)$, $y_d = e^{-P(t)}$ is a solution of $y' + p(t)y = 0$ by plugging in and checking that the equation is satisfied. Our goal is to show that multiples of this one solution give all solutions. Suppose g is an arbitrary solution of this ODE. Fix t_0 and consider the following IVP:

$$y' + p(t)y = 0,\ y(t_0) = g(t_0),$$

where we use the solution g itself to define the initial condition. Then g is certainly a solution to this IVP. But we can verify that

$$y = g(t_0)e^{P(t_0)}e^{-P(t)}$$

is *also* a solution of this IVP, and so by uniqueness of solutions to IVPs we see that g and this y must be the same function, so that

$$g(t) = g(t_0)e^{P(t_0)}e^{-P(t)}.$$

Hence any solution of $y' + p(t)y = 0$ is indeed a multiple of $y_d = e^{-P(t)}$ as claimed.

The point is that we can verify that $y_d = e^{-P(t)}$ generates all solutions to the given ODE without relying on the fact that we can solve this equation explicitly, by instead by relying on uniqueness. This idea is not so important in this case since, as we know, we can indeed solve any first-order linear ODE explicitly, but it will more important when we later consider second- or higher-order equations.

Approximating solutions. In situations, say arbitrary nonlinear ODEs, where it is not possible to find a solution explicitly, we then rely on methods for approximating solutions. This is not something we'll delve into too deeply in this course, since we can use a computer to do this work for us, but we will at least give a sense as to where such techniques come from.

First, we can use the idea we described last time behind the proof of the existence and uniqueness theorem. Recall that the point is that we can rephrase a given IVP as a fixed-point problem:

$$y' = f(y,t),\ y(t_0) = y_0 \iff y(t) = y_0 + \int_{t_0}^{t} f(y(s),s)\,ds \iff y = Ty$$

where $T : \{\text{functions}\} \to \{\text{functions}\}$ is the mapping defined by

$$(Ty)(t) = y_0 + \int_{t_0}^{t} f(y(s),s)\,ds.$$

The fixed-point which gives the solution to the IVP arises as a limit of a certain sequence of functions: take any function f_0 to start with, and consider the functions obtained by repeatedly applying T:

$$g_0,\ g_1 = Tf_0,\ g_2 = T^2 f_0,\ g_3 = T^3 f_0,\ \ldots.$$

To be clear, the $(n+1)$-st function in this list is obtained from the n-th as follows:

$$g_{n+1}(t) = y_0 + \int_{t_0}^{t} f(g_n(s),s)\,ds.$$

It turns out (this is the part which requires a study of metric spaces in real analysis) that the functions $g_0, g_1, g_2, \ldots$ *converge* as $n \to \infty$ to the function y which is the unique solution of the given IVP. Thus, the sequence of functions g_n provide better and better approximations to this solution: g_1 is one approximation to y, g_2 is a better approximation, and so on. Hence by using an appropriate g_n we can approximate the solution y we want. This method of approximating is called *Picard iteration*.

Alternatively, we can use tangent-line approximations to approximate the solution we want as follows. The key is that the ODE $y' = f(y,t)$ tells us what the slope of the graph of a solution should be at any point. Consider an initial condition $y(t_0) = y_0$. At the point (t_0, y_0), the slope of the solution satisfying this IVP should be

$$y'(t_0) = f(y_0, t_0).$$

Hence the tangent line to this solution looks like

$$y = y_0 + f(y_0, t_0)(t - t_0).$$

Near t_0, this tangent line should provide a good approximation to the desired solution, so that if we move along the tangent line a bit we will not stray too far from the solution we want. So, let us move along this tangent line a small amount to a point (t_1, y_1), and consider now the tangent line to the graph of the solution satisfying the new initial condition: $y(t_1) = y_1$:

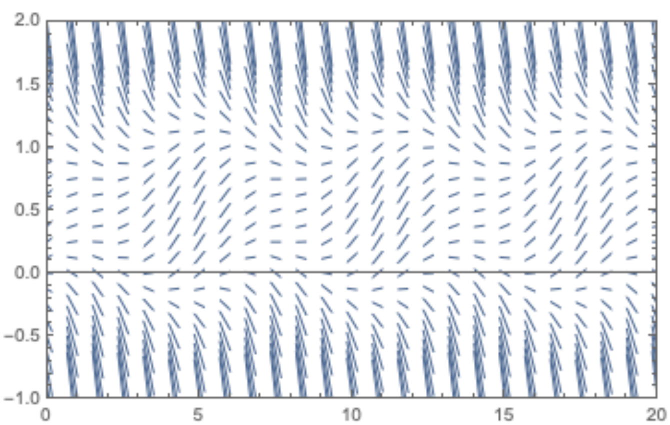
This tangent line looks like

$$y = y_1 + f(y_1, t_1)(t - t_1),$$

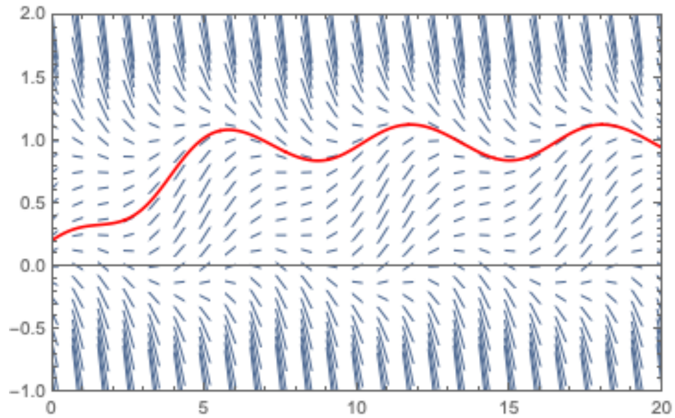
and again near t_1 this should provide a good approximation to the solution satisfying $y(t_1) = y_1$. Now we move along this new tangent line a bit to a new point (t_2, y_2) and consider a new tangent line approximation, and so on and os on. The point is that the tangent line segments we end up forming in this way are providing some approximation to the solution of the initial IVP $y(t_0) = y_0$ we considered:

If the steps we took from t_0 to t_1 to t_2 etc were too large, this would not be a good approximation, but if we take very small steps along the way we actually won't be too far off. This technique is known as *Euler's Method*, and is also widely used.

Slope fields. The fact that an ODE gives us the slope of the tangent line to an actual solution at any point is useful when it comes to deriving information about solutions in situations where an explicit solution is not readily available. Consider a plot where at each point (t, y) in the ty-plane we draw a little line segment with slope equal to the value of $f(y, t)$:



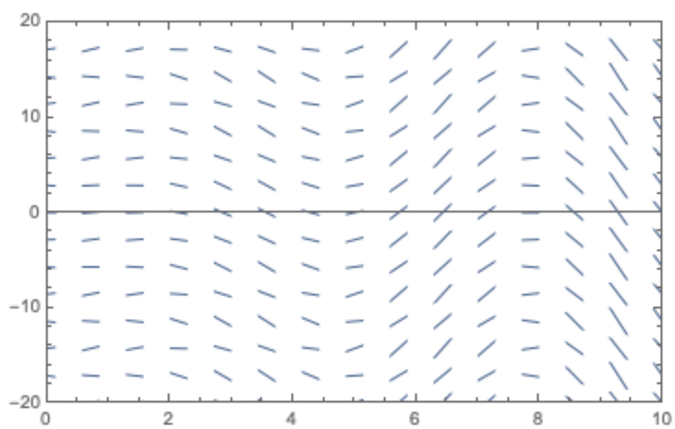
Such a plot is called the *slope field* of the ODE $y' = f(y, t)$, since it visually describes the slopes of possible solutions. The graph of any solution of $y' = f(y, t)$ must then be tangent to these segments at each point through which it passes:



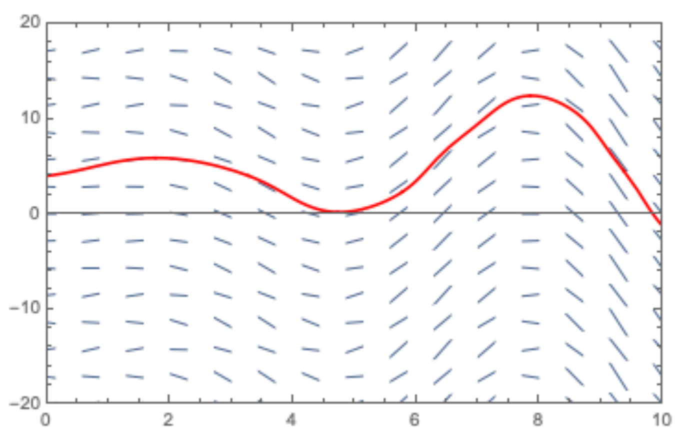
The point is that, even with only the slope field available, much information about solutions can be derived, and indeed, we can make reasonable drawings of their graphs by following the tangents.

Example 1. Consider the ODE $y' = y$. ***TO BE FINISHED***

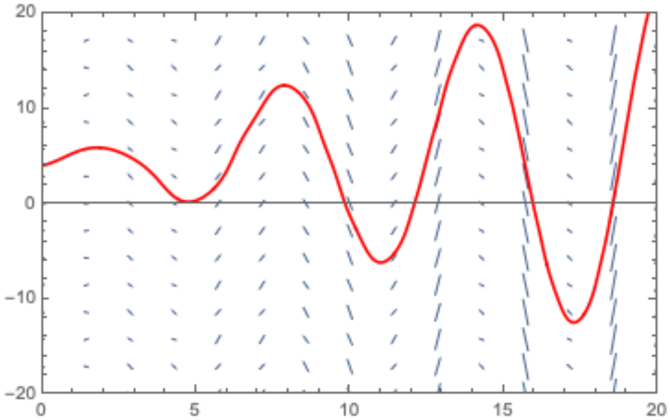
Example 2. The slope field of $y' = -\sin y + t\cos t$ looks like:



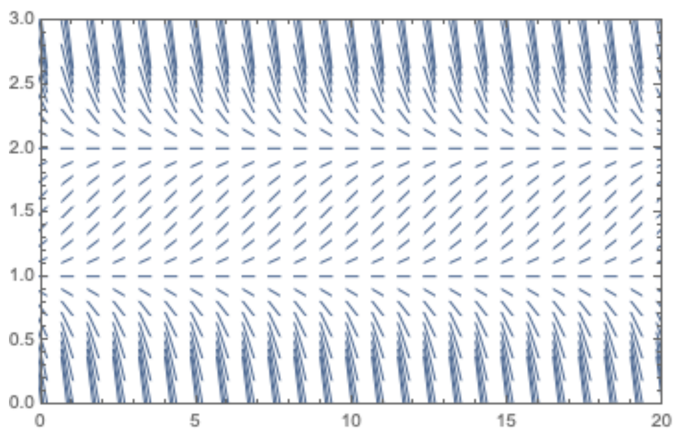
This already indicates some possible sinusoidal behavior among solutions. Indeed, the solution satisfying $y(0) = 4$ is:



and the sinusoidal behavior becomes even more apparent if we zoom out:



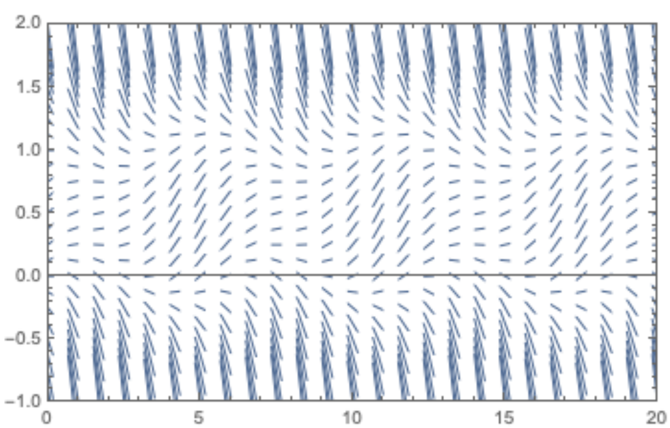
Example 3. The slope field $y' = 3y - y^2 - 2$ looks like:



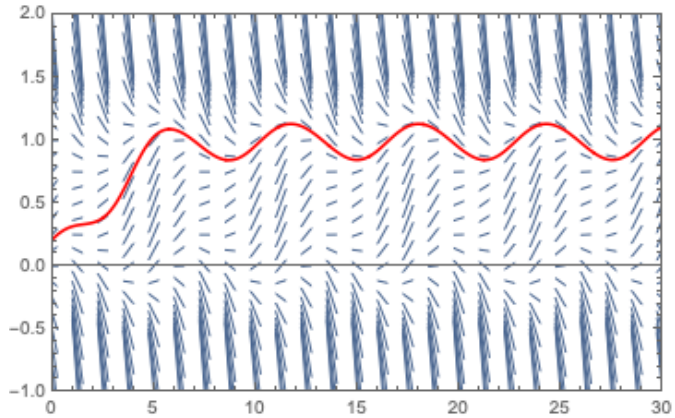
From this we can easily identify the constant solutions $y = 1$ and $y = 2$, as well as the qualitative behavior of all other solutions: those with $y(0) > 2$ appear to shrink down towards $y = 2$; those with $1 < y(0) < 2$ appears to increase towards $y = 2$; and those with $y(0) < 1$ appear to decay rapidly.

Lecture 7: Separable Equations

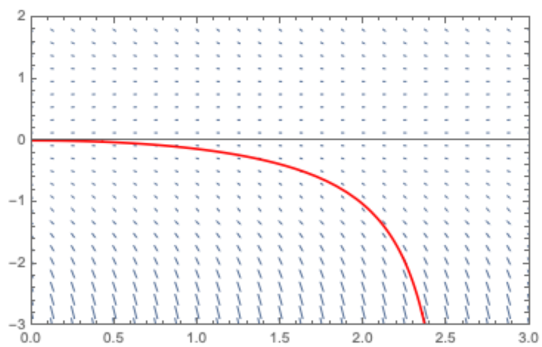
Warm-Up. Consider the ODE $y' = y - y^2 - 0.2 \sin t$. Its slope field looks like:



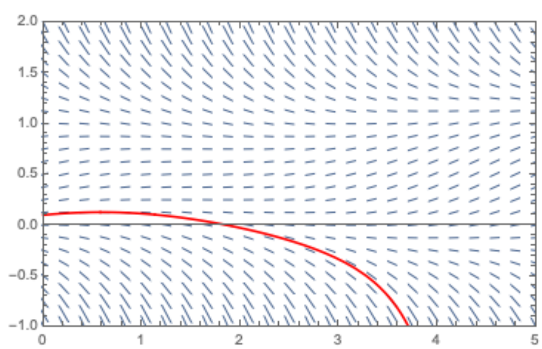
and we want to determine the behavior of some solutions from this alone. There are some tangents which appear to move along something like a sine or cosine curve, so it seems that some solutions will be asymptotically periodic:



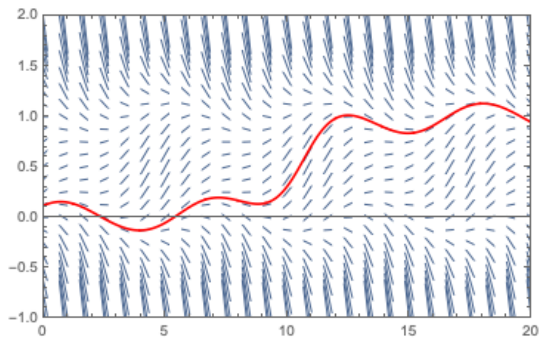
But this is not always the case: note for, say, the initial condition $y(0) = 0$, following the tangents seems to give a solution which will die off to $-\infty$:



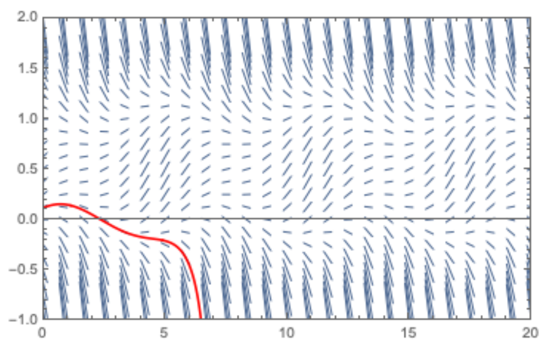
(This picture is zoomed in more as opposed to the previous ones, but the ODE is the same.) This is also the case for some small initial values of $y(0)$:



For the solutions which do remain bounded, for some it appears we get ones which perhaps dip below the initial value $y(0)$ at some point, only to be pulled back up:



(This is the solution satisfying $y(0) = 0.117$.) Yet for $y(0)$ just a bit smaller than this, once the solution dips below the value of $y(0)$, the downward pull is too strong so that the it dies off:



(This is the solution satisfying $y(0) = 0.116$.)

The point then is that there appears to be a point at which solutions change from being bounded to rapidly decreasing to $-\infty$, or more precisely a value y_0 so that for $y(0) > y_0$ our solution remains bounded, but for $y(0) < y_0$ it dies off. A question one can ask is whether we can determine this value of y_0 in a more concrete way, from only the original ODE and without having the explicit solution available? We'll come back to this later. (It you start plotting these solutions with the aid of a computer, the value of y_0 will lie between 0.116 and 0.117, as the two drawings above show.)

Separable equations. And now we return to equations which can be solved explicitly. In certain cases, it might be possible to write a given ODE in the form:

$$N(y)y' + M(t) = 0,$$

where we can group all terms involving y together and all terms involving t together. Such an equation is said to be *separable* and can be solved explicitly, at least to the extent that explicit antiderivatives for $N(y)$ with respect to y and $M(t)$ with respect to the t. Given an explicit antiderivative $G(y)$ of $N(y)$ with respect to y, the equation above becomes

$$(G(y))' + M(t) = 0$$

and anti-differentiating gives

$$G(y) + \int M(t)\,dt = 0,$$

which *implicitly* defines y as a function of t. Let us see this in action.

Example 1. Consider the ODE $y' = y$. Of course we know that all solutions are of the form $y(t) = Ce^t$, which we can derive using an integrating factor, but let us see how to derive this by viewing this as a separable equation. First, note that $y = 0$ is a constant solution, and is the only solution which can attain the value 0 anywhere by uniqueness of solutions. Thus any other solution y will be nowhere zero, so for such solutions we can "separate the variables" by dividing through by y:

$$\frac{y'}{y} = 1.$$

So, all y-terms are grouped together and all t-terms (viewing the constant 1 as a such a term in this case) are grouped together. The left side is the derivative of $\ln|y|$:

$$(\ln|y|)' = 1,$$

so anti-differentiating gives

$$\ln|y| = t + C \text{ for } D \text{ a constant.}$$

This implicitly defines y as a function of t, and in this case we can go further to explicitly obtain y as a function of t.

Exponentiating both sides gives

$$|y| = e^t e^C,$$

so $y = \pm e^D e^t$. Setting $C = \pm e^D$, we get that the nonzero solutions are of the form

$$y = Ce^t \text{ where } C \text{ is a nonzero constant.}$$

Note it was important to use $\ln|y|$ as the antiderivative of $\frac{y'}{y}$ instead of simply $\ln y$ in order to guarantee that we cover solutions with C being negative; we would have missed these with $\ln y$ alone. Finally, the expression we obtained $y = Ce^t$ *does* also make sense when $C = 0$, in which case we get the constant solution we already pointed out above. So, $y(t) = Ce^t$ for C an arbitrary constant does give all solutions. (To be clear, the separation of variables method yields solutions $y = Ce^t$ with $C \neq 0$, but it just so happens that the missing solution of $y = 0$ can also be written in this form once we allow $C = 0$ as well.)

Example 2. Consider the IVP

$$2tyy' = 1 + y^2, \; y(2) = 3.$$

Since we are looking for a solution defined at $t = 2$, we will consider only t which are close to this, meaning that we will assume $t > 0$. Since $1 + y^2 > 0$, we can separate variables to rewrite this equation as

$$\frac{2yy'}{1 + y^2} = \frac{1}{t}.$$

The left-side is the derivative of $\ln(1 + y^2)$, so anti-differentiating gives:

$$\ln(1 + y^2) = \ln|t| + C$$

for C constant. (Note $1 + y^2$ is positive, so there is no need to use $|1 + y^2|$ instead.) We can exponentiate both sides to get:

$$1 + y^2 = e^{\ln|t|}e^C = e^C|t|,$$

and thus $y^2 = D|t| - 1$ where $D := e^C > 0$. This implicit equation gives two possible explicit values for y:

$$y(t) = \sqrt{D|t| - 1} \text{ or } y(t) = -\sqrt{D|t| - 1}.$$

Given the initial condition $y(2) = 3$ we want satisfied, with $y(2)$ positive, we need the first form, in which case we can find D via:

$$3 = y(2) = \sqrt{2D-1} \implies 9 = 2D - 1, \text{ so } D = 5.$$

Thus the solution of the given IVP is $y(t) = \sqrt{5|t|-1} = \sqrt{5t-1}$, which is defined for $t \geq \frac{1}{5}$.

Example 3. We have previously looked at the equation

$$y' = 3y - y^2 - 2,$$

which gave a nonlinear model for fish populations with harvesting. We have seen that $y = 1$ and $y = 2$ are equilibrium constant solutions, and we have determined the asymptotic behavior of other solutions, based on whether the initial population $y(0)$ is larger than 2, between 1 and 2, or smaller than 1. Let us now find all solutions explicitly, and verify this behavior.

The equilibrium solutions are the only ones which can attain the value 1 or 2, by uniqueness. Thus we will assume that y is never 1 nor 2. In this case the given equation is separable:

$$\frac{y'}{3y - y^2 - 2} = 1.$$

To find the antiderivative of the left-side, we use the method of partial fractions. That is, since $3y - y^2 - 2 = -(y-1)(y-2)$, we look for constants A, B such that

$$\frac{1}{3y - y^{-}2} = \frac{A}{y-1} + \frac{B}{y-2}.$$

After clearing denominators, this becomes:

$$1 = -A(y-2) - B(y-1).$$

Setting $y = 1$ shows that $A = 1$, and setting $y = 2$ shows that $B = -1$. Thus

$$\frac{y'}{3y - y^2 - 2} = \frac{y'}{y-1} - \frac{y'}{y-2},$$

so $\ln|y-1| - \ln|y-2|$ is an antiderivative of the left-side of

$$\frac{y'}{3y - y^2 - 2} = 1.$$

Hence anti-differentiating gives

$$\ln|y-1| - \ln|y-2| = t + C, \text{ or } \ln\left|\frac{y-1}{y-2}\right| = t + C.$$

Exponentiating gives

$$\left|\frac{y-1}{y-2}\right| = e^C e^t,$$

so

$$\frac{y-1}{y-2} = \pm e^C e^t = De^t \text{ for } D \neq 0.$$

Solving for y explicitly thus gives

$$y(t) = \frac{2De^t - 1}{De^t - 1}$$

as the non-equilibrium solutions. Actually, setting $D = 0$ gives $y = 1$, so the constant solution $y = 1$ can actually be incorporated into the form above. Thus all solutions of $y' = 3y - y^2 - 2$ are

$$y = 2 \text{ or } y = \frac{2De^t - 1}{De^t - 1} \text{ for } D \text{ constant.}$$

Now, we verify that the non-constant solutions—those of the second form above for $D \neq 0$—do have the qualitative behavior we previously described. Using the fact that

$$y(0) = \frac{2D - 1}{D - 1},$$

we can work that $y(0) > 2$ if and only if $D > 1$, $1 < y(0) < 2$ if and only if $D < 0$, and $y(0) < 1$ if and only if $0 < D < 1$. (For $D = 1$ the solution does not exist at $t = 0$, and so this is not among the initial conditions we are considering.) When $1 < y(0) < 2$ or $y(0) > 2$, the denominator $De^t - 1$ in our solution is never zero, since $D < 0$ in the former case and $D > 1$ in the latter. Thus these solutions do not experience any singularity for $t > 0$ and

$$\lim_{x \to \infty} \frac{2De^t - 1}{De^t - 1} = \lim_{x \to \infty} \frac{2D - e^{-t}}{D - e^{-t}} = \frac{2D}{D} = 2$$

as expected. However, for $y(0) < 1$, so for $0 < D < 1$, there exists a value of t_1 such that $De^{t_1} - 1 = 0$, so as t increases beyond $t = 0$ these solutions will decay down to $-\infty$ since

$$\lim_{t \to t_1^-} \frac{2De^t - 1}{De^t - 1} = -\infty.$$

Hence these solutions, as expected, do not approach a limiting value and decay in finite time.

Lecture 8: Predator-Prey Model

Warm-Up. Consider the ODE

$$-yy' - x = 0,$$

where y is now a function of x, so that y' denotes $\frac{dy}{dx}$. We can solve this using separation as follows:

$$-yy' = x, \text{ so } -\frac{1}{2}y^2 = \frac{1}{2}x^2 + C \text{ for a constant } C$$

after anti-differentiating. This gives $x^2 + y^2 = D$ (for D a constant) as an implicit equation for y in terms of x, and we can then find y as an explicit function of terms of x, choosing a positive or negative square root as needed depending on a given initial value.

Now, this implicit equation describes a circle in the xy-plane, and from previous courses we know that the values for (x, y) along this circle can be parametrized using equations of the form

$$x(t) = R\cos t, \ y(t) = R\sin t$$

for a constant R. So, the question is, if we think about x and y *both* as functions of some additional parameter t, is there a way to rephrase the given ODE

$$-yy' - x = 0$$

in terms of $x(t)$ and $y(t)$ instead, meaning in terms of their derivatives *with respect to* t? To not confuse the various derivatives we will consider, let us now reserve the prime notation y', x' for derivatives with respect to t, and write $\frac{dy}{dx}$ explicitly for the derivative of y with respect to x, so that our original ODE is

$$-y\frac{dy}{dx} - x = 0, \text{ or } \frac{dy}{dx} = -\frac{x}{y}.$$

But we *do* know how to express $\frac{dy}{dx}$ in terms of $y'(t)$ and $x'(t)$: the chain rule gives

$$y'(t) = \frac{dy}{dx}x'(t), \text{ so } \frac{dy}{dx} = \frac{x'(t)}{y'(t)}.$$

Thus in order to have $\frac{dy}{dx} = -\frac{x}{y}$, we can take for instance

$$y' = x \text{ and } x' = -y.$$

Hence, the non-linear ODE $-yy' - x = 0$ can be turned into a pair of linear ODEs:

$$-yy' - x = 0 \iff \begin{cases} x' & = -y \\ y' & = x \end{cases}$$

where we parametrize both x and y. Note that the parametrizations $x(t) = R\cos t$ and $y(t) = R\sin t$ we gave above do indeed satisfy this system $x' = -y, y' = x$, so that from solutions of this system we can recover solutions of the original ODE. We will discuss systems of linear ODEs, and their solutions, in more detail later on.

Phase portraits. Given a system of ODEs for two functions x and y, we can plot the relation between x and y in the xy-plane. The resulting picture is called the *phase portrait* of the system, and each individual point (x, y) in it describes on possible state the system can be in. For instance, for the system above:

$$\begin{aligned} y' &= x \\ x' &= y \end{aligned}$$

the relation between x and y is given by $x^2+y^2 = D$ (the implicit equation derived in the Warm-Up), so that the phrase portrait consists of circles of varying radii:

To be clear, these circles do not describe the *graphs* of x and y as functions of t, but rather the relationship between x and y themselves. From this we can see, for instance, that the solutions x and y are both periodic with respect to t, since horizontally x moves back forth between $-R$ and R, as does y vertically. Of course, we have explicit solutions $x(t) = R\cos t, y(t) = R\sin t$ available and can thus tell that they are periodic, but the point is that this is something we can determine from the phase portrait alone even if explicit solutions were not available.

Lotka-Volterra. We seek now to model two interacting populations, one of a predator species and one of a prey species. (Of course, using "predator" and "prey" is not meant to indicate that this only applies to literal predator and prey animal species. Any more abstract relationship—such as that between two companies or industries—in which one entity can be characterized as a "predator" and the other as "prey" would work just as well. So, "population" could also refer to something else, like market share.) Denote the predator and prey populations at time t by $x(t)$ and $y(t)$ respectively. As a first step towards modeling these populations, we can imagine that were there no interaction between the two, the predator population should *decrease* (no prey present to be eaten) while the prey population should *increase* (no predators present to eat), so a first-order approximation to these populations might be given by equations of the form:

$$x' = -ax \quad \text{and} \quad y' = cy$$

where a, c are positive constants. (So, we assume that with no interactions, the rate at which populations change is proportional to the population, and even if there were births to consider in the predator case or deaths in the prey case, deaths outweigh births for predators and births outweigh deaths for prey so that the coefficients a, c are indeed positive.)

But now we incorporate interactions between the two species. Specially, we assume the following:

> the rate of change in populations due to interactions at a given time is proportional to the product of the two populations at that time.

Thus, x' and y' should each be proportional to xy. In particular, x should respond positively to interactions, so the proportionality constant in this case should be positive, while y should respond negatively, so the proportionality constant should be negative:

$$x' = bxy, \; y' = -kxy \quad \text{where } b, k > 0.$$

Hence all together we get the following system, known as the *Lotka-Volterra* model:

$$\begin{aligned} x' &= -ax + bxy \\ y' &= cy - kxy \end{aligned}$$

with $a, b, c, k > 0$. Again, the linear terms represent intrinsic change not due to interactions, while the non-linear terms represent the interactions.

First observations. First we note that these equations have a pair of constant solutions: $(x, y) = (0, 0)$ and $(x, y) = (\frac{c}{k}, \frac{a}{b})$. Certainly $x = 0, y = 0$ satisfy the given equations. If $x \neq 0$, in order to get a constant solution the first equation (after dividing through by x) requires that

$$0 = -a + by, \text{ so } y = \frac{a}{b}.$$

Similarly, if $y \neq 0$, the x coordinate of a constant solution must satisfy

$$0 = c - kx, \text{ so } x = \frac{c}{k}$$

from the second equation. Thus $(\frac{c}{k}, \frac{a}{b})$ is the only nonzero equilibrium solution: populations with these initial values will remain constant for all time.

Second, note that $x = Ce^{-at}, y = 0$ and $x = 0, y = De^{ct}$, where C, D are constants, are solutions. In the phase portrait—the plot of x vs y—these solutions take up the x- and y-axes respectively. This implies, by uniqueness of solutions (which also applies to such a *system* of ODEs, as we will elaborate on later) that no other solutions can cross these axes, so in particular a solution which initially beings with positive populations in the first quadrant will remain in the first quadrant for all time.

Finally, consider a solution in the first quadrant $(x(t), y(t))$ which at some point intersects itself: $(x(t_0), y(t_0)) = (x(t_1), y(t_1))$ for some t_0, t_1. The first point can be viewed as specifying an initial condition, so the uniqueness of solutions the behavior of $(x(t), y(t))$ after t_0 will be the same as the behavior after t_1, since this initial condition is also satisfied as this latter point. This says that, visually, $(x(t), y(t))$ forms a *closed* orbit, meaning that once it intersects itself it just retraces the exact same curve. These correspond to *periodic* populations: the populations of predator and prey might not remain constant, but eventually they return back to where they were before and the behavior repeats from this point on.

Volterra's first law. The observation about periodic solutions above is actually indicative of what happens to *all* solutions in the first quadrant. That is, any solution with $(x(t_0), y(t_0))$ in the first quadrant will form a closed orbit, a result which is known as *Volterra's first law*. The proof of this comes from turning the Lotka-Volterra equations into a single separable equation.

If we think of y as being a function of x (the *Implicit Function Theorem* from real analysis implies that at least one of y or x can be viewed locally as being a function of the other), we have that

$$\frac{dy}{dx} = \frac{y'(t)}{x'(t)} = \frac{cy - kxy}{-ax + bxy}.$$

But this is a separable equation for y and x:

$$-ax\frac{dy}{dx} + bxy\frac{dy}{dx} = cy - kxy \implies \left(b - \frac{a}{y}\right)\frac{dy}{dx} = \frac{c}{x} - k.$$

Anti-differentiating gives:

$$by - a\ln|y| = c\ln|x| - kx + D$$

where D is constant. This equation actually defines a closed curve in the xy-plane, which says that a non-constant solution indeed defines a closed (periodic) orbit. We won't verify that this equation defines a closed curve here, but if interested you can see some of the work done in the book, both in this section and in one of the exercises.

Volterra's second law. Next we are interested in understanding *average* populations of predator and prey. Consider a solution $(x(t), y(t))$ to the Lotka-Volterra equations of period T with $x, y > 0$. The average population of the predator species over one cycle is given by the value of

$$\frac{1}{T}\int_0^T x(t)\,dt.$$

But the point is that we can compute this even if we do not know $x(t)$ explicitly, by using the fact that $(x(t), y(t))$ satisfies in particular the second Lotka-Volterra equation: $y' = cy - kxy$. This gives $x = \frac{c}{k} - \frac{1}{k}\frac{y'}{y}$, so

$$\frac{1}{T}\int_0^T x(t)\,dt = \frac{1}{T}\int_0^T \frac{c}{k}\,dt - \frac{1}{Tk}\int_0^T \frac{y'}{y}\,dt = \frac{c}{k} - \frac{1}{Tk}(\ln y(T) - \ln y(0)) = \frac{c}{k}$$

where we use periodicity to see that $y(T) = y(0)$, so $\ln y(T) = \ln y(0)$. The upshot is that, no matter the initial populations, the average population of predators along any one cycle is always the same value of $\frac{c}{k}$.

Similarly, for the prey species the first Lotka-Volterra equation gives $y = \frac{a}{b} + \frac{1}{b}\frac{x'}{x}$, so the average population of prey over one cycle is:

$$\frac{1}{T}\int_0^T y(t)\,dt = \frac{1}{T}\int_0^T \frac{a}{b}\,dt + \frac{1}{bT}\int_0^T \frac{x'}{x}\,dt = \frac{a}{b} + \frac{1}{bT}(\ln x(T) - \ln x(0)) = \frac{a}{b},$$

where periodicity is used to get $x(T) = x(0)$. Thus the average population of prey is always $\frac{a}{b}$ over any cycle, regardless of the initial populations.

Of course, $x = \frac{c}{k}$ and $y = \frac{a}{b}$ are precisely the equilibrium solutions, which makes sense: if the average populations are going to be the same over any cycle, then they should all be equal to the ones obtained along the equilibrium cycle (which consists of a single point), which are $\frac{c}{k}$ and $\frac{a}{b}$ since the populations along the equilibrium cycle are constant. The nice thing is that we were able to derive this from the structure of the Lotka-Volterra equations themselves, without the use of an explicit form of the solutions.

Lecture 9: Long-Term Behavior

Warm-Up. Consider the Lotka-Volterra predator-prey model with *harvesting*, where we assume the rate of population (of either predator or prey) lost due to harvesting is proportional to the population itself. This gives the equations:

$$\begin{aligned} x' &= -ax + bxy - H_1 x \\ y' &= cy - kxy - H_2 y \end{aligned}$$

for some positive constants H_1, H_2. The goal is to understand the effect of harvesting on the predator and prey populations.

TO BE FINISHED

Extending solutions. As we've seen, solutions to an IVP may only exist on some interval around the initial input. A natural question to ask then is how large such an interval might be. That is, if we know a solution exists on a given interval, can we *extend* that solution to a larger interval?

Suppose $f(y,t)$ and $\frac{\partial f}{\partial y}$ are continuous on some rectangle, and take a point (t_0, y_0) in that rectangle. We know that the IVP

$$y' = f(y,t),\ y(t_0) = y_0$$

has a unique solution on some interval $[t_0 - \delta, t_0 + \delta]$. But now consider an initial condition *at* the point $t_1 = t_0 + \delta$, so we get the IVP

$$y' = f(y,t),\ y(t_1) = \text{whatever this value is.}$$

Then *this* also has a solution which exists on some interval $[t_1 - \delta_1, t_1 + \delta_1]$. This interval overlaps a bit with the previous interval, so by uniqueness the solution to this first IVP and this second IVP must be the same on this overlap, so that we can splice both solutions together to now get a solution defined on the larger interval $[t_0 - \delta, t_1 + \delta_1]$:

We can now take the new right endpoint of this interval to get a new IVP which has a unique solution, extending our given solution beyond $t_1+\delta_1$, and so on and so on as long as we remain within the given rectangle. We can do the same with left endpoints, extending our solution "backwards" in time:

We conclude that we can extend our solution all the way to the edges of the rectangle on which $f(y,t)$ and $\frac{\partial f}{\partial y}$ are continuous. For instance, a while back we considered the IVP

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

TO BE FINISHED

Solutions which escape to ∞ or stop. We have also seen that solutions can escape to infinity in finite time or simply stop.***TO BE FINISHED***

Bounding solutions. Another long-term behavior we have seen in some examples is that of being *bounded*. One question to ask is whether we can determine that a solution will be bounded, and if so, whether the bound can be extracted from the ODE alone. We will give one answer to this, in the simple case of a first-order linear IVP. Other arguments are possible for various other IVPs, and in particular there was an non-linear example of this in one of the discussion problems.

Consider a first-order linear ODE:

$$y' + p(t)y = q(t)$$

where $p(t), q(t)$ are continuous. Suppose that $p(t)$ is positive and bounded away from zero, so that there exists p_0 such that $P(t) \geq p_0 > 0$ for all $t \geq 0$. Suppose also that $q(t)$ is bounded, so that there exists $M > 0$ such that $|q(t)| \leq M$ for $t \geq 0$. Then any solution $y(t)$ of this ODE is bounded as follows:

$$|y(t)| \leq y(0) + \frac{M}{p_0}.$$

This result is usually called *BIBO* for short, which stands for "bounded input, bounded output". That is, if we can provide appropriate bounds on the "input" data $p(t), q(t)$ then we get a bound on the "output" solution $y(t)$.

TO BE FINISHED

Lecture 10: Autonomous Equations

Warm-Up. We give examples showing that neither hypothesis of BIBO (the bounded-input, bounded-output theorem) can be dropped.

$$y' + y = t, \ t \geq 0$$

$$y' + \frac{1}{(t+1)^2} y = e^{\frac{1}{t+1}}, \ t \geq 0$$

TO BE FINISHED

Autonomous equations. We now focus more closely on certain types of equations: *autonomous* ODEs. These are ODEs of the form

$$y' = f(y)$$

with no explicit dependence on t. (Of course, these still implicitly depend on t, since $y = y(t)$ is still a function of t.) We've seen a few equations of this type before, and they have various nice properties which are worth fleshing out.

In particular, we can speak of *equilibrium*, or constant, solutions. These are precisely the zeroes of the function $f(y)$, since for a constant solution we have $y' = 0$, and so a constant function satisfies $y' = f(y)$ if and only if $f(y) = 0$. These are called "equilibrium" solutions for two reasons: first, a solution which begins with an initial value given by an equilibrium must remain that same value throughout; and second, all other solutions must either blow-up to ∞, die-off to $-\infty$, or approach (when the solution is bounded) one of these equilibrium solutions. The intervals over which $f(y)$ is positive or negative describes what happens to other solutions in relation to these equilibria, as we will see.

Example 1. Here is one example we have actually looked at before:

$$y' = 3y - y^2 - 2 = -(y-1)(y-2).$$

The equilibria solutions are $y = 1$ and $y = 2$. For $y > 2$, $f(y) = -y(y-1)(y-2)$ is negative so solutions which initially start above $y = 2$ decrease towards $y = 2$; for $1 < y < 2$, $f(y)$ is positive so solutions in this range increase towards $y = 2$; and for $y < 1$, $f(y)$ is negative so solutions here die-off to $-\infty$. We'll keep track of this increase/decrease behavior with arrows, and to make this more easily noticeable we'll do this in a separately-drawn line:

The line on the left is called the *state line* of this autonomous ODE. Since solutions move "towards" $y = 2$, we call the equilibrium $y = 2$ an *attractor*, while since solutions move "away from" $y = 1$, we call $y = 1$ a *repeller*. Note that this is fully encoded by the state line, with arrows point towards attractors and away from repellers.

Example 2. Consider the following autonomous ODE:

$$y' = (y - 1)(y - 2)(y - 3).$$

The equilibrium solutions are $y = 1$, $y = 2$, and $y = 3$. For $y > 3$, $f(y) = (y - 1)(y - 2)(y - 3)$ is positive; for $2 < y < 3$, $f(y)$ is negative; for $1 < y < 2$ it is positive, and for $1 < y$ it is negative. Hence the state line and various solutions look like:

These ODE has two repellers ($y = 3$ and $y = 1$) and one attractor ($y = 2$).

Example 3. Something new happens with the equilibria of

$$y' = (y-1)^2(y-2).$$

These has two equilibria, $y = 1$ and $y = 2$, and while $y = 2$ is a repeller, $y = 1$ is neither an attractor nor a repeller. Indeed, for $1 < y < 2$, $f(y) = (y-1)^2(y-2)$ is negative, but for $y < 1$ it is still negative!

(This happens due to the $(y-1)^2$, which is always nonnegative regardless of y.) The equilibrium $y = 1$ attracts on one side (above) but repels on the other (below), so we call it (for lack of a better word) an *attractor/repeller*.

Long-term behavior. ***TO BE FINISHED***

Logistic equation. A particularly important autonomous ODE is one we've referred to before: the *logistic equation*. These is the ODE which characterizes a population with linear growth (births), quadratic decay (deaths), and some possible constant harvesting/restocking:

$$y' = ay - by^2 + H.$$

It is more common to rewrite this in the following way, after setting $a = r$ and $b = \frac{r}{K}$:

$$y' = ry - \frac{r}{K}y^2 + H = ry(1 - \tfrac{y}{K}) + H.$$

Here we take r, K to be positive, and H to be positive or negative depending on whether we restock or harvest respectively. The reason for writing the logistic equation this way is that K can be given some meaning, as follows.

The logistic equation without harvesting (nor restocking) is $y' = ry(1 - \frac{y}{K})$ and has equilibrium solutions $y = 0$ and $y = K$. For $y > K$, $y' = f(y) = ry(1 - \frac{y}{K})$ is negative since $yK > 1$, and for $0 < y < K$, y' is positive since $\frac{y}{K} < 1$. For $y < 0$, y' is positive again, but these solutions have no physical meaning since population cannot be negative. The point is that $y = K$ is an attractor, and so K has the interpretation as the *carrying capacity* of the population, which is the optimal population the background environment can support. Any initial positive population will in the long-run approach this carrying capacity population. The r coefficient in the logistic equation measures the "intrinsic" growth of the population, not due to the environment.

Explicit solutions. The explicit non-equilibria solution of the non-harvested logistic equation

$$y' = ry(1 - \tfrac{y}{K})$$

can be found using separation and partial fractions. We will leave the details of this to elsewhere (the book or discussion problems), but they end up being of the form:

$$y(t) = \frac{KCe^{rt}}{1 + Ce^{rt}}$$

for C a constant. From this one can determine that for $y(0) > 0$, $\lim_{t\to\infty} y(t) = \frac{KC}{C} = K$, as expected. For the logistic equation with harvesting, a general solution is tougher to write down, but can be done without too much difficulty in concrete examples, like $y' = 3y - y^2 - 2$, which we previously worked out.

The behavior of equilibria. Consider the following logistic equation, with variable harvesting parameter c:

$$y' = 3y - y^2 - c.$$

We seek to determine the behavior of the equilibria as c varies. The equilibrium solutions are found by solving

$$3y - y^2 - c = 0, \text{ so they are } y = \frac{-3 \pm \sqrt{9 - 4c}}{-2}.$$

Now, when $9 - 4c < 0$, so for $c > \frac{9}{4} = 2.25$, there are no real equilibria; for $c = \frac{9}{4}$, there is one equilibrium; and for $c < \frac{9}{4}$ there are two equilibria. Thus as c varies, we go from having no equilibria to having one to having two, or from having two, to one, to none depending on whether we view c as decreasing to the left or increasing to the right. Since the nature of the equilibria changes at $c = \frac{9}{4}$, we call this value a *bifurcation point*: we have one equilibria "bifurcating" into two, at least if we view c as decreasing from right to left. We'll study bifurcations more carefully next time.

Lecture 11: Bifurcations

Warm-Up. We find and classify the equilibrium solutions of $y' = (e^y - 2)\sin y$. ***TO BE FINISHED***

Logistic (saddle-node) bifurcations. We return now the general logistic equation, with harvesting parameter:

$$y' = ry(1 - \tfrac{y}{K}) + H.$$

(Technically, when $H > 0$ we should actually interpret it as "restocking" instead of "harvesting", but we'll refer to it as the "harvesting parameter" either way.) Our goal is to understand the behavior of equilibrium solutions as the parameter H varies.

Concretely, equilibrium solutions arise when $ry(1 - \frac{y}{K}) + H = 0$, and so are explicitly given by

$$y = \frac{-r \pm \sqrt{r^2 + \frac{4rH}{K}}}{-\frac{2r}{K}}.$$

When $H < -\frac{rK}{4}$, there are no real equilibria since $r^2 + \frac{4rH}{K}$ is negative; for $H = -\frac{rK}{4}$ there is only one equilibrium solution (here $r^2 + \frac{4rH}{K} = 0$), which is $y = \frac{K}{2}$; and for $H > -\frac{rK}{4}$ there are two equilibria since $r^2 + \frac{4rH}{K} > 0$ has two square roots. Thus $H = -\frac{rK}{4}$ is the *bifurcation point* of the logistic equation, meaning it is the value of the parameter H where the number and nature of equilibria changes. (If you had used $-H$ as the parameter in the logistic equation, the bifurcation would occur at $H = \frac{rK}{4}$ instead.)

For $H < -\frac{rK}{4}$, where there are no equilibria, the value of $ry(1 - \frac{y}{K}) + H$ is negative, ***TO BE FINISHED***

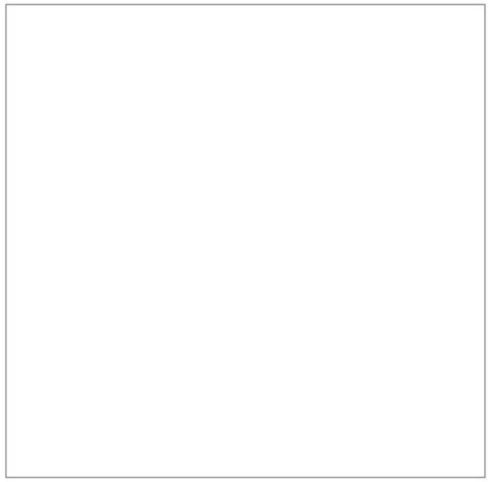
When $H = -\frac{rK}{4}$, the unique equilibrium $y = \frac{K}{2}$ is an attractor/repeller. Indeed, for an initial value $y(0) > \frac{K}{2}$, say for instance $y(0) = K$, the value of $y' = ry(1 - \frac{y}{K}) - \frac{rK}{4}$ is negative, while for $y(0) < \frac{K}{2}$, say $y(0) = 0$, y' is still negative. Hence $y = \frac{K}{2}$ attracts solutions above and repels solutions below in this case:

The point is that whether or not a population is sustainable with harvesting $H = -\frac{rK}{4}$ depends heavily on the initial population and there is no room for error: if $y(0)$ is above $\frac{K}{2}$, the population will not die-off and will approach the equilibrium $\frac{K}{2}$ in the long-term, while if $y(0)$ is even ever-so-slightly below $\frac{K}{2}$, the population will die-off. So, this is better than having no equilibria, although still not ideal since there is a question as to whether or not we should indeed consider populations which start off larger than the equilibrium (should we allow initial sizes beyond the carrying capacity?), even if they will tend towards the equilibrium in the long-run.

With harvesting parameter $H > -\frac{rK}{4}$ larger than that at the bifurcation, we have a more relaxed range of initial populations which lead to sustainable outcomes. We get two equilibria in these cases:

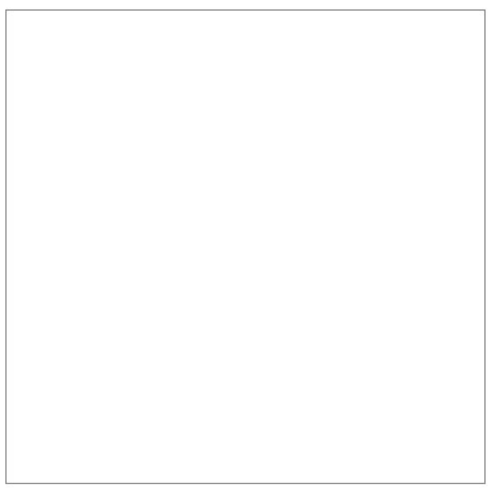
$$y = \frac{r - \sqrt{r^2 + \frac{4rH}{k}}}{\frac{2r}{K}} \quad \text{and} \quad y = \frac{r + \sqrt{r^2 + \frac{4rH}{k}}}{\frac{2r}{K}}$$

where the vertical distance between them grows as H moves further beyond $-\frac{rK}{4}$:



This growing distance indicates that we have a greater allowable range of initial values $y(0)$ which lead to sustainable populations, which are all still always below the larger limiting equilibrium. So, we do not have to worry about initial sizes which are too large beyond the carrying capacity: which can we start off below this carrying capacity and still maintain sustainability as long as we remain above the lower equilibrium. Harvesting parameters which are larger than $-\frac{rK}{4}$ (so little harvesting, or much restocking) seem to give the best controlled behavior overall.

We can summarize this discussion in a *bifurcation diagram*, which plots the relation between the equilibria and the parameter H as follows:



Pitchfork bifurcations. Consider now the following autonomous ODE:

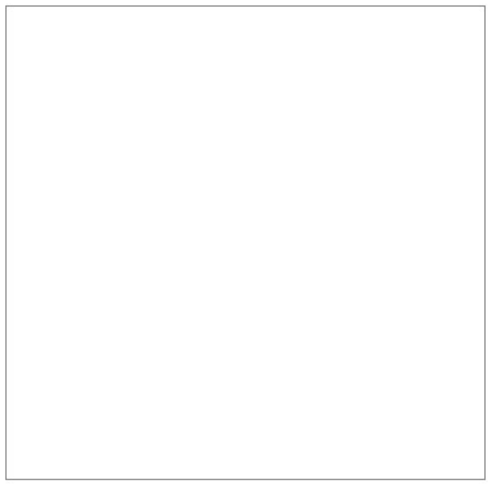
$$y' = -(c + y^2)y.$$

The equilibrium solutions are $y = 0$ and $y = \pm\sqrt{-c}$. So, here there is always at least one equilibrium, and then two more when $c < 0$. The bifurcation thus occurs at $c = 0$. For $c \geq 0$, the only equilibrium is $y = 0$, and for $y(0) > 0$ we have $y' < 0$, while for $y(0) < 0$ we have $y' > 0$. Thus $y = 0$ attracts solutions in this case.

For $c < 0$, we have the following analysis:

$$\begin{aligned} y(0) < -\sqrt{-c} &\implies y' > 0 \\ -\sqrt{-c} < y(0) < 0 &\implies y' < 0 \\ 0 < y(0) < \sqrt{-c} &\implies y' > 0 \\ \sqrt{-c} < y(0) &\implies y' < 0. \end{aligned}$$

Thus $y = -\sqrt{-c}$ attracts, $y = 0$ repels, and $y = \sqrt{-c}$ attracts, so the bifurcation diagram looks like:



Such a bifurcation is called a *pitchfork* bifurcation.

Transcritical bifurcations. For a final example, consider

$$y' = cy - y^2 = y(c - y).$$

There is always at least one equilibrium, $y = 0$, and a second at $y = c$ when $c \neq 0$. When $c = 0$, $y = 0$ is an attractor/repeller since $y' < 0$ for either $y(0) < 0$ or $y(0) > 0$. For $c < 0$, we have the following:

$$\begin{aligned} y(0) < c &\implies y' < 0 \\ c < y(0) < 0 &\implies y' > 0 \\ 0 < y(0) &\implies y' < 0 \end{aligned}$$

so in this case $y = c$ repels and $y = 0$ attracts. However, for $c > 0$ we have:

$$\begin{aligned} y(0) < 0 &\implies y' < 0 \\ 0 < y(0) < c &\implies y' > 0 \\ c < y(0) &\implies y' < 0 \end{aligned}$$

so here $y = 0$ repels and $y = c$ attracts. The point is that the two equilibria switch behaviors (attracting vs repelling) at the bifurcation $c = 0$:

This is known as a *transcritical* bifurcation.

Lecture 12: Springs & 2nd Order ODEs

Warm-Up.

The motion of a spring. We now move on to studying 2nd order equations, in particular (eventually) ones of the form

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

where f is a function of three variables encoding some relation between y, y', and t. For instance:

$$y'' = \sin y + 3y' + ty + t^2$$

is an example. As before, the goal is to understand functions y which satisfy such an equation.

For now, we will focus on *linear* second-order equations, and in particular ones with constant coefficients. A common interpretation of such equations comes from modeling the motion of a spring. Suppose a spring is anchored to a certain spot and that a mass is attached to the other end. If we pull the mass in one direction and let go, the spring will pull at the mass and cause it to undergo some motion:

Denote by $y(t)$ the position of the spring at time t, which we measure as its displacement away from its resting position $y = 0$—we want an equation characterizing $y''(t)$, which is the acceleration of the moving spring. In physics, the force exerted by a moving body is equal to its mass times its acceleration: $F) = my''$. *Hooke's Law* for springs says that the restorative force which wants to pull the spring back after it has been extended is proportional to the displacement from its resting position:

$$my'' = -ky$$

where k is known as the *spring constant* and depends on the spring in question. (So this term gives rise to an "internal" or "intrinsic" motion.) The negative sign comes from the fact that the restorative force works in the direction opposite the displacement: if the mass is pulled to the left, the force pulls it to the right, while if the mass is pulled to the right, the force pulls it left.

But there may be other factors to consider as well. For instance, if the spring is moving through the air, the mass will experience air resistance, which causes a *damping* effect. The molecules making up the mass collide with air molecules, which causes them to lose energy as time goes on, and causes the spring to slow down. The force caused by such a damping effect is proportional to the velocity of the moving mass, so the faster the mass moves the greater the effect:

$$my'' = -dy'$$

for some *damping constant* d. Again, the negative is due to the direction: if the spring moves right, the drag caused by air resistance is pushing to the left, and vice-versa. Putting the motion due to Hooke's Law and that due to damping together gives:

$$my'' = -dy' - ky.$$

Moreover, there could possibly be some driving term to consider, say caused by some external phenomena which also guides the motion of the spring:

$$my'' = -dy' - ky + r(t).$$

Rewriting this ODE by moving all terms involving y to the left and dividing by m puts it into a more standard form:

$$y'' + \frac{d}{m}y' + \frac{k}{m}y = \frac{1}{m}r(t).$$

So, this is an example of a second-order linear driven ODE, in this case with "constant coefficients".

Example.

Orbits.

2nd-order linear ODEs. In general, we consider second-order linear ODEs of the form

$$y'' + p(t)y' + q(t)y = r(t)$$

where $p(t), q(t), r(t)$ are continuous. Even if not modeling the motion of a spring, we'll refer to $p(t)$ as a "damping" term, and $q(t)$ as some "internal/intrinsic" term. As we did in the first-order case, we can be much more explicit about what we mean by saying that this is a "linear" equation. Define the "differential operator" $L : \{\text{functions}\} \to \{\text{functions}\}$ by setting

$$L(y) = y'' + p(t)y' + q(t)y,$$

so that the ODE we considering can be written as $L(y) = r(t)$. To say that L is *linear* means that it behaves in the following way with respect to sums and scalar multiples:

- $L(y_1 + y_2) = L(y_1) + L(y_2)$ for any functions y_1, y_2, and
- $L(cy) = cL(y)$ for any function y and scalar c.

This can be verified by explicitly computing both sides in each to see that equality holds. The fact that the operation of computing a first- or second-order derivative itself has such linearity properties is the key point.

And just as in the first-order case, the linearity of L gives us the following fact, which is the crucial point to understanding the structure of solutions to such ODEs: any solution of $L(y) = r(t)$ can be written as $y_h + y_d$ where y_h is a solution of the non-driven (or *homogeneous*) equation $L(y) = 0$ and y_d is a particular solution of the driven equation $L(y) = r(t)$ under consideration. The proof of this fact is exactly the same as it was the first-order case, so you can go back and check the appropriate previous lecture to see the details. This says that in order to find all solutions of $L(y) = r(t)$, we should focus on finding *one* solution, possibly via some ad-hoc means, and on finding *all* solutions of the analogous non-driven equation. We'll carry out this plan for various cases in the coming days.

Lecture 13: 2nd Order Solutions

Warm-Up.

Existence & uniqueness. As we saw before, in order to obtain a unique solution to a second-order IVP, we need to specify both an initial value $y(t_0) = y_0$ for a solution *and* an initial value $y'(t_0) = v_0$ for its derivative at the same t_0; specifying $y(t_0) = y_0$ alone is not enough, essentially because a second-order ODE allows for two "degrees of freedom", and setting $y(t_0) = y_0$ only cuts that down by one. (In other words, knowing the second derivative of a function only characterizes that function up to two arbitrary coefficients.)

But now, how can we guarantee that even with $y(t_0)$ and $y'(t_0)$ both specified, a solution exists *and* that it is unique? The answer comes from using what we know about existence and uniqueness for first-order equations (or a slight generalization thereof), and the fact that a second-order equation can always be turned into a pair of first-order equations. Consider a second-order linear ODE of the form:

$$y'' + p(t)y' + q(t)y = r(t).$$

Introduce a new function v by setting $v = y'$, so that v keeps track of "velocity". Then $v' = y''$, so we can turn the second-order ODE above into a first-order ODE for v:

$$v' = y'' = -p(t)y' - q(t)y + r(t) = -p(t)v - q(t)y + r(t).$$

Including the ODE $y' = v$ which came from the definition of v gives the following system of first-order equations:

$$\begin{aligned} y' &= v \\ v' &= -p(t)v - q(t)y + r(t). \end{aligned}$$

Of course, we can go backwards: given this system of equations, we can recover the original second-order equation for y by replacing v with y'. Moreover, a pair of initial conditions $y(t_0) = y_0, y'(t_0) =$

v_0 for y and its derivative gives initial conditions $y(t_0) = y_0, v(t_0) = v_0$ for y and v. So, we conclude that any second-order linear IVP is equivalent to a first-order system of IVPs:

$$\begin{cases} y'' + p(t)y' + q(t)y = r(t) \\ y(t_0) = y_0, \ y'(t_0) = v_0 \end{cases} \qquad \text{is equivalent to} \qquad \begin{cases} y' = v \\ v' = -p(t)v - q(t)y + r(t) \\ y(t_0) = y_0, \ v(t_0) = v_0. \end{cases}$$

How does this help us formulate an existence and uniqueness theorem for second-order IVPs? By exploiting the fact that the previous first-order existence and uniqueness theorem we looked at works just as well (modified appropriately) for a first-order IVP for *vector-valued* functions. Consider a vector-valued function $\mathbf{y}(t)$ with component functions $y(t)$ and $v(t)$:

$$\mathbf{y}(t) = \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}.$$

The first-order system above then becomes a single first-order ODE for this vector function, which we can write using matrices as:

$$\begin{bmatrix} y'(t) \\ v'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q(t) & -p(t) \end{bmatrix} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ r(t) \end{bmatrix},$$

or more succinctly as:

$$\mathbf{y}' = A\mathbf{y} + \mathbf{r}(t)$$

where A is the 2×2 matrix above and $\mathbf{r}(t) = \left[\begin{smallmatrix} 0 \\ r(t) \end{smallmatrix}\right]$ is the vector-valued driving term. The linearity of this second-order ODE is reflected in the fact that the non-driving term $A\mathbf{y}$ is given by multiplication by a matrix. More generally, one might consider a vector-valued first-order ODE of the form

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, t)$$

where $\mathbf{f}$ some other vector-valued function. The first-order equations $y' = f(y, t)$ we considered previously are simply the special case where $\mathbf{y}$ and $\mathbf{f}$ are have "vector values" in $\mathbb{R}^1$, which means they are actually just scalar-valued. The initial conditions $y(t_0) = y_0, v(t_0) = v_0$ can be turned into a vector-valued initial condition: $\mathbf{y}(t_0) = \mathbf{y}_0$ where $\mathbf{y}_0 = \begin{bmatrix} y_0 \\ v_0 \end{bmatrix}$.

The same type of argument we gave for the existence and uniqueness of first-order IVPs—rephrasing as a fixed-point problem and applying some general fixed-point theorem—also applies to such vector-valued first-order IVPs:

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, t), \ \mathbf{y}(t_0) = \mathbf{y}_0,$$

so we get an existence and uniqueness theorem for these equations, and in turn then an existence and uniqueness theorem for our second-order IVPs. To be clear, the claim is that if $p(t), q(t)$, and $r(t)$ are continuous, then for any t_0, y_0, and v_0 there exists a solution of the IVP

$$y'' + p(t)y' + q(t)y = r(t), \ y(t_0) = y_0, \ y'(t_0) = v_0$$

on an interval around t_0, and that solution is unique.

Structure of solutions. So, solutions to the types of equations we are considering exist, and the problem now is to find them. As stated last time, the linearity of these equations implies that all solutions are of the form

$$y_h + y_d$$

where y_h is a solution of the corresponding non-driven equation and y_d is a particular solution of the equation at hand. If our equation looks like $L(y) = r(t)$, where L is a second-order linear differential operator, we thus look for one solution of $L(y) = r(t)$ and all solutions of $L(y) = 0$. We will leave finding particular solutions to later, and focus now on understanding solutions of $L(y) = 0$.

The key fact is the following: if y_1, y_2 are two linearly independent solutions of $L(y) = 0$, then any solution of $L(y) = 0$ is of the form $c_1y_1 + c_2y_2$ for some scalars c_1, c_2. The notion of "linear independence" is the same as that seen in linear algebra, only now applied to functions and not just vectors in $\mathbb{R}^n$. For the case of two functions, this just means that neither function is a multiple of the other. We will soon see how we can characterize this notion of independence in terms of a function known as the *Wronskian*, which is useful as well for determining independence of more than two functions. The expression $c_1y_1 + c_2y_2$ is called a *linear combination* of y_1 and y_2 (again borrowing a term from linear algebra), and we say that the set of solutions of $L(y) = 0$ is *spanned* by y_1 and y_2. Since this space of solutions is spanned by two independent solutions, we also say that it is "2-dimensional". (If you have taken a more abstract course in linear algebra, the set of solutions of $L(y) = 0$ is a *vector space* of dimension 2, with y_1, y_2 forming a *basis*.) This is a higher-dimensional version of what we saw for first-order linear non-driven ODEs: the solutions of $y' + p(t)y = 0$ are all multiples of $e^{-P(t)}$, where $P(t)$ is an antiderivative of $p(t)$, so the function $e^{-P(t)}$ alone spans the space of all solution, and hence this space is 1-dimensional.

Thus, the problem of finding all solutions of $L(y) = 0$ comes down to finding two independent solutions. We will see that, at least for constant coefficient equations, this can always be done explicitly.

Linear independence and the Wronskian. Let us now clarify a bit more what is meant by saying that two functions y_1, y_2 are linearly independent. The quick definition is that neither y_1 nor y_2 is a scalar multiple of the other. Another way of saying this, which is perhaps a more common phrasing, is that the only scalars c_1, c_2 satisfying $c_1y_1 + c_2y_2 = 0$ are $c_1 = 0 = c_2$. (If there were nonzero scalars satisfying this equation, it would be possible to isolate one function on one side to express as a multiple of the other.)

So, suppose instead that y_1 and y_2 are dependent, so that there exists a scalars c_1, c_2, not both zero, such that $c_1y_1 + c_2y_2 = 0$. This equation and the one obtained after differentiating both sides gives the following pair of equations:

$$\begin{aligned} c_1y_1(t) + c_2y_2(t) &= 0 \\ c_1y_1'(t) + c_2y_2'(t) &= 0. \end{aligned}$$

But this can be written in matrix form as:

$$\begin{bmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The key point is that here $\left[\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right]$ is a nonzero vector, and the only way such a matrix equation can have a nonzero solution is for the 2×2 matrix on the left to not be invertible, meaning that it has determinant zero. Thus, if y_1, y_2 are dependent, the determinant

$$\det \begin{bmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{bmatrix} = y_1(t)y_2'(t) - y_1'(t)y_2(t)$$

must be zero for all t. This determinant is called the *Wronskian* of y_1 and y_2, and is denoted by $W(y_1, y_2)$:

$$W(y_1, y_2) = y_1y_2' - y_1'y_2.$$

Thus, dependence of y_1, y_2 implies $W(y_1, y_2) = 0$, and so by taking the contrapositive we find that if the Wronskian is nonzero at at least one point, then y_1, y_2 must be independent. This gives a clean way of checking independence of solutions of a given second-order ODE. Now, it is important to clarify precisely what we are saying: if there exists t_0 such that $W(y_1, y_2)(t_0) \neq 0$, then y_1, y_2 are independent, but in general knowing that $W(y_1, y_2)$ is always zero does NOT guarantee that y_1 and y_2 are independent. That is, for functions in general, it is possible to be linearly independent and yet have Wronskian which is identically zero. However, as we will see next time, for functions which a priori satisfy a second-order linear non-driven ODE, this cannot happen: in this case, independence is equivalent to having a nonzero Wronskian. This will come from the fact that the Wronskian itself satisfies as first-order ODE, as we'll see.

Why 2-dimensional? It remains to justify the fact that if $y_1(t), y_2(t)$ are two linearly independent solutions of $L(y) = 0$, they do indeed span the space of all solutions. Suppose $g(t)$ is an arbitrary solution of $L(y) = 0$. From this we can formulate the following IVP:

$$L(y) = 0,\ y(t_0) = g(t_0),\ y'(t_0) = g'(t_0).$$

But now, we can find scalars c_1, c_2 such that

$$c_1 y_1(t_0) + c_2 y_2(t_0) = g(t_0) \quad \text{and} \quad c_1 y_1'(t_0) + c_2 y_2'(t_0) = g'(t_0).$$

Indeed, this pair of equations can be expressed in matrix form as

$$\begin{bmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} g(t_0) \\ g'(t_0) \end{bmatrix},$$

and since the matrix on the left is invertible (its determinant is the Wronskian of the linearly independent solutions y_1, y_2, so it is nonzero), there exists a solution for $\left[\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right]$.

Thus $c_1 y_1(t) + c_2 y_2(t)$ is a solution of $L(y) = 0$ which satisfies the IVP stated above with initial values derived from g, so since g itself also satisfies this IVP, we have

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

by uniqueness. Hence every solution of $L(y) = 0$ is a linear combination of y_1, y_2 as claimed.

Finding solutions. The search is now on for linearly independent solutions of a second-order linear non-driven ODE with constant coefficients:

$$y'' + ay' + by = 0.$$

We first look for solutions of the form $y = e^{rt}$. The idea to consider such functions comes from knowing that, based on the given ODE, y should somehow be related to a multiple of its derivative and a multiple of its second derivative, and exponentials are simple functions which do have this property. Plugging in, we see that $y = e^{rt}$ satisfies $y'' + ay' + by = 0$ if and only if

$$r^2 e^{rt} + are^{rt} + be^{rt} = 0.$$

Since e^{rt} is nonzero, this holds if and only if $r^2 + ar + b = 0$, so we find that the solutions of this quadratic equation describe the exponential solutions of the given ODE.

The equation $r^2 + ar + b = 0$ is called the *characteristic equation* of $y'' + ay' + by = 0$. If this quadratic has distinct real roots r_1, r_2, then

$$y = e^{r_1 t} \quad \text{and} \quad y = e^{r_2 t}$$

are two solutions of $y'' + ay' + by = 0$. The Wronskian of these two is:

$$W(y_1, y_2) = \det \begin{bmatrix} e^{r_1 t} & e^{r_2 t} \\ r_1 e^{r_1 t} & r_2 e^{r_2 t} \end{bmatrix} = (r_2 - r_1) e^{r_1 t} e^{r_2 t},$$

which is nonzero since $r_1 \neq r_2$. Hence the solutions above are linearly independent, so

$$y(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

describes all solutions of $y'' + ay' + by = 0$ when the characteristic equation has distinct real roots.

Example. Consider the ODE

$$y'' - y' - 2y = 0.$$

The characteristic equation is $r^2 - r - 2 = 0$, which has roots $r = -1, 2$. Thus $y = e^{-t}$ and $y = e^{2t}$ are linearly independent solutions, so

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

gives all solutions of $y'' - y' - 2y = 0$.

Lecture 14: Constant Coefficients

Warm-Up. We verify that for solutions y_1, y_2 of a second-order linear non-driven ODE:

$$y'' + p(t)y' + q(t)y = 0,$$

the Wronskian $W(y_1, y_2) = y_1 y_2' - y_1' y_2$ is either always zero or never zero. The derivative of the Wronskian is:

$$W' = y_1' y_2' + y_1 y_2'' - y_1'' y_2 - y_1' y_2' = y_1 y_2'' - y_1'' y_2.$$

Since y_1, y_2 each satisfies the second-order ODE above, we have:

$$y_1'' = -p(t)y_1' - q(t)y_1 \quad \text{and} \quad y_2'' = -p(t)y_2' - q(t)y_2.$$

Thus:

$$W' = y_1[-p(t)y_2' - q(t)y_2] - [-p(t)y_1' - q(t)y_1]y_2 = -p(t)[y_1 y_2' - y_1' y_2] = -p(t)W.$$

Hence the Wronskian of y_1, y_2 satisfies a first-order linear non-driven ODE, so it looks like $W = ce^{-\int p(t)\,dt}$ and is thus always zero or never zero depending on the value of c.

Phase portraits. Recall that we can visualize solutions of a second-order ODE in two ways: either via a plot of the solution itself versus t, or via a plot of its *orbits* instead, which are the curves traced out by y and y' in the yy'-plane. Using a plot of orbits has the advantage that solution curves will not intersect, since only one function can satisfy a specified set of initial values for y and y' simultaneously. (This is in contrast with plots of y vs t, where specifying an initial value for y alone does not uniquely characterize a solution, so solution curves here might intersect. We only get uniqueness if we further specify an initial *tangent* direction.) The plot of orbits is called the *phase portrait* (or also *orbital portrait*) of the ODE. The name comes from physics, where the data of y and y' together (for instance, position and velocity) characterizes what's called the *phase* of the system in question.

Consider the ODE $y'' - y' - 2y = 0$ from last time, whose general solution was

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

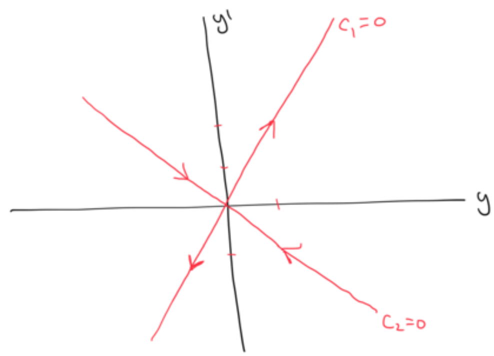
Differentiating gives

$$y' = -c_1 e^{-t} + 2c_2 e^{2t}.$$

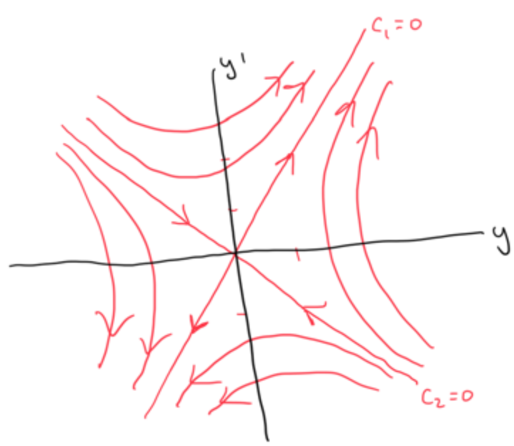
Thus the orbits are the curves with the following parametrized form:

$$\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} c_1 e^{-t} + c_2 e^{2t} \\ -c_1 e^{-t} + 2c_2 e^{2t} \end{bmatrix} = c_1 e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Those solutions with $c_2 = 0$ thus trace out the line spanned by $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, while those with $c_1 = 0$ trace out the line spanned by $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. Moreover, those solutions with $c_2 = 0$ in the long-term approach the origin due to the e^{-t} factor while those with $c_1 = 0$ move away from the origin since e^{2t} gets larger and larger, so we indicate this behavior with arrows:



Now, any other solution combines both of these effects, but in the long-term is dominated by the e^{2t} factor since e^{-t} heads to 0. Thus, any other solution will asymptotically approach those with $c_1 = 0$, meaning that visually they get closer to the line spanned by $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. If we go backwards in time by taking $t \to -\infty$, the opposite happens: the e^{-t} term dominates, so solutions "start off" behaving more like those with $c_2 = 0$. Thus the phase portrait looks like:



Specifying an initial value picks out a specific point in this plane, and the long-term behavior of the corresponding unique solution is fully characterized by this picture. In this case, we call the equilibrium solution $y = 0$ (corresponding to the orbits which consists of the single origin point) a *saddle node* equilibrium. We'll talk about more general types of nodes and equilibriums which can arise when we discuss phase portraits for systems of ODEs, which will subsume this second-order theory. In particular, there are slightly different but related pictures when our characteristic equation has distinct real roots of the *same* sign, as we'll see later.

Complex roots. Next we consider the case where the characteristic equation $r^2 + ar + b = 0$ of

$$y'' + ay' + by = 0$$

has complex non-real solutions. Complex roots of a real quadratic equation come in complex conjugate pairs: if $\alpha + i\beta$ is a root, so is $\alpha - i\beta$. From the root $r = \alpha + i\beta$, we get

$$y = e^{(\alpha+i\beta)t}$$

as a complex solution of the given ODE. But of course, what we are really after are *real solutions*. Using the fact that $e^{i\beta t} = \cos\beta t + i\sin\beta t$ when β is real, the complex solution above can be written as

$$y = e^{\alpha t}e^{i\beta t} = e^{\alpha t}(\cos\beta t + i\sin\beta t) = e^{\alpha t}\cos\beta t + ie^{\alpha t}\sin\beta t.$$

The point now is that the real and imaginary parts of a complex solution of a real linear ODE themselves satisfy that same ODE! That is, if $y = u(t) + iv(t)$ satisfies $L(y) = 0$ for L a real linear differential operator, then:

$$0 = L(u(t) + iv(t)) = L(u) + iL(v),$$

which since $L(u)$ and $L(v)$ are real-valued, requires that $L(u) = 0$ and $L(v) = 0$. Thus in our case, we find that

$$y_1 = e^{\alpha t}\cos\beta t \quad \text{and} \quad y_2 = e^{\alpha t}\sin\beta t$$

are real solutions of $y'' + ay' + by = 0$. Their Wronskian is:

$$W(y_1, y_2) = \det\begin{bmatrix} e^{\alpha t}\cos\beta t & e^{\alpha t}\sin\beta t \\ \alpha e^{\alpha t}\cos\beta t - e^{\alpha t}\beta\sin\beta t & \alpha e^{\alpha t}\sin\beta t + e^{\alpha t}\beta\cos\beta t \end{bmatrix} = e^{2\alpha t}\beta,$$

which is nonzero since $\beta \neq 0$ given that $\alpha + i\beta$ is a non-real root of the characteristic equation. Hence all solutions of the given ODE are of the form

$$y = c_1e^{\alpha t}\cos\beta t + c_2e^{\alpha t}\sin\beta t.$$

Example. Consider the ODE

$$y'' + 2y' + 10y = 0.$$

The characteristic equation is $r^2 + 2r + 10 = 0$, which has roots $r = -1 \pm 3i$. Thus we get $e^{(-1+3i)t}$ as a complex solution to this ODE, which we can write as

$$e^{(-1+3i)t} = e^{-t}e^{i3t} = e^{-t}(\cos 3t + i\sin 3t).$$

After extracting real and imaginary parts, we find that all solutions to $y'' + 2y' + 10y = 0$ are of the form

$$y = c_1e^{-t}\cos 3t + c_2e^{-t}\sin 3t.$$

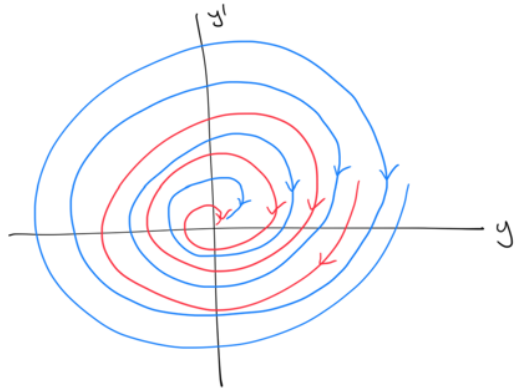
To visualize the phase portrait, we compute:

$$y' = -c_1e^{-t}\cos 3t - 3c_1e^{-t}\sin 3t - c_2e^{-t}\sin 3t + 3c_2e^{-t}\cos 3t.$$

Thus the orbits are parameterized as follows:

$$\begin{aligned}\begin{bmatrix} y \\ y' \end{bmatrix} &= \begin{bmatrix} c_1e^{-t}\cos 3t + c_2e^{-t}\sin 3t \\ -c_1e^{-t}\cos 3t - 3c_1e^{-t}\sin 3t - c_2e^{-t}\sin 3t + 3c_2e^{-t}\cos 3t \end{bmatrix} \\ &= c_1e^{-t}\begin{bmatrix} \cos 3t \\ -\cos 3t - 3\sin 3t \end{bmatrix} + c_2e^{-t}\begin{bmatrix} \sin 3t \\ -\sin 3t + 3\cos 3t \end{bmatrix}.\end{aligned}$$

The pairs of equations in brackets describe ellipses, and the factors of e^{-t} alter the radii by making them get smaller and smaller as t increases, so that in the long-term orbits will spiral in towards the origin. The phase portrait thus looks (roughly) something like:



The origin in this case is called a *spiral* node.

We won't fully justify the fact that the equations in brackets above describe ellipses, but it can be done using linear algebra: the first pair can be written as

$$\begin{bmatrix} \cos 3t \\ -\cos 3t - 3\sin 3t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} \cos 3t \\ \sin 3t \end{bmatrix},$$

and since $\left[\begin{smallmatrix} \cos 3t \\ \sin 3t \end{smallmatrix}\right]$ parametrizes a unit circle, the figure of interest can be determined by understanding the geometric effect of the linear transformation defined by the matrix $\left[\begin{smallmatrix} 1 & 0 \\ -1 & -3 \end{smallmatrix}\right]$, which ends up being some combination of a scaling, reflection, and shear. The same thing applies to orbits for more general second-order ODEs whose characteristic equations have complex roots, where the phase portrait ends up consisting of ellipses or spirals, depending on whether or not there is an exponential term present.

Repeated roots. Finally we consider the case of a second-order linear non-driven ODE with constant coefficients whose characteristic equation has a repeated real root. If $r^2 + ar + b = 0$ has a repeated root r (or in other words a root of multiplicity 2), then the methods described so far only give one independent solution e^{rt}; using the second "root" r gives the same function, and so not one which is linearly independent from the first. So, to look for a second independent solution, we must look elsewhere.

The fact is that in this scenario, $y = te^{rt}$ is also a solution of $y'' + ay' + by = 0$. The idea to consider such a function is still motivated by the fact that its derivatives can be related to itself in a clear way, meaning that its derivatives still involve the same type of expression. However, we should be clear: $y = te^{rt}$ is a solution *only* in the repeated root case, and the same function is *not* a solution in either of the cases considered previously. We'll give a proof of this at the start of next time, but for now take it as given and push on.

So, in the case of a repeated root, we first get the solution $y = e^{rt}$ we had before, but also now $y = te^{rt}$. The Wronskian of these is:

$$W(e^{rt}, te^{rt}) = \det \begin{bmatrix} e^{rt} & te^{rt} \\ re^{rt} & (1+rt)e^{rt} \end{bmatrix} = e^{2rt}.$$

which is nonzero. Hence these solutions are linearly independent, so the general solution of the second-order ODE in this case is

$$y = c_1 e^{rt} + c_2 te^{rt}.$$

Example. Consider the ODE

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

The characteristic equation $r^2 - 6r + 9 = 0$ has repeated root $r = 3$, so the general solution is

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

TO BE FINISHED

Higher-order equations. ***TO BE FINISHED***

Lecture 15: Simple Harmonic Motion

Lecture 16: Undetermined Coefficients

Warm-Up. We determine the values of ω_0 for which the following simpler harmonic oscillator driven ODE has a periodic solution:

$$y'' + \omega_0^2 y = 2\cos 5t.$$

The non-driven equation has general solution given by:

$$y = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t.$$

If we look for a particular solution of the driven equation of the form:

$$y_d = A\cos 5t + B\sin 5t,$$

we get the following requirement:

$$(-25A\cos 5t - 25B\sin 5t) + \omega_0^2(A\cos 5t + B\sin 5t) = 2\cos 5t,$$

which turns into the pair of requirements

$$-25A + \omega_0^2 A = 2 \quad \text{and} \quad -25B + \omega_0^2 B = 0.$$

Thus if $\omega_0 \neq 5$, we get $B = 0$ and $A = \frac{2}{\omega_0^2 - 25}$. Hence the general solution to our driven ODE in this case is

$$y = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t + \frac{2}{\omega_0^2 - 25}\cos 5t.$$

Now, in order for such a solution to be periodic, the periods of the $\cos\omega_0 t, \sin\omega_0 t$ terms have to sync-up with the period of the $\cos 5t$ term. The period of the first pair of terms if $\frac{2\pi}{\omega_0}$, while period of the latter term is $\frac{2\pi}{5}$, so we need some integer multiples of each of these to agree:

$$n\frac{2\pi}{\omega_0} = m\frac{2\pi}{5} \text{ for some integers } n, m.$$

But this requires that $\frac{\omega_0}{5} = \frac{n}{m}$ be a rational number, which happens precisely when ω_0 is rational. Thus we get so far that the given ODE has a periodic solution for rational $\omega_0 \neq 5$.

But for $\omega_0 = 5$ this reasoning does not apply: the guess for y_d used above does not work since this actually satisfies the non-driven equation, and so cannot possibly satisfy the driven equation. So, in this case we look for a particular solution of a different form, say:

$$y_d = At\cos 5t + Bt\sin 5t.$$

(We'll discuss all possible forms to consider soon enough, at least when the driving term consists only of polynomial, exponential, and trigonometric expressions.) We have:

$$\begin{aligned} y_d' &= A\cos 5t - 5At\sin 5t + B\sin 5t + 5Bt\cos 5t \\ y_d'' &= -10A\sin 5t - 25At\cos 5t + 10B\cos 5t - 25Bt\sin 5t \end{aligned}$$

Thus the requirement needed in order for y_d to satisfy our ODE when $\omega_0 = 5$ is:

$$(-10A\sin 5t - 25At\cos 5t + 10B\cos 5t - 25Bt\sin 5t) + 25(At\cos 5t + Bt\sin 5t) = 2\cos 5t,$$

which gives

$$-10A = 0 \quad \text{and} \quad 10B = 2.$$

Thus $A = 0, B = \frac{1}{5}$, so the general solution to our ODE in this case is:

$$y = c_1\cos 5t + c_2\sin 5t + \frac{1}{5}t\sin 5t.$$

Due to the $t\sin t$ term, such a solution will never be periodic, so there are no periodic solutions when $\omega_0 = 5$. Thus, in summary, $y'' + \omega_0^2 y = 2\cos 5t$ has a periodic solution if and only if ω_0 is a rational number different from 5.

Polynomial driving example. We now work towards finding particular solutions of second-order ODEs of the form:

$$y'' + ay' + by = r(t).$$

A first observation is that if the driving term $r(t)$ is made up of multiple summands, we need only focus on one-at-a-time. That is, if we want a solution of

$$L(y) = r_1(t) + r_2(t) + \cdots + r_n(t)$$

where L is a second-order linear differential operator, we can find solutions y_i for the equations

$$L(y) = r_i(t)$$

with single driving terms one-at-a-time, and then add them together: linearity then implies that

$$L(y_1 + \cdots + y_n) = L(y_1) + \cdots + L(y_n) = r_1(t) + \cdots + r_n(t),$$

so $y_1 + \cdots + y_n$ will indeed be a solution of our ODE with driving term $r_1(t) + \cdot + r_n(t)$.

We will only work this out for a special class of driving terms, where the first types we consider are polynomial driving terms. It is easiest to see how this works in explicit examples which will illustrate the general approach. So, consider for now the following ODE:

$$y'' - 2y' - 15y = t + 2t^2.$$

The non-driven equation has solution $y_h = c_1 e^{-3t} + c_2 e^{5t}$. For a particular solution of the driven equation, let us guess that we can find one of the following form:

$$y_d = A_0 + A_1 t + A_2 t^2.$$

This guess comes from wanting a function whose derivative and second derivative will reproduce the t and t^2 portions of the driving term. In order for this be solution, we need:

$$2A_2 - 2(A_1 + 2A_2 t) - 15(A_0 + A_1 t + A_2 t^2) = t + 2t^2.$$

Comparing coefficients of t^k on both sides gives the following requirements:

$$\begin{aligned} 2A_2 - 2A_1 - 15A_0 &= 0 \\ -4A_2 - 15A_1 &= 1 \\ -15A_2 &= 2, \end{aligned}$$

from which we can now solve for A_2, A_1, A_0. In this case we get:

$$A_2 = -\frac{2}{15},\ A_1 = -\frac{7}{15^2},\ A_0 = -\frac{46}{15^3},$$

so the general solution to this driven ODE is

$$y = c_1 e^{-3t} + c_2 e^{5t} - \frac{46}{15^3} - \frac{7}{15^2} t - \frac{2}{15} t^2.$$

But consider now a slight modification of our ODE:

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

In this case a particular guess of $y_d = A_0 + A_1 t + A_2 t^2$ does not work: the y'' and y' completely eliminate any t^2 term, as we see in

$$y_d'' - 2y_d' = 2A_2 - 2(A_1 + 2A_2 t)$$

so that we cannot match up with the required driving term. This was not an issue in the previous example since the y in the ODE itself still maintains a t^2 term, but this y is missing from this new equation. The solution here is to instead consider a guess of the form

$$y_d = A_1 t + A_2 t^2 + A_3 t^3,$$

where we need a cubic guess in order to get t^2 from $y'' - 2y'$. Note that we no longer need to consider a constant term in this guess, since $y'' - 2y'$ will not make use of this constant term anyway. This guess is a particular solution when:

$$(2A_2 + 6A_3 t) - 2(A_1 + 2A_2 + 3A_3 t^2) = t + 2t^2,$$

and by comparing coefficients of t^k we again get a system of equations which can be solved for A_3, A_2, A_1. We'll omit the resulting algebra here.

If instead we had the following ODE:

$$y'' = t + 2t^2,$$

then we would need to consider a quartic (degree 4) guess in order to make y'' retain a quadratic term. In this case our guess can omit constant and degree 1 terms since these play no role in y'', so our guess would be $y_d = A_2t^2 + A_3t^3 + A_4t^4$. Or said another way, in this case y should be obtained simply by anti-differentiating $t + 2t^2$ twice, and such a process does indeed result in a quadratic.

Polynomial driving in general. The same procedure works for any polynomial driving term:

- for an ODE with y present, take y_d to be of the same degree as the driving term;
- for an ODE with y' present but not y, take t_d to be of one degree higher and omit the constant;
- for an ODE with no y' nor y, take t_d to be of two degrees higher, omitting degrees 0 and 1.

The technique being used here is called the *method of undetermined coefficients* since it comes down to working out some unknown coefficients via some algebra.

Operator identities. We will next consider driving terms which are products of polynomial and exponential functions, but first it will be useful to become accustomed to doing certain computations using the algebra of *operator identities*. The point is that when polynomial-exponential driving, it will be necessary to compute derivatives and second derivatives of terms like $t^n e^{rt}$, and the language of differential operators will make this simpler to do.

We denote by $D = \frac{d}{dx}$ the differential operator which sends a function to its derivative, so that a second-order expression like:

$$y'' + ay' + by \text{ can be written as } (D^2 + aD + b)y.$$

To be clear, expanding the final expression gives $D^2y + aDy + by$, where D^2 means to compute the second derivative and D the first. If we denote the characteristic polynomial of the second-order ODE $y'' + ay' + by$ by $p(t)$:

$$p(t) = t^2 + at + b,$$

then we use $p(D)$ to denote $D^2 + aD + b$, where we simply informally replace t by D in $p(t)$. That is, $p(D)$ is the differential operator defined by

$$p(D)[y] = (D^2 + aD + b)[y] = y'' + ay' + by.$$

Here then are the operator identities we will use:

- $p(D)[e^{rt}] = p(r)e^{rt}$. This comes from the fact that each time we differentiate e^{rt} we bring a factor of r down in front, and so altogether computing the derivatives required in $p(D)$ results in a coefficient of $p(s)$;
- $p(D)[h(t)e^{rt}] = e^{rt}p(D + r)[h(t)]$. To clarify, $p(D + r)$ denotes the operator obtained by replacing t by $D + r$ in the polynomial $p(t) = t^2 + at + b$. This identity comes from the product rule: $D[h(t)e^{rt}] = h'(t)e^{rt} + h(t)re^{rt} = e^{rt}D[h(t)] + e^{rt}rh(t) = e^{rt}(D + r)[h(t)]$, and similarly when applying D^2, so that overall we end up applying $p(D + r)$ to $h(t)$;

- $p(D+r) = D^2 + (2r+a)D + p(r)$. This comes from some algebra:

$$(D+r)^2 + a(D+r) + b = D^2 + 2rD + r^2 + aD + ar + b = D^2 + (2r+a)D + p(r)$$

  where $p(r) = r^2 + ar + b$. The point is that applying $p(D+r)$ to a function is the same as applying D^2 to it, applying D and multiplying by $2r+a$, multiplying by $p(r)$, and then adding everything together. How to actually use this will become clear in examples.

Example. Let us consider the following ODE:

$$y'' - 2y' - 15y = t^3 e^{4t}.$$

In looking for a particular solution, we look for one of the following form:

$$y_d = (A_0 + A_1 t + A_2 t^2 + A_3 t^3)e^{4t},$$

where the guess for the cubic portion comes from the t^3 in the driving term. Plugging this into our ODE amounts to applying $p(D) = D^2 - 2D - 15$ to it, and this is where the operator identities derived above will come in handy.

We first have:

$$p(D)[y_d] = e^{4t} p(D+4)[A_0 + A_1 t + A_2 t^2 + A_3 t^3]$$

using the second operator identity with $h(t) = A_0 + A_1 t + A_2 t^2 + A_3 t^3$. Now, the third identity gives:

$$p(D+4) = D^2 + (2 \cdot 4 - 2)D + p(4) = D^2 + 6D - 7$$

where we use $p(t) = t^2 - 2t - 15$ with an "a" value of -2. Thus:

$$p(D+4)[h(t)] = (2A_2 + 6A_3 t) + 6(A_1 + 2A_2 t + 3A_3 t^2) - 7(A_0 + A_1 t + A_2 t^2 + A_3 t^3)$$

where all the work of computing the required derivatives is hidden within the identities we already worked out beforehand, which is what makes these identities so useful! Thus, in order for our guess to actually be a solution, we need $p(D+4)[h(t)]$ to agree with the t^3 in front of e^{4t} in the driving term, so we need:

$$(2A_2 + 6A_3 t) + 6(A_1 + 2A_2 t + 3A_3 t^2) - 7(A_0 + A_1 t + A_2 t^2 + A_3 t^3) = t^3,$$

which results in the equations:

$$\begin{aligned} 2A_2 + 6A_1 - 7A_0 &= 0 \\ 6A_3 + 12A_2 - 7A_1 &= 0 \\ 18A_3 - 7A_2 &= 0 \\ -7A_3 &= 1. \end{aligned}$$

These can now be solved for A_3, A_2, A_1, A_0, thus giving a particular solution to our ODE, and hence—after solving the non-driven equation—the general solution. We will omit the process of solving for the undetermined coefficients here, and indeed we will omit doing so going forward: the process is straightforward, but messy with not-so-nice numbers, and the more enlightening aspects are the forms the guesses should take and the use of operator identities anyway.

Lecture 17: Variation of Parameters

Warm-Up. We solve $y'' + 2y' - 15y = (t + 2t^2)e^{3t}$. The non-driven equation has general solution $y_h = c_1 e^{-5t} + c_2 e^{3t}$. Now, we first look for a general solution of the form:

$$y_d = (A_0 + A_1 t + A_2 t^2)e^{3t}.$$

Using some operator identities, for $p(t) = t^2 + 2t - 15$ we have:

$$p(D)[y_d] = e^{3t} p(D+3)[A_0 + A_1 t + A_2 t^2] = e^{3t}(D^2 + 8D)[A_0 + A_1 t + A_2 t^2].$$

But now we run into a problem: applying the operator $D^2 + 8D$ to this polynomial will not result in a t^2 term, as is needed to match up with the driving term. This is analogous to something we saw with polynomial driving terms alone last time, so we should actually use a polynomial of one degree higher in our guess:

$$y_d = (A_1 t + A_2 t^2 + A_3 t^3)e^{3t}.$$

This happened because the exponential e^{3t} in the driving term *is itself* a solution of the non-driven equation, since 3 is in fact a root of the characteristic equation, so that $p(3) = 0$ in the operator identity for $p(D+3) = D^2 + (2a+3)D + p(3)$.

We our new guess, we have:

$$\begin{aligned} p(D+3)[A_1 t + A_2 t^2 + A_3 t^3] &= (D^2 + 8D)[A_1 t + A_2 t^2 + A_3 t^3] \\ &= (2A_2 + 6A_3 t) + 8(A_1 + 2A_2 t + 3A_3 t^2). \end{aligned}$$

Thus in order for this y_d to actually be a solution, we need:

$$(2A_2 + 6A_3 t) + 8(A_1 + 2A_2 t + 3A_3 t^2) = t + 2t^2,$$

which gives the requirements:

$$2A_2 + 8A_1 = 0 \quad 12A_3 + 16A_2 = 1 \quad 24A_3 = 2,$$

from which we can solve for A_3, A_2, A_1 and get our particular solution.

If we had ended up with $p(D+r) = D^2$ having no D nor constant term, we would have needed to use in y_d a polynomial of two degrees higher than the one in the driving term. This occurs when the characteristic polynomial has a repeated real root which agrees with the factor r which shows up in the exponential e^{rt} part of the driving term.

Polynomial-exponential driving. To summarize what we now know for driving terms which are formed by multiplying polynomials $h(t)$ and exponentials e^{rt}, we have:

- when r is not a root of the characteristic polynomial, take y_d to be a product of e^{rt} and a polynomial of the same degree as $h(t)$;
- when r is a root but not a repeated root, use a polynomial of one degree higher and omit constant terms;
- when r is a repeated root, use a polynomial of two degrees higher and omit the degree 0 and degree 1 terms.

Rather than memorizing what to do when, just keep in mind that it all comes down to what $p(D+r)$ looks like, so compute this as a first step: the first case above corresponds to $p(D+r)$ having a constant term present; the second to D present but no constant term; and the third to no D nor constant term present, only D^2.

Polynomial-exponential-trigonometric driving. Finally we consider driving terms formed by products of polynomials, exponentials, and trigonometric (i.e. sine or cosine) functions. But all the work is essentially already done, since (and this is the point) we can consider such trig functions as arising from complex exponentials. Again, an example will make this clear.

Consider the following ODE:

$$y'' + 2y' - 15y = t^3 e^{3t} \cos 2t.$$

The key idea is to view $e^{3t} \cos 2t$ as being the real part of $e^{(3+2i)t} = e^{3t}(\cos 2t + i \sin 2t)$. So, we instead look for a particular solution of the same ODE but with a complex driving term:

$$y'' + 2y' - 15y = t^3 e^{(3+2i)t},$$

and in the end take the real part of this solution to get a solution to the equation we actually want. Using operator identities, for $p(t) = t^2 + 2t - 15$, we have:

$$\begin{aligned} p(D + (3+2i)) &= D^2 + (2[3+2i] + 2)D + p(3+2i) \\ &= D^2 + (8+4i)D + (-4+16i). \end{aligned}$$

Thus we should look for a particular complex solution of the form

$$y_d = (A_0 + A_1 t + A_2 t^2 + A_3 t^3) e^{(3+2i)t}.$$

In order for this to be a solution, we need:

$$(D^2 + (8+4i)D + (-4+16i))(A_0 + A_1 t + A_2 t^2 + A_3 t^3) = t^3,$$

which becomes

$$(2A_2 + 6A_3 t) + (8+4i)(A_1 + 2A_2 t + 3A_3 t^2) + (-4+16i)(A_0 + A_1 t + A_2 t^2 + A_3 t^3) = t^3.$$

Comparing coefficients of t^k gives four equations, which can then be solved for A_3, A_2, A_1, A_0. The resulting values will be complex, and will give a complex particular solution y_d, whose real part is then a particular solution to our original equation. As stated earlier, we'll omit the details of working out the full solution—the process is straightforward, but messy.

Variation of parameters. So, we can now solve any second-order linear ODE with constant coefficients and polynomial-exponential-trigonometric driving term. (That's quite a mouthful!) Still, the method we've described, that of undetermined coefficients, can be tedious to carry out, and does not shed too much light on what the underlying theory actually is. In addition, this only applies to constant coefficient equations, so says nothing about more general ODEs.

To get a sense for what is actually going on behind the scenes. consider a second-order linear ODE with possibly non-constant coefficients:

$$y'' + p(t)y' + q(t)y = r(t).$$

Suppose we know of two linearly independent solutions y_1, y_2 to the non-driven equation, so that the non-driven equation has general solution

$$y_h = c_1 y_1(t) + c_2 y_2(t).$$

(Of course, so far we only know how to find y_1, y_2 in the case of constant coefficients, but we'll see some more general examples in a bit.) If we now consider the driving term and look for a particular solution, we make the bold guess that a particular solution can be obtained from the form of y_h by allowing the coefficients c_1, c_2 to actually *vary*; that is, we seek a particular solution of the form

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

where $c_1(t), c_2(t)$ are to-be-determined *functions* and not simply constants. This technique is called the method of *variation of parameters*, since we are allowing the "parameters" c_1, c_2 to vary.

We want to determine when this guess will indeed be a solution. First we compute:

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

All we are looking for is some choice of $c_1(t), c_2(t)$ which will give a solution, so we can impose whatever constraints on these we want as long as they lead to a valid solution. So, we impose the condition that:

$$c_1' y_1 + c_2' y_2 = 0.$$

The point is that this will simply some of our computations, since y_d' is now simply

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

Then we get

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

Since y_1, y_2 are themselves solutions of the non-driven ODE, we have:

$$y_1'' = -p(t)y_1' - q(t)y_1 \text{ and } y_2'' = -p(t)y_2' - q(t)y_2,$$

so y_d'' above becomes:

$$y_d'' = c_1' y_1' + c_1(-p(t)y_1' - q(t)y_1) + c_2' y_2' + c_2(-p(t)y_2' - q(t)y_2).$$

Thus $y_d'' + p(t)y_d' + q(t)y_d$ becomes:

$$\begin{aligned} y_d'' + p(t)y_d' + q(t)y_d &= c_1' y_1' + c_1(-p(t)y_1' - q(t)y_1) + c_2' y_2' + c_2(-p(t)y_2' - q(t)y_2 \\ &\quad + p(t)(c_1 y_1' + c_2 y_2' + q(t)(c_1 y_1 + c_2 y_2) \\ &= c_1' y_1' + c_2' y_2' \end{aligned}$$

after everything else cancels out. Hence, $y_d = c_1 y_1 + c_2 y_2$, with the constraint that $c_1' y_1 + c_2' y_2 = 0$, satisfies $y'' + p(t)y' + q(t) = r(t)$ when

$$c_1' y_1' + c_2' y_2' = r(t).$$

This equation together with the constraint can be viewed as a pair of linear equations for the unknowns c_1', c_2':

$$c_1' y_1 + c_2' y_2 = 0$$

$$c_1'y_1' + c_2'y_2' = r(t)$$

which can be solved using an inverse matrix:

$$\begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} \begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \begin{bmatrix} 0 \\ r(t) \end{bmatrix} \rightsquigarrow \begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ r(t) \end{bmatrix}.$$

Note that the matrix used is invertible since its determinant, which is precisely the Wronskian of y_1 and y_2, is nonzero given that y_1, y_2 are assumed to be independent. Using this inverse we get:

$$\begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \frac{1}{y_1y_2' - y_1'y_2} \begin{bmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{bmatrix} \begin{bmatrix} 0 \\ r(t) \end{bmatrix} = \frac{1}{W(y_1, y_2)} \begin{bmatrix} -y_2r(t) \\ y_1r(t) \end{bmatrix}.$$

Thus after integrating, we find that

$$c_1(t) = \int \frac{-y_2r(t)}{W(y_1, y_2)}\, dt \quad \text{and} \quad c_2(t) = \int \frac{y_1r(t)}{W(y_1, y_2)}\, dt.$$

Hence, for these functions c_1, c_2, the function $y_d = c_1(t)y_1(t) + c_2(t)y_2(t)$ is indeed a particular solution of our second-order driven ODE. We'll see next time that this description of the particular solution does in fact lead to a better conceptual description of what is going on in general.

Example. Consider the ODE

$$y'' - 2y' + y = \frac{e^t}{t+1},$$

which is not one we can solve using the method of undetermined coefficients due to the nature of the driving term. But we do know that the non-driven equation has the following independent solutions:

$$y_1 = e^t \quad \text{and} \quad y_2 = te^t$$

since the characteristic equation $r^2 - 2r + 1 = 0$ has a repeated root of $r = 1$.

The method of variation of parameters now gives a way to find a particular solution to our driven ODE. The Wronskian of the independent non-driven solutions above is:

$$W(y_1, y_2) = \det \begin{bmatrix} e^t & te^t \\ e^t & (1+t)e^t \end{bmatrix} = e^{2t}.$$

We thus compute:

$$\begin{aligned} c_1(t) &= \int -\frac{te^t(\frac{e^t}{t+1})}{e^{2t}}\, dt = \int -\frac{t}{t+1}\, dt = \int -\left(1 - \frac{1}{t+1}\right) dt = -t + \ln|t+1| \\ c_2(t) &= \int \frac{e^t(\frac{e^t}{t+1})}{e^{2t}}\, dt = \int \frac{1}{t+1}\, dt = \ln|t+1|. \end{aligned}$$

Hence

$$y_d = \ln|t+1|e^t + (\ln|t+1| - t)te^t$$

is a particular solution of $y'' - 2y' + y = \frac{e^t}{t+1}$, so the general solution is

$$y = c_1e^t + c_2te^t + \ln|t+1|e^t + (\ln|t+1| - t)te^t.$$

(If we look only for solutions with $t > 0$, or even $t > 1$, as would be needed to ensure continuity of the driving term, we can ignore the absolute value.)

Lecture 18: Green's Functions

Warm-Up 1. We think through the process of solving the following ODE:

$$y'' - 7y' + 6y = t\sin 2t - e^{-t}\cos 3t.$$

TO BE FINISHED

Warm-Up 2. We solve the following ODE:

$$t^2y'' - ty' - 3y = t \text{ for } t > 0.$$

We will use variation of parameters, but since this does not have constant coefficients, we need a new way of finding linearly independent solutions of the non-driven form. This particular equation is an example of what's called an *Euler equation*, where the power of t that occurs in each coefficient agrees with the order of the derivative which is taken. The point is that for such equations we can look for solutions of the form $y = t^n$: each derivative we take decreases the power of t which occurs, but which is then compensated for by multiplying by the power of t needed to make all resulting terms still have degree n. In our case, we get:

$$t^2y'' - ty' - 3y = t^2[n(n-1)t^{n-2}] - t[nt^{n-1}] - 3t^n = (n^2 - 2n - 3)t^n.$$

Thus, $y = t^n$ is a solution of $t^2y'' - ty' - 3y = 0$ when $n^2 - 2n - 3 = 0$, so when $n = -1, 3$. Hence $y_1 = t^{-1}$ and $y_2 = t^3$ are two solutions of our non-driven equation, and these are independent since neither is a multiple of the other, or as can be checked using the Wronskian, which is $4t$. The non-driven equation thus has general solution

$$y_h = \frac{c_1}{t} + c_2t^3.$$

Now we can compute:

$$c_1(t) = \int -\frac{t^3 \cdot t}{4t}\,dt = \int -\frac{1}{4}t^4\,dt = -\frac{t^5}{20}$$
$$c_2(t) = \int \frac{t^{-1} \cdot t}{4t}\,dt = \int \frac{1}{4t}\,dt = \frac{\ln t}{4}.$$

Then

$$y_d = \left(-\frac{t^5}{20}\right)t + \left(\frac{\ln t}{4}\right)t^3$$

is a particular solution of our Euler ODE, so the general solution is

$$y = \frac{c_1}{t} + c_2t^3 - \frac{t^4}{20} + \frac{t^3\ln t}{4}.$$

Inverting differential operators. So now we can solve second-order linear ODEs using variation of parameters, at least to the extent that the integrals defining $c_1(t), c_2(t)$ are actually computable. Still, this method also seems to be kind of ad-hoc, coming seemingly out of nowhere.

To obtain a more conceptual understanding of what is underlying everything we've done, we seek to rephrase the problem of solving a linear ODE of terms of *inversion*. Given a second-order linear differential operator L, a particular solution of some driven equation will satisfy

$$L(y_d) = r(t).$$

We can think about the process of finding y_d as that of *producing* y_d *from* $r(t)$. But this process, which turns $r(t)$ into y_d, amounts to "inverting" the operator L:

$$L(y_d) = r(t) \rightsquigarrow y_d = L^{-1}[r(t)]$$

since this "inverse" is the thing which will take $r(t)$ as input and output (i.e. produce) the particular solution y_d as a result. This is analogous to working with inverse functions in general: inverting a function f amounts to recovering x from y in an equation $f(x) = y$, in which case $x = f^{-1}(y)$.

Now, we should be careful here: the operator L does not have a literal inverse in the usual sense, in particular since the equation $L(y_d) = r(t)$ will have multiple solutions (i.e. a given driven ODE will have many different particular solutions), whereas for a literal inverse there should be only one. Nevertheless, thinking of the process of solving a linear ODE as that of "inverting" a differential operator will be a very useful perspective.

Integral kernels. So, how do we invert differential operators? Intuitively, inverting something which involves differentiation should in turn involve some kind of integration. So, we might expect that the inverse of a linear differential operator should be what's called an *integral operator*, which is a linear operator sending functions to functions defined by some type of integral. In particular, given a two-variable function $K(t,s)$, we can define an integral operator via the following:

$$r(t) \mapsto \int_{t_0}^{t} K(t,s)r(s)\,ds.$$

To be clear, this is the operator which takes a function $r(t)$ as input and outputs the function $\int_{t_0}^{t} K(t,s)r(s)\,ds$—with t as the independent variable—obtained by integrating r against f. In this integral, we run through all possible values of $r(s)$ (with s varying as the variable of integration) and use the remaining variable t in $K(t,s)$ to characterize the input of the resulting function. The function $K(t,s)$ is called the *kernel* of the resulting operator.

To put this into the right context, I claim that what we've done here is *precisely* an infinite-dimensional analog of matrix multiplication! Indeed, suppose $A = (a(i,j))$ is a matrix, with entry $a(i,j)$ in the i-th row and j-th column. Multiplying by a vector $\mathbf{x}$ gives a result like:

$$A\mathbf{x} = \begin{bmatrix} a(1,1)x_1 + \cdots + a(1,n)x_n \\ a(2,1)x_1 + \cdots + a(2,n)x_n \\ \vdots \\ a(m,1)x_1 + \cdot + a(m,n)x_n \end{bmatrix}.$$

The i-th entry in the result is the sum

$$\sum_j a(i,j)x_j$$

where the first coordinate i is fixed and the second j varies. Now, replace $A = (a(i,j))$ by a two-variable function $K(t,s)$ and $\mathbf{x}$ by a function $r(t)$. If we replace summation by its infinite analog of integration, the sum above turns into

$$\int K(t,s)r(s)\,ds$$

where we fix the "coordinate" t and allow s to vary among all possible values this "coordinate" can take. As this second coordinate varies, so too do the values of $r(s)$ just as the values of x_j do in the matrix-vector product above. The resulting value describes the "t-th" entry in the result, which is the value of the resulting function at t, just as the result of $\sum_j a(i,j)x_j$ describes the i-th coordinate of the result. So, if we interpret $K(t,s)$ as a type of "infinite-by-infinite" sized matrix and $r(t)$ as a type of "infinite" vector, then the result of integral operator defined above applied to $r(t)$ is indeed the analog of multiplying a matrix by a vector in this scenario. (To be more precise, we should of $K(t,s)$ as an "uncountably-sized" matrix and of $r(s)$ as an "uncountable" vector, whose entires are indexed by the real numbers.)

The idea that the inverse of L should be defined by an infinite analog of a matrix is due to the fact that L is linear: the inverse of a linear operator should itself be linear (at least, this is true in the finite-dimensional setting of linear algebra), and linear things are in fact defined by matrices. All of this can be made much more precise than what we're describing here, but this is good enough for our purposes.

Green's functions. So, given a second-order linear differential operator L, we can hope to find a function $K(t,s)$ which we can use to construct an "inverse" for L, allowing us to construct a particular solution y_d to a driven ODE from the driving term $r(t)$:

$$L(y_d) = r(t) \rightsquigarrow y_d = \int_{t_0}^{t} K(t,s)r(s)\,ds.$$

In this setting, $K(t,s)$ is called the *Green's function*, or *Green's kernel*, of L, and provides a conceptually cleaner way of characterizing particular solutions.

What still remains is to find $K(t,s)$ explicitly, but I claim that this is precisely what the method of variation of parameters gives! Suppose $L(y) = y'' + p(t)y' + q(t)$ is our differential operator. Variation of parameters gives:

$$y_d = y_1(t)\int \frac{-y_2(s)r(s)}{W(y_1,y_2)(s)}\,ds + y_2(t)\int \frac{y_1(s)r(s)}{W(y_1,y_2)(s)}\,ds$$

where we now use s for the variable of integration to distinguish it from t. Since both integrals are taken with respect to s, we can bring $y_1(t)$ and $y_2(t)$ inside, to get:

$$\begin{aligned} y_d &= \int \frac{-y_1(t)y_2(s)r(s)}{W(y_1,y_2)(s)}\,ds + \int \frac{y_2(t)y_1(s)r(s)}{W(y_1,y_2)(s)}\,ds \\ &= \int \frac{-y_1(t)y_2(s)r(s) + y_2(t)y_1(s)r(s)}{W(y_1,y_2)(s)}\,ds \\ &= \int \left(\frac{-y_1(t)y_2(s) + y_2(t)y_1(s)}{W(y_1,y_2)(s)}\right) r(s)\,ds. \end{aligned}$$

Thus, the two-variable function $K(t,s)$ which satisfies $y_d = \int K(t,s)r(s)\,ds$ as required of a Green's function for L is

$$K(t,s) = \frac{-y_1(t)y_2(s) + y_2(t)y_1(s)}{W(y_1,y_2)(s)}.$$

It turns out that this function is independent of the specific linearly independent non-driven solutions we used: if f_1, f_2 is another pair of linearly independent solutions of $L(y) = 0$, the Green's function $K(t,s)$ computed using f_1, f_2 will be the same as the one computed using y_1, y_2, which will come from the fact that each of f_1, f_2 can be written as linear combinations of y_1, y_2, and vice-versa. We'll omit this verification.

Example. Consider the linear differential operator defined by $L = D^2 + \frac{5}{t}D + \frac{3}{t^2}$. Then $L(y) = 0$ corresponds (after multiplying through by t^2) to the Euler equation

$$t^2y'' + 5ty' + 3y = 0.$$

The function $y = t^n$ is a solution of this when $n(n-1) + 5n + 3 = 0$, so when $n = -1, -3$. Thus $y_1 = t^{-1}$ and $y_2 = t^{-3}$ are two linearly independent solutions of $L(y) = 0$. The Wronskian of these solutions is: $y_1 y_2' - y_1' y_2 = \frac{1}{t}\frac{-3}{t^4} - \frac{-1}{t^2}\frac{1}{t^3} = -\frac{2}{t^5}$.

The Green's function for $L = D^2 + \frac{5}{t}D + \frac{3}{t^2}$ is thus:

$$\begin{aligned} K(t,s) &= \frac{-y_1(t)y_2(s) + y_2(t)y_1(s)}{W(y_1,y_2)(s)} \\ &= \frac{-\frac{1}{t}\frac{1}{s^3} + \frac{1}{t^3}\frac{1}{s}}{-\frac{2}{s^5}} \\ &= \frac{s^2}{2t} - \frac{s^4}{2t^3}. \end{aligned}$$

The point is then that a particular solution of $L(y_d) = r(t)$ can always be found as the value of the integral $\int (\frac{s^2}{2t} - \frac{s^4}{2t^3}) r(s)\, ds$, which will still have t as a parameter.

Lecture 19: Nonlinear Pendulums

TO BE FINISHED

Lecture 20: The Laplace Transform

TO BE FINISHED

Lecture 21: Linear Differential Systems

Warm-Up. ***TO BE FINISHED***

Differential Systems. We now move to studying *systems* of differential equations, also termed *differential systems*. We have already seen three examples of these before:

- the scenario towards the start of the quarter where some mass of uranium radioactively decayed into thorium, which itself experienced radioactive decay, resulting in one equation which described the rate of change of the mass of uranium and another the change in the mass of thorium, the latter of which depended on the mass of uranium as well;
- the Lotka-Volterra predator-prey equations, characterizing the rates of change in the populations of predator and prey, each dependent on both populations; and
- the pair of first-order equations which arise when converting a second-order linear equation for y by introducing a new variable $v = y'$.

Our final goal for the quarter is to study such systems more systematically.

To be clear, we consider *vector-valued* functions $\mathbf{x}(t)$ whose components are the real-valued functions we care about. A differential system then looks like

$$\mathbf{x}'(t) = f(\mathbf{x}, t)$$

where f is a vector-valued function as well, with variables $x_1, \dots, x_n, t$ where $\mathbf{x} = (x_1, \dots, x_n)$ in terms of components. The left-side $\mathbf{x}'(t)$ simply denotes the vector obtained by differentiating each entry of $\mathbf{x}(t)$:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \rightsquigarrow \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}.$$

(It will be convenient to write vectors as columns, as in a linear algebra course.) The single equation

$$\mathbf{x}'(t) = f(\mathbf{x}, t)$$

then encodes n total ODEs (if $\mathbf{x}$ is an n-dimensional vector)—one for $x_1(t)$, one for $x_2(t)$, and so on—obtained by equating a component of the left side with the corresponding component on the right. We'll see that phrasing this collection of n equations as a single vector equation makes things conceptually nicer.

Linear Systems. For now, until next quarter, we will focus exclusively on *linear* differential systems. (This covers the first and third examples above, but not the Lotka-Volterra model, unless we look at its *linearization* instead.) Linearity is phrased in the same way as we've seen before: the system $\mathbf{x}'(t) = f(\mathbf{x}, t)$ is linear when the differential operator L defined by

$$L[\mathbf{x}(t)] = \mathbf{x}'(t) - f(\mathbf{x}, t)$$

is linear in the sense that $L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$ and $L(c\mathbf{x}) = cL(\mathbf{x})$. As with previous linear equations, the linearity property implies that all solutions can be formed by adding a particular driven solution to arbitrary non-driven solutions, and so the work comes down to finding these two types of solutions.

Practically, a linear system is one which can be written in the form:

$$\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{F}(t)$$

where $A(t)$ is a *matrix-valued* function, which you can think of as being a matrix with functions as entries. If $A(t)$ concretely has entries $a_{ij}(t)$ (i corresponds to the row and j to the column), then the linear system above corresponds to the following collection of equations, where we denote by $f_1, \dots, f_n$ the components of $\mathbf{F}(t)$:

$$\begin{aligned} x_1'(t) &= a_{11}(t)x_1(t) + \cdots + a_{1n}(t)x_n(t) + f_1(t) \\ x_2'(t) &= a_{21}(t)x_1(t) + \cdots + a_{2n}(t)x_n(t) + f_2(t) \\ &\vdots \\ x_n'(t) &= a_{n1}(t)x_1(t) + \cdots + a_{nn}(t)x_n(t) + f_n(t) \end{aligned}$$

The right sides come from computing the product of the matrix $A(t)$ with the vector $\mathbf{x}(t)$, and adding the vector $\mathbf{F}(t)$.

Example. Consider the following linear differential system:

$$\begin{aligned} x_1' &= 7x_1 + 2x_2 \\ x_2' &= -4x_1 + x_2. \end{aligned}$$

In matrix form, this can be written as $\mathbf{x}' = \left[\begin{smallmatrix} 7 & 2 \\ -4 & 1 \end{smallmatrix}\right] \mathbf{x}$. From the first equation, we can express x_2 in terms of x_1 as $x_2 = \frac{1}{2}x_1' - \frac{7}{2}x_1$. Then $x_2' = \frac{1}{2}x_1'' - \frac{7}{2}x_1'$, so the second equation becomes

$$\frac{1}{2}x_1'' - \frac{7}{2}x_1' = -4x_1 + \left(\frac{1}{2}x_1' - \frac{7}{2}x_1\right), \text{ or } x_1'' - 8x_1' + 15x_1 = 0.$$

But this is now a second-order equation for x_1 with constant coefficients, which we already know how to solve: the characteristic equation has roots 3 and 5, so the general solution is

$$x_1(t) = c_1 e^{3t} + c_2 e^{5t}.$$

Then, using the expression for x_2 in terms of x_1, we get:

$$x_2 = \frac{1}{2}x_1' - \frac{7}{2}x_1 = -2c_1 e^{3t} - c_2 e^{5t}.$$

If we write our solution in vector form, we get the following nice expression:

$$\mathbf{x}(t) = \begin{bmatrix} c_1 e^{3t} + c_2 e^{5t} \\ -2c_1 e^{3t} - c_2 e^{5t} \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{5t}.$$

This technique is fine for two-dimensional systems, but is a bit tedious and becomes even more so when considering higher-dimensional systems. If we had three equations for three unknown functions, we would have to use one equation to express one function in terms of the other two, use this to rewrite the remaining two equations in terms of only two functions, then do what we did above to turn this into a higher-order equation for one function alone, and so on. But, of course, we'll see that this is not necessary, once we develop a cleaner approach. In this two-dimensional example, we could then specify some initial values, which will single out specific c_1 and c_2, and we could also consider some driving terms, but again we'll eventually see a better way of doing this in more generality.

A general approach. As stated earlier, solutions of a linear system $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{F}(t)$ are of the form:

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_d(t)$$

where $\mathbf{x}_h$ is a solution of the non-driven equation $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ and $\mathbf{x}_d(t)$ is a particular solution of the driven equation. Moreover, it turns out, for a reason similar to that we saw for second-order linear equations, that all solutions of the non-driven system can be obtained once we have enough linearly independent solutions. In other words, the space of solutions to an n-dimensional system $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ is itself n-dimensional, and so is spanned by n linearly independent solutions. We won't repeat the entire argument here, but the idea is that if $\phi_1(t), \ldots, \phi_n(t)$ are linearly independent solutions ("linear independence" can be characterized using a version of the *Wronskian*), then a solution of the form

$$c_1\phi_1(t) + \cdots + c_n\phi_n(t)$$

can be found for *any* specified initial condition, and so by uniqueness of solutions these expressions must describe all possible solutions. (Yes, there is an existence and uniqueness theorem for differential systems just as for single equations, which we alluded to earlier when discussing where existence and uniqueness for solutions of second-order linear ODEs came from. The idea behind the proof is similar to that for first-order single ODEs.)

The first goal is thus to find such linearly independent solutions. We'll see this possible to do explicitly at least for the case of a *constant coefficient matrix A*. Thus we consider a system of the form

$$\mathbf{x}'(t) = A\mathbf{x}(t)$$

where A is a matrix with constant entries. When $A = \begin{bmatrix} a \end{bmatrix}$ is 1×1, this equation is simply $x' = ax$, where we know that the solutions are multiples of an exponential. Thus by analogy, we might first guess that solutions of $\mathbf{x}' = A\mathbf{x}$ will also be related to certain exponentials. In particular, we can ask when a vector function of the form

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v} = e^{\lambda t}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} e^{\lambda t} v_1 \\ e^{\lambda t} v_2 \end{bmatrix},$$

where $\mathbf{v}$ is a constant vector, is a solution. In this case,

$$\mathbf{x}' = \begin{bmatrix} \lambda e^{\lambda t} v_1 \\ \lambda e^{\lambda t} v_2 \end{bmatrix} = \lambda e^{\lambda t}\mathbf{v},$$

so $x = e^{\lambda t}\mathbf{v}$ satisfies $\mathbf{x}' = A\mathbf{x}$ if and only if

$$\lambda e^{\lambda t}\mathbf{v} = A(e^{\lambda t}\mathbf{v}) = e^{\lambda t}A\mathbf{v}, \text{ or equivalently } \lambda\mathbf{v} = A\mathbf{v}$$

since $e^{\lambda t}$ is nonzero. But this final equation says that $\mathbf{v}$ is an *eigenvector* of A with *eigenvalue* λ, so the key observation is that there is a relation between solutions of $\mathbf{x}' = A\mathbf{x}$ and eigenvalues/eigenvectors of A, which we'll explore next time.

Lecture 22: Eigenvector Solutions

Warm-Up. We solve the following system of linear ODEs:

$$\begin{aligned} x_1' &= x_1 + x_2 \\ x_2' &= -5x_1 - x_2. \end{aligned}$$

For now, we proceed by a brute-force method, which will soon be made clearer using eigenvectors. The first equation gives $x_2 = x_1' - x_1$, which after substituting into the second gives the following second-order equation for x_1:

$$x_1'' - x_1' = -5x_1 - (x_1' - x_1).$$

To be clear, the left side is x_2', and the right side is $-5x_1 - x_2$. This final equation can be written as

$$x_1'' + 4x_1 = 0,$$

which, using what we know about second-order equations, has solution

$$x_1 = c_1 \cos 2t + c_2 \sin 2t.$$

Then

$$x_2 = x_1' - x_1 = c_1(-2\sin 2t - \cos 2t) + c_2(2\cos 2t - \sin 2t),$$

so we have solved our system.

Real distinct eigenvalues. As we saw at the end of last time, there is a relation between solutions of $\mathbf{x}' = A\mathbf{x}$, for a matrix A with constant entries, and eigenvalues/eigenvectors of A. To be precise, the observation was that $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ is a solution of $\mathbf{x}' = A\mathbf{x}$ if and only if $\mathbf{v}$ is an eigenvector of A with eigenvalue λ. Thus, if our goal is to find enough linearly independent solutions of the non-driven equation $\mathbf{x}' = A\mathbf{x}$, we can start by finding eigenvectors of A.

For now we focus on two-dimensional systems described by 2×2 matrices. As a first step, if A has distinct eigenvalues of λ_1, λ_2, then we are done: if $\mathbf{v}_1, \mathbf{v}_2$ denote eigenvectors corresponding to the distinct eigenvalues λ_1, λ_2 respectively, we get two solutions:

$$e^{\lambda_1 t}\mathbf{v}_1 \quad \text{and} \quad e^{\lambda_2 t}\mathbf{v}_2$$

which will be linearly independent since $\mathbf{v}_1$ and $\mathbf{v}_2$ are (it is a fact from linear algebra that eigenvectors corresponding to different eigenvalues are always linearly independent), and so these two alone will span the space of all solutions.

Example. Consider the example

$$\mathbf{x}' = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix} \mathbf{x}$$

we looked at last time. Then we used the approach of expressing one function in terms of the other, and then deriving a second-order equation that this latter function must satisfy. But this is not necessary: the matrix $\left[\begin{smallmatrix} 7 & 2 \\ -4 & 1 \end{smallmatrix}\right]$ has eigenvalues 3 and 5, so if we find eigenvectors for each we will be able to write down the general solution right away. One possible eigenvector for 3 is $\left[\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right]$, and a possible eigenvector for 5 is $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, so we get

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{3t} \quad \text{and} \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{5t}$$

as linearly independent solutions. Thus the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{5t}.$$

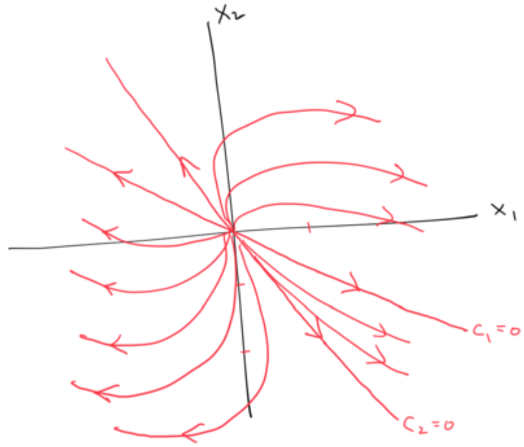
Note that this is precisely the same form we found last time using the previous approach! Of course, this is no accident, and I hope you will agree that the eigenvector approach is much cleaner.

Phase portraits. As was the case for second-order equations, we can discuss *phase portraits* for first-order systems. Here, the *orbits* are the curves traced out by the parametric equations encoded within a solution $\mathbf{x}(t)$, so the phase portrait is a plot of these curves in the x_1x_2-plane.

In the example above we had the following general solution:

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{5t}.$$

Those solutions with $c_2 = 0$ trace out the line spanned by $\left[\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right]$, while those with $c_1 = 0$ trace out the line spanned by $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. (Technically, the origin is its own orbit, corresponding to the constant equilibrium solution $\mathbf{x} = \mathbf{0}$.) Due to both the e^{3t} and e^{5t} terms, solutions along *both* of these lines move away from the origin. Long-term, as $t \to \infty$, the e^{5t} dominates the e^{3t} term, which means that other orbits in general asymptotically approach a direction parallel to $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, since solutions behave more and more like the second term $c_2 \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right] e^{5t}$. All orbits, going backwards in time as $t \to -\infty$, approach the origin closer to the direction of $\left[\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right]$ since for $t \to -\infty$ the e^{3t} term dominates since it goes to zero more slowly than e^{5t}. Thus, the phase portrait looks like:



The origin here is called an *improper node*, or more precisely an *unstable* improper node, where the instability is due to the fact that solutions move away from the equilibrium at the origin. (Arrows going the other way, when solutions approach the origin, is the case of a *stable* node.)

Complex (non-real) eigenvalues. Next we consider the case where A has complex eigenvalues with nonzero imaginary part. We specify "nonzero imaginary part" since real eigenvalues can also be considered to be complex, only with zero imaginary part. Going forward, when we say "complex eigenvalues" we will always means ones which are not real.

So, suppose A has eigenvalues $\alpha \pm i\beta$. (It is also a fact from linear algebra that complex eigenvalues for a real matrix come in complex conjugate pairs.) Take one, say $\alpha + i\beta$. For this then we can find a *complex eigenvector* $\mathbf{v}$, which we can express in terms of a real and imaginary part $\mathbf{v} = \mathbf{a} + i\mathbf{b}$ where $\mathbf{a}, \mathbf{b}$ are real vectors. We thus get

$$e^{(\alpha+i\beta)t}(\mathbf{a} + i\mathbf{b})$$

as a complex solution of $\mathbf{x}' = A\mathbf{x}$. But, the real and imaginary parts of *this* complex solution will then be *real* solutions of $\mathbf{x}' = A\mathbf{x}$, which will in fact be linearly independent, so they span the space of all solutions. This will become clear when looking at an example in a second, but the point is that if $\mathbf{x}(t) = \mathbf{u}(t) + i\mathbf{v}(t)$ is a complex solution of $\mathbf{x}' = A\mathbf{x}$, we have:

$$\mathbf{u}'(t) + i\mathbf{v}'(t) = A(\mathbf{u}(t) + i\mathbf{v}(t)) = A\mathbf{u}(t) + iA\mathbf{v}(t),$$

so equating real and imaginary parts on both sides gives $\mathbf{u}' = A\mathbf{u}$ and $\mathbf{v}' = A\mathbf{v}$, which is why $\mathbf{u}$ and $\mathbf{v}$ themselves are real solutions of the given system.

Example. Consider the system from the Warm-Up:

$$\mathbf{x}' = \begin{bmatrix} 1 & 1 \\ -5 & -1 \end{bmatrix} \mathbf{x}.$$

The given matrix A has eigenvalues $\pm 2i$. For the eigenvalue $2i$, we have:

$$A - (2i)I = \begin{bmatrix} 1-2i & 1 \\ -5 & -1-2i \end{bmatrix}.$$

Here is a nice fact: a vector satisfying $(A - 2iI)\mathbf{v} = \mathbf{0}$ can be obtained simply by swapping the entries in the first row and changing the sign of one of them:

$$\mathbf{v} = \begin{bmatrix} 1 \\ -1+2i \end{bmatrix}.$$

This comes from the fact that when row-reducing the second row will become zero, since this matrix is meant to be noninvertible, so all we need are v_1, v_2 satisfying $(1-2i)v_1 + v_2 = 0$, and this swapping method always (in the 2×2 case) gives values which work.

Thus we get

$$\begin{bmatrix} 1 \\ -1+2i \end{bmatrix} e^{2it} = \begin{bmatrix} 1 \\ -1+2i \end{bmatrix} (\cos 2t + i \sin 2t)$$

as a complex solution of our system. Extracting real and imaginary parts gives

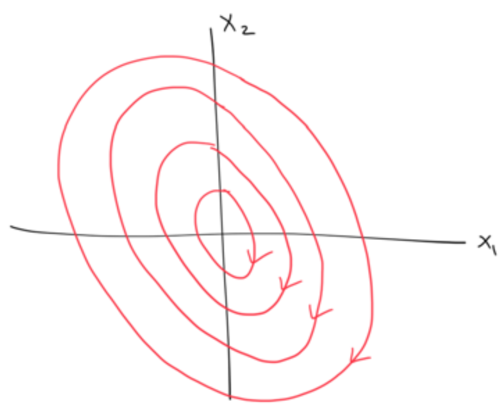
$$\begin{bmatrix} \cos 2t \\ -\cos 2t - 2\sin 2t \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \sin 2t \\ -\sin 2t + 2\cos 2t \end{bmatrix}$$

as two independent real solutions, so the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} \cos 2t \\ -\cos 2t - 2\sin 2t \end{bmatrix} + c_2 \begin{bmatrix} \sin 2t \\ -\sin 2t + 2\cos 2t \end{bmatrix}.$$

Comparing with what we did in the Warm-Up, this is precisely the same solution obtained there.

The parametric equations within the first vector in the solution, or the second, describe ellipses for the same reason as what we saw in some second-order examples previously: these equations can be obtained from $\left[\begin{smallmatrix} \cos 2t \\ \sin 2t \end{smallmatrix}\right]$, which describe a circle, via some combination of rotation, reflection, scaling, or shearing. Thus the orbits in this case are all ellipses, so the phase portrait looks like:



The origin in this case is called a *center point*. Sometimes we throw in the adjective *neutrally stable* center since, although not technically "stable" since we do not approach the origin, the solutions remain bounded as closed (periodic) orbits.

Spirals vs centers. In the example above, we got ellipses overall since there were no exponential terms which affected distance to the origin. Instead, if we had obtained solutions of the form

$$\mathbf{x}(t) = c_1 e^{2t} \begin{bmatrix} \cos 2t \\ -\cos 2t - 2\sin 2t \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} \sin 2t \\ -\sin 2t + 2\cos 2t \end{bmatrix}$$

the orbits would look like spirals, so the origin would be a *spiral* point; in this case unstable, but it would stable if we had exponentials with negative exponents instead. Centers arise when these exponential factors are missing.

Lecture 23: More on Eigenvector Solutions

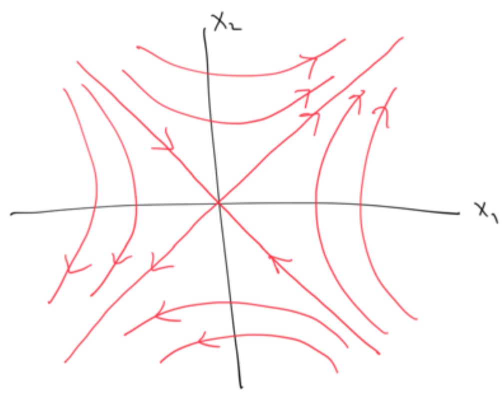
Warm-Up 1. We solve the system

$$\mathbf{x}'(t) = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \mathbf{x}(t).$$

The given matrix has eigenvalues $-2, 4$, with possible eigenvectors given by $\left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ respectively. Thus the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t}.$$

Two orbits are the lines spanned by the two eigenvectors, and more general orbits asymptotically approach the line corresponding to $c_1 = 0$ as $t \to \infty$ since the e^{4t} dominates and e^{-2t} goes to 0. Going backwards in time, the e^{-2t} dominates, so solutions approach the line corresponding to $c_2 = 0$ when $t \to -\infty$. The phase portrait is thus:



The origin here is an unstable *saddle* node. Notice we saw a similar picture in some second-order examples previously, which is explained by the fact that this first-order system theory subsumes the previous second-order theory, as we'll clarify next time.

Warm-Up 2. Now consider

$$\mathbf{x}' = \begin{bmatrix} -1 & -5 \\ 1 & -1 \end{bmatrix} \mathbf{x}.$$

The defining matrix A has eigenvalues $-1 \pm i\sqrt{5}$. For the eigenvalue $-1 + i\sqrt{5}$, we have:

$$A - (-1 + i\sqrt{5})I = \begin{bmatrix} -1 - (-1 + i\sqrt{5}) & -5 \\ 1 & -1 - (-1 + i\sqrt{5}) \end{bmatrix} = \begin{bmatrix} -i\sqrt{5} & -5 \\ 1 & -i\sqrt{5} \end{bmatrix}.$$

A possible complex eigenvector is thus given by $\left[\begin{smallmatrix} -5 \\ i\sqrt{5} \end{smallmatrix}\right]$, so

$$\begin{bmatrix} -5 \\ i\sqrt{5} \end{bmatrix} e^{(-1+i\sqrt{5})t} = \begin{bmatrix} -5 \\ i\sqrt{5} \end{bmatrix} e^{-t}(\cos\sqrt{5}t + i \sin\sqrt{5}t).$$

is a complex solution. Thus, using real and imaginary parts, we get following the general real solution:

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -5\cos\sqrt{5}t \\ -\sqrt{5}\sin\sqrt{5}t \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} -5\sin\sqrt{5} \\ \sqrt{5}\cos\sqrt{5}t \end{bmatrix} e^{-t}.$$

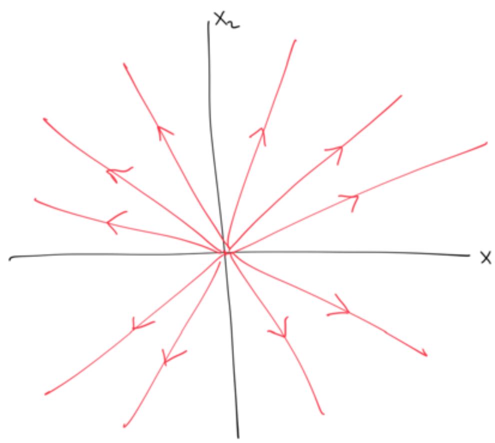
The e^{-t} factor causes solutions to approach $\mathbf{0}$ in the long-term, so the origin is a stable spiral node and the phase portrait consists of orbits spiraling in towards the origin.

Repeated eigenvalues. Finally we come to the case of a 2×2 matrix A with a repeated real eigenvalue λ, or in other words an eigenvalue of multiplicity 2. One possibility is that even with only one eigenvalue, we can still find two linearly independent eigenvectors, so that we can still find two linearly independent solutions of $\mathbf{x}' = A\mathbf{x}$ which will then span all solutions.

For instance, consider the system $\mathbf{x}' = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right]\mathbf{x}$. The given matrix only has 2 as an eigenvalue, but nonetheless we can find two linearly independent eigenvectors, for instance $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$. (In fact, in this case any vector in $\mathbb{R}^2$ is an eigenvector.) These give

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{2t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{2t} = e^{2t} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

as the general solution. The values of c_1, c_2 are determined by an initial condition, so we can see that the non-constant orbits are open rays emanating from the origin, obtained by taking all positive multiples of a fixed initial vector $\left[\begin{smallmatrix} c_1 \\ c_2 \end{smallmatrix}\right]$. The phase portrait thus looks like:



and the equilibrium at the origin is called a *star* node, which in this case is unstable. In the 2×2 case, such a node—coming from two linearly independent eigenvectors with the same eigenvalue—only arises with matrices which are multiples of the identity.

Deficient eigenspaces. But for 2×2 matrices with repeated eigenvalues which are not multiples of the identity, or in other words for 2×2 matrices which are not *diagonalizable*, an alternate approach is needed. If A is a 2×2 matrix with an eigenvalue λ of multiplicity 2 with only a one-dimensional eigenspace (sometimes referred to as being a *deficient* eigenspace), we only get one linearly independent solution of the form

$$e^{\lambda t}\mathbf{v}.$$

Thus, to get a second linearly independent solution we must look beyond ones of this form alone.

We saw a similar phenomenon arise with second-order linear equations with constant coefficients whose characteristic polynomial had a repeated root. In that case, the answer was to also consider solutions of the form te^{rt}. (In fact, we'll soon see that this phenomenon for second-order equations is just a special case of what we'll now develop for differential systems applied to the system obtained by converting a second-order ODE into a pair of first-order ODEs.) So now, we take a similar approach and consider next solutions of the form

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v} + te^{\lambda t}\mathbf{w}.$$

(Using $\mathbf{x} = te^{\lambda t}\mathbf{w}$ alone as a guess won't work: this satisfies $\mathbf{x}' = A\mathbf{x}$ only when $e^{\lambda t}\mathbf{w} + t\lambda e^{\lambda t}\mathbf{w} = te^{\lambda t}A\mathbf{w}$, but this requires that $e^{\lambda t}\mathbf{w}$ be zero, which requires that $\mathbf{w}$ be zero.) The guess above satisfies $\mathbf{x}' = A\mathbf{x}$ if and only if the following holds:

$$\lambda e^{\lambda t}\mathbf{v} + e^{\lambda t}\mathbf{w} + t\lambda e^{\lambda t}\mathbf{w} = A(e^{\lambda t}\mathbf{v} + te^{\lambda t}\mathbf{w}).$$

After distributing A on the right side and dividing by the nonzero term $e^{\lambda t}$, this turns into

$$\lambda\mathbf{v} + \mathbf{w} + t\lambda\mathbf{w} = A\mathbf{v} + tA\mathbf{w}.$$

Comparing constant terms and coefficients of t on both sides shows that $\mathbf{x}(t) = e^{\lambda t}\mathbf{v} + te^{\lambda t}\mathbf{w}$ satisfies $\mathbf{x}' = A\mathbf{x}$ if and only if

$$A\mathbf{v} = \lambda\mathbf{v} + \mathbf{w} \quad \text{and} \quad A\mathbf{w} = \lambda\mathbf{w}.$$

Thus by finding such $\lambda, \mathbf{w}, \mathbf{v}$, we get a second solution to $\mathbf{x}' = A\mathbf{x}$, which will be linearly independent from the first (still to-be-checked), and then we can form the general solution.

Generalized eigenvectors. The second equation derived above says that $\mathbf{w}$ should be an eigenvector of A with eigenvalue λ. But $\mathbf{v}$ is not quite an eigenvector, since $A\mathbf{v}$ does not equal a multiple of $\mathbf{v}$, but rather a multiple of $\mathbf{v}$ plus the eigenvector $\mathbf{w}$. The equation $A\mathbf{v} = \lambda\mathbf{v} + \mathbf{w}$ which $\mathbf{v}$ should satisfy can be rewritten as

$$(A - \lambda I)\mathbf{v} = \mathbf{w}$$

after subtracting $\lambda\mathbf{v}$ from both sides and factoring out the common $\mathbf{v}$. Since $\mathbf{w}$ is an eigenvector, we have $(A - \lambda I)\mathbf{w} = 0$ (this is just a rewritten form of $A\mathbf{w} = \lambda\mathbf{w}$), so we get that

$$(A - \lambda I)^2\mathbf{v} = (A - \lambda I)\mathbf{w} = \mathbf{0}.$$

A vector having such a property is called a *generalized eigenvector* of A corresponding to the eigenvalue λ.

To be precise, $\mathbf{u}$ is a *generalized eigenvector* of A with eigenvalue λ if $(A-\lambda I)^k\mathbf{u} = \mathbf{0}$ for some k. Those generalized eigenvectors which satisfy this condition for $k = 1$ are just ordinary eigenvectors, but for more general generalized eigenvectors a higher power of $A - \lambda I$ may be necessary to result in a value of $\mathbf{0}$. If $\mathbf{u}$ is a generalized eigenvector and k is the smallest power needed in order to make $(A - \lambda I)^k\mathbf{u} = \mathbf{0}$, then the vectors obtained by repeatedly applying $A - \lambda I$ up to the point before you start getting $\mathbf{0}$:

$$\mathbf{u},\ (A - \lambda I)\mathbf{u},\ (A - \lambda I)^2\mathbf{u},\ \ldots,\ (A - \lambda I)^{k-1}\mathbf{u}$$

are always linearly independent (this is a linear algebra fact we won't prove here), and this list is called the *Jordan chain* generated by $\mathbf{u}$. Note that the final vector in this chain, $(A - \lambda I)^{k-1}\mathbf{u}$ is an ordinary eigenvector of A since applying $A - \lambda I$ to it results in zero. In this language, $\mathbf{x} = e^{\lambda t}\mathbf{v} + te^{\lambda t}\mathbf{w}$ is a solution of $\mathbf{x}' = A\mathbf{x}$ in the repeated eigenvalue case when $\mathbf{v}$ is a generalized eigenvector of A corresponding to λ and $\mathbf{v}, \mathbf{w}$ (recall $\mathbf{w} = (A - \lambda I)\mathbf{v}$) are the vectors in the Jordan chain it generates.

Thus from a Jordan chain $\mathbf{v}, \mathbf{w} = (A-\lambda I)\mathbf{v}$ corresponding to λ, we get two solutions to $\mathbf{x}' = A\mathbf{x}$, namely:

$$e^{\lambda t}\mathbf{w} \quad \text{and} \quad e^{\lambda t}\mathbf{v} + te^{\lambda t}\mathbf{w}.$$

The Wronskian of these two solutions, which is simply the determinant of the matrix which has these as its columns, is:

$$\det \begin{bmatrix} e^{\lambda t}w_1 & e^{\lambda t}v_1 + te^{\lambda t}w_1 \\ e^{\lambda t}w_2 & e^{\lambda t}v_2 + te^{\lambda t}w_2 \end{bmatrix} = e^{\lambda t} \det \begin{bmatrix} w_1 & v_1 \\ w_2 & v_2 \end{bmatrix},$$

which is nonzero since $\mathbf{w}, \mathbf{v}$ are linearly independent. Hence these solutions are linearly independent, so they span the space of all solutions to $\mathbf{x}' = A\mathbf{x}$.

Example 1. We consider the system

$$\mathbf{x}' = \begin{bmatrix} -1 & 1 \\ -4 & 3 \end{bmatrix} \mathbf{x}.$$

The given matrix A has a repeated eigenvalue of 1, which only gives rise to one linearly independent eigenvector, such as $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. Thus so far we only have

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^t$$

as a solution. To get a second, we look for a generalized eigenvector $\mathbf{v}$ satisfying $(A - I)\mathbf{v} = \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, which becomes

$$\begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix} \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

One possible solution is $\mathbf{v} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, so

$$e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + te^t \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is a second solution to our system, linearly independent from the first. Thus the general solution is

$$\mathbf{x}(t) = c_1 e^t \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \left(e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + te^t \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right)$$

Next we determine the phase portrait. Solutions corresponding to $c_2 = 0$ trace out the line spanned by $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, and in this case this is the only "eigendirection", or in other words the only orbit which looks like a line. The node at the origin is unstable since solutions blow-up as $t \to \infty$ due to the e^t term. But asymptotically, the te^t term gives the dominant behavior, which means that orbits will asymptotically approach a direction parallel to the single eigenline.

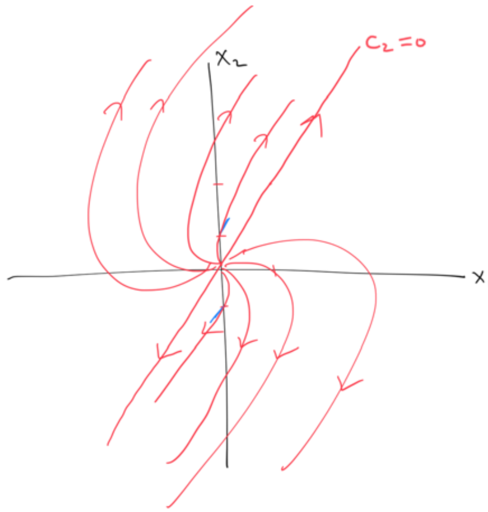
Furthermore, we can determine which way the orbits go by imagining their tangent vectors along, say, the y-axis. For instance, the given ODE says that the tangent vector to a solution at $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ should be

$$\mathbf{x}' = \begin{bmatrix} -1 & 1 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

which means that a solution will emanate from the origin, curve to hit this tangent vector at $y = 1$ on the y-axis, and then go on getting more and more parallel to the line spanned by $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. Similarly, at $\left[\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right]$ the tangent vector should be

$$\mathbf{x}' = \begin{bmatrix} -1 & 1 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \end{bmatrix}.$$

Thus, the phase portrait looks like:



The origin in this case is called an unstable *deficient node*.

Example 2.

Higher-dimensional systems.

Lecture 24: Matrix Exponentials

Warm-Up.

Subsumes second-order theory. We've alluded to a connection between second-order linear ODEs with constant coefficients and linear systems with constant coefficients multiple times, so now we will make this relation clear. Given an ODE of the form

$$y'' + ay' + by = 0,$$

by introducing a new variable $v = y'$, we know we can rewrite the given ODE as the following first-order system:

$$\begin{bmatrix} y' \\ v' \end{bmatrix} = \begin{bmatrix} v \\ -by - av \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} y \\ v \end{bmatrix}.$$

The characteristic polynomial (whose roots are eigenvalues) of the matrix defining this system is:

$$\det \begin{bmatrix} -\lambda & 1 \\ -b & -a-\lambda \end{bmatrix} = \lambda^2 + a\lambda + b,$$

and the point is that this is the same as the characteristic polynomial of the second-order ODE. Thus, the roots of the characteristic equation are the eigenvalues of the matrix defining the corresponding first-order system, and the different cases we saw before for second-order equations—distinct roots, complex roots, repeated root—are the same as the ones which occur for linear systems. Moreover, the form which solutions take—exponentials, exponential times sine or cosine, t times an exponential—are the same as those for systems. The phase portraits are even the same! So, in summary, theory of first-order systems completely contains within it the entire theory of second-order constant coefficient linear equations, which explains many of similarities we've seen.

Solution matrices. If $\mathbf{x}_1, \mathbf{x}_2$ are two independent solutions of a system $\mathbf{x}' = A\mathbf{x}$, we can take them as the columns of a matrix $X(t)$, often called a *solution matrix* of the system:

$$X(t) = \begin{bmatrix} \mathbf{x}_1(t) & \mathbf{x}_2(t) \end{bmatrix}.$$

Since each column satisfies $\mathbf{x}' = A\mathbf{x}$, this matrix satisfies the following property:

$$X'(t) = \begin{bmatrix} \mathbf{x}_1'(t) & \mathbf{x}_2'(t) \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1(t) & A\mathbf{x}_2(t) \end{bmatrix} = A \begin{bmatrix} \mathbf{x}_1(t) & \mathbf{x}_2(t) \end{bmatrix} = AX(t).$$

In other words, the matrix-valued function $X(t)$ satisfies the *matrix differential equation*

$$X'(t) = AX(t).$$

For example, consider an example we looked at previously:

$$\mathbf{x}' = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \mathbf{x} \text{ with general solution } \mathbf{x}(t) = c_1 e^{-2t} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Taking $c_1 = 1, c_2 = 0$ and $c_1 = 0, c_2 = 1$ gives two solutions we can use to form a solution matrix:

$$X(t) = \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix}.$$

Then

$$X'(t) = \begin{bmatrix} 2e^{-2t} & 4e^{4t} \\ -2e^{-2t} & 4e^{4t} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} X(t).$$

In general, if $X(t)$ is a solution matrix of $\mathbf{x}' = A\mathbf{x}$ and B is any constant matrix, then $X(t)B$ is also a solution matrix. One way to see this is to note that the columns of $X(t)B$ will just be linear combinations of the columns of $X(t)B$, and so will still be solutions of the given system. Alternatively, we can verify that $X(t)B$ still satisfies the necessary matrix differential equation: we have $X'(t) = AX(t)$, so

$$(X(t)B)' = X'(t)B = AX(t)B,$$

which means that the derivative of $X(t)B$ is A times $X(t)B$ itself.

A cleaner approach. We know how to solve any non-driven system $\mathbf{x}' = A\mathbf{x}$ with constant-coefficient matrix A. But the approach we took, finding eigenvalue and eigenvectors, is not so clean and does not shed much light on what is going on behind the scenes. So we look for something which puts everything in the proper context.

As a guide, consider the case of a first-order non-driven ODE with constant coefficients: $y' = ay$. Here we know the solution to be $y = ce^{at}$. Since the type of equation $\mathbf{x}' = A\mathbf{x}$ are of the same form—derivative is a "constant" multiple of the function itself—we might expect that naively the solution should also be of the same form, so something like $\mathbf{x} = e^{At}\mathbf{c}$. Pushing things further, we will soon consider driven linear systems like $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$, and since the solution to the analogous single-equation ODE $y' = ay + f(t)$ looks like

$$y = ce^{at} + ce^{at} \int e^{-at} f(t)\, dt$$

via the method of integrating factors, we might expect that the solution to our driven system should also be expressible in a similar way, using something like e^{At} and integral involving e^{-At}.

But, of course, none of this makes sense unless we know what "e^{At}" is actually supposed to mean. After all, A here denotes a matrix, and At the matrix obtained by multiplying A by the scalar t, so what on earth does it mean to take e to the power of a matrix? We cannot literally multiply e by itself a "matrix-number" of times, so e^{At} must mean something else. But if can indeed find a way to make sense of this, we can hope to obtain a cleaner approach towards expressing the solutions of first-order linear systems as a result.

Matrix exponentials. To motivate the definition of e^{At}, or even just e^A, we will soon give, we think back to alternate interpretations of e^x when x is a real number. If we don't want to think about this as meaning "multiply e by itself so many times", we can instead view it as the result of the following infinite summation:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

But notice this expression still makes sense if we replace x by a matrix A! Indeed, the resulting expression just involves computing powers of A, which is something we know how to do.

So, we will define the *matrix exponential* e^A to be precisely the following infinite summation:

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}A^n$$

where we interpret A^0 as the identity matrix I. Now, as is the case when dealing with any type of "infinite sum", we have to think about the notion of *convergence*, or in other words whether this infinite sum makes sense as producing a finite value. We will not go into the details here, but it turns out that this can be made fully precise: there is a sense in which we can measure the "absolute value" of a matrix, and then use this to argue that the infinite sum above does converge for any matrix A. For those of you who are familiar with notion of the *operator norm* $|A|$ of a matrix A, the key point is that the following inequality holds:

$$|e^A| \leq e^{|A|}$$

and the fact that $e^{|A|}$ exists (meaning that the series defining it converges) for all real numbers $|A|$ implies that the infinite sum of matrices defining e^A exists (or converges) as well. (The inequality above comes from an infinite version of the triangle inequality.)

It is now fair to ask why it is that the infinite sum defining e^A above deserves to be called an "exponential"? One answer is that we came up with it using the infinite series expression for e^x, but it there a more satisfying reason? Sure! Consider the following definition of e^{At} for t a real variable:

$$e^{At} = \sum_{n=0}^{\infty} \frac{1}{n!}(At)^n = \sum_{n=0}^{\infty} \frac{t^n}{n!}A^n.$$

Differentiating with respect to t (in which case you treat the matrix A as a constant) gives:

$$\frac{d}{dt}\left(\sum_{n=0}^{\infty} \frac{t^n}{n!}A^n\right) = \sum_{n=0}^{\infty} \frac{d}{dt}\left(\frac{t^n}{n!}A^n\right) = \sum_{n=1}^{\infty} \frac{nt^{n-1}}{n!}A^n = A\sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!}A^{n-1} = A\sum_{n=0}^{\infty} \frac{t^n}{n!}A^n.$$

But this says that the derivative of e^{At} is Ae^{At}, which is precisely the property you expect an exponential to have! Said another way, this shows that the matrix-valued function $X(t) = e^{At}$

satisfies the matrix matrix differential equation $X'(t) = AX(t)$, and since $e^{A0} = e^0 = I$ (the second 0 here denotes the zero matrix), $X(t) = e^{At}$ satisfies the matrix IVP

$$X'(t) = AX(t),\ X(0) = I,$$

which in the case of a single equation:

$$y'(t) = ay(t),\ y(0) = 1$$

does uniquely characterize exponentials. Indeed, saying that e^{At} is the unique matrix-valued function which satisfies the matrix IVP $X' = AX, X(0) = I$ is what the book and other sources take to be the *definition* of e^{At}, from which the series definition can then be derived.

Computing exponentials. So, we can now characterize the general solution of the linear system $\mathbf{x}' = A\mathbf{x}$ by $\mathbf{x}(t) = e^{At}\mathbf{c}$ where $\mathbf{c}$ is an arbitrary constant vector. Indeed, the computations above show that

$$\mathbf{x}'(t) = Ae^{At}\mathbf{c} = A\mathbf{x}(t),$$

so $\mathbf{x} = e^{At}\mathbf{c}$ does satisfy this first-order system, and varying $\mathbf{c}$ will produce solutions satisfying any possible initial value, so by uniqueness $\mathbf{x} = e^{At}\mathbf{c}$ does give all solutions. In particular, since $e^{A0} = I$, the solution which satisfies the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is just $\mathbf{x}(t) = e^{At}\mathbf{x}_0$, again mimicking what happens for single first-order equations.

But none of this is helpful in practice if we cannot actually compute such matrix exponentials, and the series definition does not make it clear that this can actually be done. But, we'll see that these exponentials are fairly straightforward (although possibly tedious at times) to compute. For a first approach, we use the characterization of e^{At} as the unique solution of the matrix IVP

$$X'(t) = AX(t),\ X(0) = I.$$

Take $Y(t)$ to be any solution matrix. Then $Y(t)B$ is also a solution matrix for any constant matrix B, so we simply seek B which will satisfy the initial condition $Y(0) = I$. This turns into the requirement that

$$X(0)B = I,\ \text{so } B = X(0)^{-1}.$$

(Note that $X(0)$ is invertible since the columns of $X(t)$ are linearly independent.) Thus $X(t)X(0)^{-1}$ is the solution of the matrix IVP above, so we get that

$$e^{At} = X(t)X(0)^{-1}$$

for any choice of solution matrix $X(t)$.

Example. Consider the system

$$\mathbf{x}' = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \mathbf{x}$$

we looked at earlier, where we came up with the following solution matrix:

$$X(t) = \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix}.$$

For $A = \left[\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right]$, we have:

$$e^{At} = X(t)X(0)^{-1}$$

$$= \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1}$$
$$= -\frac{1}{2} \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
$$= -\frac{1}{2} \begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix}.$$

Thus the general solution of this system can be written as:

$$\mathbf{x}(t) = e^{At}\mathbf{c} = -\frac{1}{2} \begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

which if you multiply and simplify does agree with the form of the general solution we found before.

The benefit of using this specific form of the solution is that, as mentioned earlier, it makes IVP's immediate to solve: the solution which satisfies $\mathbf{x}(0) = \mathbf{x}_0$ is precisely $\mathbf{x}(t) = e^{At}\mathbf{x}_0$, coming from the fact that $e^{A0} = I$. If we instead use the previous form of the general solution we found:

$$\mathbf{x}(t) = c_1 e^{-2t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c + 2e^{4t} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

finding the solution satisfying $\mathbf{x}(0) = \mathbf{x}_0$ would involve some messy algebra needed to find the appropriate values of c_1, c_2. So, matrix exponentials make expressing solutions conceptually cleaner.

Lecture 25: Driven Systems

Warm-Up. Consider the system

$$\mathbf{x}' = \begin{bmatrix} -1 & 1 \\ -4 & 3 \end{bmatrix} \mathbf{x}.$$

We looked at this previously as an example of a matrix with a repeated eigenvalue, and found the general solution to be

$$\mathbf{x}(t) = c_1 e^t \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \left(e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + te^t \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right).$$

Thus using the solutions with $c_1 = 1, c_2 = 0$ and $c_1 = 0, c_2 = 1$, we get the solution matrix

$$X(t) = \begin{bmatrix} e^t & te^t \\ 2e^t & e^t + 2te^t \end{bmatrix}.$$

Thus, for $A = \left[\begin{smallmatrix} -1 & 1 \\ -4 & 3 \end{smallmatrix}\right]$, we get:

$$e^{At} = X(t)X(0)^{-1}$$
$$= \begin{bmatrix} e^t & te^t \\ 2e^t & e^t + 2te^t \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} e^t & te^t \\ 2e^t & e^t + 2te^t \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} e^t - 2te^t & te^t \\ -4te^t & e^t + 2te^t \end{bmatrix}.$$

The solution of the IVP with $\mathbf{x}(0) = \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$ is thus:

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) = \begin{bmatrix} e^t - 2te^t & te^t \\ -4te^t & e^t + 2te^t \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

Computing exponentials via similarity. Computing matrix exponentials using solution matrices is fine, but we can still give a conceptually better approach. In the end this ends up being just a reformulation of the solution matrix approach, but it highlights the role of another concept in linear algebra—that of the *Jordan form* of a matrix. We will elaborate on this in a bit.

Consider the case of a 2×2 matrix A with real and distinct eigenvalues $\lambda_1 \neq \lambda_2$. Take respective linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2$. Then we can write A in the following factored form:

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{-1}.$$

Recall from linear algebra that this happens because A is *diagonalizable*, which for a size $n \times n$ matrix means either that it has are n linearly independent eigenvectors, or equivalently that it can be written as $A = SDS^{-1}$ for an invertible matrix S and diagonal matrix D. We will use this diagonalized form $A = SDS^{-1}$ to compute e^{At}. To recall some other language from linear algebra, we say here that A is *similar* to D. We will also use the fact that that with $A = SDS^{-1}$, we have:

$$A^k = (SDS^{-1})^k = \underbrace{(SDS^{-1})(SDS^{-1})(SDS^{-1})\cdots(SDS^{-1})}_{k \text{ times}} = SD^kS^{-1}$$

since all the intermediate $S^{-1}S$ terms cancel out. This way of computing powers is what makes similarity so useful.

Using the series definition of e^A, we have:

$$\begin{aligned} e^A &= I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots \\ &= I + SDS^{-1} + \frac{1}{2}(SDS^{-1})^2 + \frac{1}{3!}(SDS^{-1})^3 + \cdots \\ &= SIS^{-1} + SDS^{-1} + \frac{1}{2}SD^2S^{-1} + \frac{1}{3!}SD^3S^{-1} + \cdots \\ &= S\left(I + D + \frac{1}{2}D^2 + \frac{1}{3!}D^3 + \cdots\right)S^{-1} \\ &= Se^DS^{-1}. \end{aligned}$$

Thus computing e^A when $A = SDS^{-1}$ is similar to D comes down to computing e^D. We also have $At = S(Dt)S^{-1}$, so $e^{At} = Se^{Dt}S^{-1}$.

But now we exploit the fact that $Dt = \left[\begin{smallmatrix} \lambda_1 t & 0 \\ 0 & \lambda_2 t \end{smallmatrix}\right]$ is a diagonal matrix. The point is that in this case, we have:

$$(Dt)^k = \begin{bmatrix} \lambda_1 t & 0 \\ 0 & \lambda_2 t \end{bmatrix}^k = \begin{bmatrix} \lambda_1^k t^k & 0 \\ 0 & \lambda_2^k t^k \end{bmatrix},$$

so we can compute:

$$\begin{aligned} e^{Dt} &= I + Dt + \frac{1}{2}(Dt)^2 + \frac{1}{3!}(Dt)^3 + \cdots \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 t & 0 \\ 0 & \lambda_2 t \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \lambda_1^2 t^2 & 0 \\ 0 & \lambda_2^k t^k \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} \lambda_1^3 t^3 & 0 \\ 0 & \lambda_2^3 t^3 \end{bmatrix} + \cdots \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} 1+\lambda_1 t + \frac{1}{2}\lambda^2 t^2 + \frac{1}{3}\lambda_1^3 t^3 + \cdots & 0 \\ 0 & 1+\lambda_2 t + \frac{1}{2}\lambda_2^2 t^2 + \frac{1}{3}\lambda_2^3 t^3 + \cdots \end{bmatrix} \\
&= \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix},
\end{aligned}
$$

where the point is that the resulting infinite summations in the upper-left and lower-right are precisely the series definitions of $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ respectively. So, computing the exponential of a diagonal matrix comes down to exponentiating each diagonal entry! Putting this together with what we did above, we get:

$$
A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{-1} \rightsquigarrow e^{At} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{-1}
$$

Example. Consider the matrix

$$
A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}
$$

we considered in an example last time. This has eigenvalues $-2, 4$, and respective eigenvectors $\left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. Thus this diagonalizes as:

$$
A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1}.
$$

Hence the exponential is:

$$
e^{At} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-2t} & 0 \\ 0 & e^{4t} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -e^{-2t} & e^{4t} \\ e^{-2t} & e^{4t} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}^{-1},
$$

and as this point we'll simply point out that this final product is precisely the $X(t)X(0)^{-1}$ we encountered last time when computing this exponential using a solution matrix, so we will end up with the same answer for e^{At}.

But this is in fact true in general: the result of

$$
\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t}\mathbf{v}_1 & e^{\lambda_2 t}\mathbf{v}_2 \end{bmatrix}
$$

is a solution matrix of $\mathbf{x}' = Ax$ since its columns are the solutions we find true the eigenvector method, and the value of this solution matrix at $t = 0$ is just $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$. Thus computing the exponential using the diagonalized expression:

$$
e^{At} = \underbrace{\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix}}_{X(t)} \underbrace{\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{-1}}_{X(0)^{-1}}
$$

amounts to the same computation $X(t)X(0)^{-1}$ in the solution matrix method. The approach using similarity with a diagonal matrix just makes the role of the diagonal matrix clear.

Jordan forms. ***TO BE FINISHED***

Constant driving. Now we move to studying driven linear systems. As stated earlier, linearity implies that all solutions are of the form

$$
\mathbf{x} = \mathbf{x}_h + \mathbf{x}_d
$$

where $\mathbf{x}_h$ is a solution of the non-driven equation and $\mathbf{x}_d$ a particular solution of the driven equation. As a first case, consider a constant-coefficient linear system with constant driving term:

$$\mathbf{x}' = A\mathbf{x} + \mathbf{F}$$

where $\mathbf{F}$ is a constant vector. The *equilibrium* constant solution occurs when $\mathbf{x}' = \mathbf{0}$, so when

$$0 = A\mathbf{x} + \mathbf{F}.$$

If A is invertible (which we are generally always assuming), this has a unique solution given by $\mathbf{x} = -A^{-1}\mathbf{F}$, which is thus the equilibrium solution of $\mathbf{x}' = A\mathbf{x} + \mathbf{F}$. Taking this as a "particular" solution, all solutions are then of the form

$$\mathbf{x}(t) = e^{At}\mathbf{c} - A^{-1}\mathbf{F}.$$

Compare this with the analogous single ODE: $y' = ay + f$ where f is constant. Using integrating factors we get the solution to be:

$$y(t) = ce^{at} + e^{at}\int e^{-at} f\,dt = ce^{at} + e^{at}\left(-\frac{1}{a}e^{-at}f\right) = ce^{at} - \frac{1}{a}f,$$

which is precisely the same as $e^{At}\mathbf{c} - A^{-1}\mathbf{F}$ when A is a 1×1 matrix. Thus the form of the solution of the higher-dimensional system does reproduce what we already know about integrating factors for first-order linear ODEs, at least for a constant driving term.

Example. Consider the system

$$\mathbf{x}' = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

The non-driven equation has general solution

$$\mathbf{x}_h(t) = -\frac{1}{2}\begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

using the matrix exponential for the defining matrix we computed previously. The equilibrium solution of the driven system is:

$$\mathbf{x}_d(t) = -\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -5/8 \\ -1/8 \end{bmatrix},$$

so the general solution is:

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_d(t) = -\frac{1}{2}\begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} -5/8 \\ -1/8 \end{bmatrix}.$$

If we want, we can just absorb the extra factor of $-\frac{1}{2}$ in front into the the values of c_1 and c_2.

Shifted nodes. ***TO BE FINISHED***

General driving. Consider now a general first-order driven linear system with constant coefficients:

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{F}(t).$$

In trying to find a particular solution, and in lieu of a clear "integrating factor" analog, we consider instead a different approach we saw before in the second-order case: variation of parameters. (Actually, we'll see that the integrating factor approach is essentially variation of parameters in disguise.) That is, the non-driven equation has general solution $\mathbf{x}_h = e^{At}\mathbf{c}$, so we now look for a particular solution of the form

$$\mathbf{x}_d = e^{At}\mathbf{c}(t)$$

where we allow the parameter $\mathbf{c}$ to vary.

In order for $\mathbf{x}_d = e^{At}\mathbf{c}(t)$ to satisfy $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$, we need the following to hold:

$$Ae^{At}\mathbf{c}(t) + e^{At}\mathbf{c}'(t) = Ae^{At}\mathbf{c}(t) + \mathbf{F}(t)$$

where the left side is obtained by differentiating $e^{At}\mathbf{c}(t)$ using the product rule. Cancelling like terms gives

$$e^{At}\mathbf{c}'(t) = \mathbf{F}(t), \text{ or } \mathbf{c}'(t) = e^{-At}\mathbf{F}(t).$$

To be clear, e^{At} is always invertible, and its inverse is $(e^{At})^{-1} = e^{-At}$. (There are multiple ways of seeing this, but we'll skip the details here. Try to work out what you get if you multiply the series defining e^{At} with the one defining e^{-At}. We should be careful of one thing though: naively you might try to say that $e^{At}e^{-At} = e^{At-At} = e^0 = I$, which is in fact true but is a bit subtle since the equality $e^Ae^B = e^{A+B}$ does not actually hold for matrices in general—it only holds for matrices which *commute*! The general relation between e^Ae^B and e^{A+B} is given by what's called the *Campbell-Baker-Hausdorff* formula, but that is not a topic for this course.)

Integrating $\mathbf{c}'(t) = e^{-At}\mathbf{F}(t)$ gives

$$\mathbf{c}(t) = \int_0^t e^{-As}\mathbf{F}(s)\,ds$$

where have chosen a specific antiderivative—the one which satisfies $\mathbf{c}(0) = \mathbf{0}$—by setting the lower bound on the integral to be 0, and we have renamed the variable of integration to be s, for reasons to be seen in a second. (The integral of a vector-valued function is defined to be the vector-valued function obtained by integrating each individual component.) The particular solution we get is

$$\mathbf{x}_d(t) = e^{At}\mathbf{c}(t) = e^{At}\int_0^t e^{-As}\mathbf{F}(s)\,ds = \int_0^t e^{At}e^{-As}\mathbf{F}(s)\,ds = \int_0^t e^{(t-s)A}\mathbf{F}(s)\,ds$$

where we can bring e^{At} inside the integration since it is dependent of s, and we use the fact that At and $-As$ commute to combine $e^{At}e^{-As}$ into e^{At-As}. Hence altogether, the general solution of $\mathbf{x}' = A\mathbf{x} + \mathbf{F}(t)$ looks like

$$\mathbf{x}(t) = \mathbf{x}_h(t) + \mathbf{x}_d(t) = e^{At}\mathbf{c} + \int_0^t e^{(t-s)A}\mathbf{F}(s)\,ds$$

An initial value $\mathbf{x}(0) = \mathbf{x}_0$ gives $\mathbf{c} = \mathbf{x}_0$, so we can write our solution as

$$\mathbf{x}(t) = \underbrace{e^{At}\mathbf{x}_0}_{\text{response to initial data}} + \underbrace{\int_0^t e^{(t-s)A}\mathbf{F}(s)\,ds}_{\text{response to external driving}}.$$

Note that this is precisely the same form we had for first-order linear ODEs of the form $y' = ay + f(t)$ at the start of the course using integrating factors:

$$y(t) = x_0e^{at} + e^{at}\int_0^t e^{-as}f(s)\,ds$$

where we can bring e^{at} inside the integration. The point is that the linear system theory subsumes everything we did previously for single equations.

Lecture 26: Steady States

Warm-Up. We solve the following driven IVP:

$$\mathbf{x}' = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 \\ e^t \end{bmatrix}, \ \mathbf{x}(0) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

The defining matrix A is the one we've used repeatedly before, and has exponential

$$e^{At} = -\frac{1}{2} \begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix}.$$

The solution is thus:

$$\begin{aligned} \mathbf{x}(t) &= e^{At} \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \int_0^t e^{(t-s)A} \begin{bmatrix} 2 \\ e^s \end{bmatrix} ds \\ &= -\frac{1}{2} \begin{bmatrix} -e^{-2t} - e^{4t} & e^{-2t} + e^{4t} \\ e^{-2t} - e^{4t} & -e^{-2t} + e^{4t} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \int_0^t -\frac{1}{2} \begin{bmatrix} -e^{-2(t-s)} - e^{4(t-s)} & e^{-2(t-s)} + e^{4(t-s)} \\ e^{-2(t-s)} - e^{4(t-s)} & -e^{-2(t-s)} + e^{4(t-s)} \end{bmatrix} \begin{bmatrix} 2 \\ e^s \end{bmatrix} ds. \end{aligned}$$

To be clear, the matrix $e^{(t-s)A}$ within the integral is obtained by replacing t by $t - s$ in the matrix exponential used in the non-driven part of the solution.

This is straightforward, albeit tedious, to compute. By "straightforward" I simply mean that computing this will involve integrating a bunch of exponential terms (recall that the integral of a vector-valued function is defined to be what you get when integrating each component one-at-a-time), which is easy to do, just with so many of them we have to take care to get everything right. At this point, using a computer to solve is probably more practically useful, but the form which the solution takes as "response to initial data" plus "response to external driving" is conceptually important nonetheless.

Constant steady states. Consider the system

$$\mathbf{x}' = \begin{bmatrix} 2 & -4 \\ 5 & -7 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 6 \\ 6 \end{bmatrix}.$$

The general solution of the non-driven equation is

$$\mathbf{x}_h(t) = c_1 e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

which is found by computing eigenvalues and eigenvectors. The driven system has equilibrium solution given by:

$$\mathbf{x}_d(t) = -\begin{bmatrix} 2 & -4 \\ 5 & -7 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = -\frac{1}{6} \begin{bmatrix} -7 & 4 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$

so the general solution of the driven system is

$$\mathbf{x}(t) = c_1 e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 4 \\ 5 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

The observation here is that, as $t \to \infty$, the general solution approaches the equilibrium solution since the e^{-2t} and e^{-3t} factors both go to 0. In this case we call the equilibrium solution $\mathbf{x}_d = \left[\begin{smallmatrix} 3 \\ 3 \end{smallmatrix}\right]$ a *constant steady state* of the system, which means that all other solutions approach it. This equilibrium is a shifted saddle node, and the phase portrait looks like:

Any system $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{F}$ with constant driving term defined by a matrix A with negative eigenvalues will have such a constant steady state. Indeed, the general solution of such a system is of the form

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 - A^{-1}\mathbf{F},$$

and if λ_1, λ_2 are both negative, the exponential factors will approach zero as $t \to \infty$, meaning that $\mathbf{x}(t)$ always approaches the equilibrium $-A^{-1}\mathbf{F}$. This doesn't happen when A has a positive eigenvalue since then the corresponding exponential factor does not approach zero. More generally, complex eigenvalues with negative real part also lead to a constant steady state (the phase portrait will be a shifted stable spiral), since such eigenvalues also result in an exponential with negative exponent and thus goes to zero.

Periodic steady states. In addition to constant steady states, other steady states to consider are those which are *periodic*. First off, if

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{F}(t)$$

is to have a periodic solution, then the driving term $\mathbf{F}(t)$ must itself be periodic. Indeed, if $\mathbf{x}(t)$ is a periodic solution of period T, then $\mathbf{x}(t+T) = \mathbf{x}(t)$, so differentiating gives $\mathbf{x}'(t+T) = \mathbf{x}'(t)$. But

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{F}(t) \quad \text{and} \quad \mathbf{x}'(t+T) = A\mathbf{x}(t+T) + \mathbf{F}(t+T) = A\mathbf{x}(t) + \mathbf{F}(t+T),$$

so if these are to be equal we need $\mathbf{F}(t) = \mathbf{F}(t+T)$, which implies that $\mathbf{F}$ is periodic.

So let us assume $\mathbf{F}(t)$ is periodic. Then we claim that the given system has a unique periodic solution. Indeed, we know we can express any solution as:

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) + \int_0^t e^{(t-s)A}\mathbf{F}(s)\,ds.$$

In order for this to be periodic of period T, we need in particular $x(0) = x(T)$, which turns into:

$$\mathbf{x}(0) = e^{AT}\mathbf{x}(0) + \int_0^T e^{(T-s)A}\mathbf{F}(s)\,ds.$$

The question is whether there exists an initial vector $\mathbf{x}(0)$ which has this property. After rewriting this equality by regrouping terms as

$$(I - e^{AT})\mathbf{x}(0) = \int_0^T e^{(T-s)A}\mathbf{F}(s)\,ds,$$

we see that there exists a unique such $\mathbf{x}(0)$ as long as $I - e^{AT}$ is invertible. But $(I - e^{AT}) = -(e^{AT} - I)$ is invertible as long as 1 is not an eigenvalue of e^{AT}; in general, if λ is an eigenvalue of A, $e^{\lambda T}$ is an eigenvalue of e^{AT}, so we want to avoid having $e^{\lambda T}$ equal 1, which means we want to avoid having λT be an integer multiple of $2\pi i$.

Thus, as long as no integer multiple of $\frac{2\pi i}{T}$ is an eigenvalue of A, a unique vector $\mathbf{x}(0)$ satisfying the equality above exists, and this is then the initial value of solution which will be periodic. If A only has real nonzero eigenvalues, or more generally complex eigenvalues with nonzero real part, then we are in the clear for sure. Moreover, if the eigenvalues of A are negative, or complex with negative real part, all solutions will asymptotically approach the periodic solution as $t \to \infty$: all solutions can be written as

$$\mathbf{x}(t) = (\text{non-driven solution}) + (\text{the periodic solution}),$$

and the non-driven solutions will approach zero due to the presence of exponential factors with negative exponents, similarly to the constant steady state case. In this case, we call the periodic solution a *periodic steady state* of the driven system.

Example. The system

$$\mathbf{x}'(t) = \begin{bmatrix} 2 & -4 \\ 5 & -7 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}$$

has periodic driving term and defining matrix with negative eigenvalues $-2, -3$. Thus this has a periodic steady state. To find it, we use the fact that all solutions are of the form

$$\mathbf{x}(t) = c_1 e^{-2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 4 \\ 5 \end{bmatrix} + (\text{periodic solution})$$

where we use the expression for the non-driven solution we gave previously. Thus, we can take *any* solution and think about its behavior as $t \to \infty$: whatever this solution asymptotically approaches must be the unique periodic solution.

With the aid of a computer (we can express the solution using the variation of parameters form involving $e^{(t-s)A}$, but the resulting integral will be tedious to compute without a computer), we can find that the solution of this system with initial value $\mathbf{x}(0) = \mathbf{0}$ is

$$\mathbf{x}(t) = \begin{bmatrix} -\frac{8}{5}e^{-3t} - \frac{14}{5}e^{-2t} + \frac{6}{5}\cos t + \frac{1}{5}\sin t \\ 2e^{-3t} - \frac{14}{5}e^{-2t} + \frac{4}{5}\cos t + \frac{2}{5}\sin t \end{bmatrix}.$$

Asymptotically, this solution behaves like

$$\mathbf{x}_d(t) = \begin{bmatrix} \frac{6}{5}\cos t + \frac{1}{5}\sin t \\ \frac{4}{5}\cos t + \frac{2}{5}\sin t \end{bmatrix},$$

so this must be the periodic steady state of this system.

Lecture 27: Transition Matrices

TO BE FINISHED