MATH 27300 (Basic Theory of Ordinary Differential Equations) Notes

Steven Labalme

December 19, 2022

Weeks

List of Figures

Week 1

Introduction to ODEs

1.1 Definitions and Scope

9/28:
- Questions:
  - When will the PDFs be made available?
- Office: Eckhart 309.
  - Office hours: MWF 3:00-4:00.
- Reader: Walker Lewis. His contact info is in the syllabus.
- Final grade is based on...
  - 2 midterms (15 pts. each; weeks 4 and 8).
  - Final exam (35 pts.).
  - HW (35 pts.).
  - Bonus problems (15 pts).
- Total points for the quarter is 115. The bonus problems usually arise from advanced math and incorporate more advanced knowledge, and we are encouraged to seek out all relevant resources as long as we write up our own solutions.
- **Ordinary differential equation**: An equation that involves an unknown function of a single variable; an equation that takes the form $F(t,y,y',\dots,y^{(n)})=0$. *Also known as* **ODE**.
  - F is a known function.
  - t is an argument (time). x is also used (when space is involved).
  - $y=y(t)$ is an unknown function.
- **Order $\boldsymbol{n}$** (ODE): An ODE for which the n^{th} derivative of y is the highest-order derivative involved (and is involved).
- ODEs are of the form $y'=f(t,y)$ or, more generally, $y^{(n)}=F(t,y,y',\dots,y^{(n-1)})$.
  - We can transform this second form into the first form via

$$
Y=\begin{pmatrix} y\\ y'\\ \vdots\\ y^{(n-1)} \end{pmatrix} \qquad\qquad f(t,y)=\begin{pmatrix} Y^2\\ Y^3\\ \vdots\\ F(t,Y^1,Y^2,\dots,Y^{n-1}) \end{pmatrix}
$$

This makes $Y' = f(t, Y)$ equal to the system of equations

$$\begin{aligned}(Y^1)' &= Y^2\\ (Y^2)' &= Y^3\\ &\vdots\\ (Y^{n-1})' &= F(t, Y^1, Y^2, \dots, Y^{n-1})\end{aligned}$$

- Think about this conversion more.

– Thus, we mainly focus on equations of the form $y' = f(t, y)$ (where y may be a scalar or vector function), because that's general enough.

- **Linear** (ODE): Any ODE that can be written in the form

$$y' = A(t)y + f(t)$$

- Because of the above, this naturally includes equations of the form

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t)y = b(t)$$

– Indeed, if we define $Y = (y, y', \dots, y^{(n-1)})$, then we may express this equation in the form

$$\begin{aligned}\underbrace{\begin{pmatrix}Y^1\\Y^2\\\vdots\\Y^n\end{pmatrix}'}_{Y'} &= \underbrace{\begin{pmatrix}Y^2\\Y^3\\\vdots\\b(t) - a_0(t)Y^1 - \cdots - a_{n-1}(t)Y^{n-1}\end{pmatrix}}_{g(t,y)}\\
&= \begin{pmatrix}Y^2\\Y^3\\\vdots\\-a_0(t)Y^1 - \cdots - a_{n-1}(t)Y^{n-1}\end{pmatrix} + \underbrace{\begin{pmatrix}0\\0\\\vdots\\b(t)\end{pmatrix}}_{f(t)}\\
&= \underbrace{\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ -a_0(t) & -a_1(t) & -a_2(t) & \cdots & -a_{n-1}(t)\end{pmatrix}}_{A(t)} \underbrace{\begin{pmatrix}Y^1\\Y^2\\\vdots\\Y^n\end{pmatrix}}_{Y} + \underbrace{\begin{pmatrix}0\\0\\\vdots\\b(t)\end{pmatrix}}_{f(t)}\end{aligned}$$

– This conversion and its implications is covered in more depth in Lecture 4.1.

- **Nonlinear** (ODE): An ODE that is not linear.
- **Autonomous** (ODE): An ODE that can be written in the form

$$y' = f(y)$$

– Remember that y can be a scalar or a vector function.

– Solutions to autonomous ODEs can start at *any* time t and still be valid.

- For example, take the scalar ODE $y' = y$. It's general solution is $y(t) = ae^{t-t_0}$ for some $a \in \mathbb{R}$ and t_0 being the start time. Importantly, notice that we can make t_0 take any value we want and $y(t)$ will still solve $y' = y$.

- **Nonautonomous** (ODE): An ODE that is not autonomous.
  - We will not investigate these in this course.
- **Initial value problem**: A problem of the form, "find $y(t)$ such that the following holds."

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}$$

  Also known as **IVP**, **Cauchy problem**.
- Locally well-posed (LWP) conditions:
  1. Existence (local in time).
  2. Uniqueness (you cannot have multiple solutions).
  3. Local stability (if you perturb your initial value or equation a little bit, you do not expect your solution to vary crazily [esp. locally]).
- Example of a nonunique ODE:
  - $y' = \sqrt{y}$, $y(0) = 0$ has solutions $y_1(t) = 0$ $(t \geq 0)$ and $y_2(t) = t^2/4$ $(t \geq 0)$.
  - We will investigate the reason later.
- Preview of the reason: **Cauchy-Lipschitz Theorem** or **Picard-Lindelöf Theorem**.
  - As long as the ODE is **Lipschitz continuous**, it's locally stable.
- **Lipschitz continuous** (function): A function f such that

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|$$

  - But in the counterexample above, the slope of the chord from 0 to $y(t)$ approaches infinity as $t \to 0$.
- **Peano Existence Theorem**: Under certain conditions, there exists a solution to a given IVP.
- **Dynamical system**: A law under which a particle evolves over time. $y' = f(t, y)$, IVP is LWP.
- If the IVP $y' = f(t, y)$, $y(t_0) = y_0$ is locally well-posed, then the map $\Phi(t, x)$ which solves

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, x) = f(t, \Phi(t, x)) \\ \Phi(0, x) = x \end{cases}$$

  is well-defined and satisfies the property

$$\Phi(t_2, \Phi(t_1, x)) = \Phi(t_1 + t_2, x)$$

  - Φ is very related to y, though how exactly is still a bit of a mystery?? Perhaps it's

$$\Phi(t, x) = y(t)$$

    where y is the solution to the IVP $y' = f(t, y)$, $y(0) = x$.
  - It appears that $\Phi(t, x)$ is related to $f_t(x)$ from Guillemin and Haine (2018), i.e., we are picking a point x and traveling along its integral curve for time t.
  - Think about $y(t) = a\mathrm{e}^{t - t_0}$ as an integral curve of the one-dimensional vector field $X(x) = x$.
  - The final property appears to express the notion that if you have a system and evolve it by time t_1 and then time t_2, that's equivalent to evolving it by time $t_1 + t_2$.

- **Steady flow**: A vector field on a manifold contained in $\mathbb{R}^2$ or $\mathbb{R}^3$ that does not vary with time.
- Let X be a vector field.
  - Trajectory of a particle: At $x \in \Omega$, the velocity of the particle should coincide with $X(x)$.
  - The differential equation $\dot{x} = X(x)$ is what we're interested in.
  - A solid shape gets shifted and deformed (imagine a chunk of water falling out of the end of a pipe). This is the **local group of transformation**.
  - Differential geometry is the purview of such things.
- Newton's law of motion $F = m \cdot a$ applied to n particles is nothing but the system of equations

$$m_i x_i'' = F_i(x_1, \dots, x_n)$$

  for $i = 1, \dots, n$.
  - Many well-known examples.
  - The best known one perhaps is that of uniform acceleration of a single particle. In this case,

$$m_0 x'' = f_0$$

    - The solution is

$$x(t) = \frac{f_0}{2m_0} t^2 + v_0 t + x_0$$

      where $x_0 = x(0)$ and $v_0 = x'(0)$ are the initial conditions.
  - A simple example is downwards motion due to gravity. Then

$$x(t) = \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} t^2 + v_0 t + x_0$$

    - The trajectory in general is a parabola.
  - Another example: The mathematical pendulum.
    - The radial directions balance ($mg \cos\theta$).
    - The tangential directions do not ($mg \sin\theta$). Thus, our ODE is

$$l \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} = g \sin\theta$$

  - One last set of examples from ecology:
    - Imagine an petri dish of infinite nutrition. The population growth of the bacteria will obey the exponential growth law

$$\frac{\mathrm{d}y}{\mathrm{d}t} = ky$$

    - Suppose we have a system capacity M. Then we obey the logistic growth law

$$\frac{\mathrm{d}y}{\mathrm{d}t} = k(M - y)$$

    - Lotka-Volterra prey-predator model: Wolf population (W) and rabbit population (R). We have

$$\begin{aligned} R' &= k_1 R - aWR \\ W' &= -k_2 W + bWR \end{aligned}$$

    - We can also introduce more species and capacities and et cetera, et cetera.
- Conclusion: Dynamical systems are everywhere, especially in physics, chemistry, and ecology.
- We can also consider long-term behavior.
  - We can have chaos, but chaos can be reasoned with using oscillation, systems that converge to oscillation, etc. We will mostly be focusing on the regular aspect of the long-term behavior.

1.2 Origin of ODEs: Boundary Value Problems

9/30:
- Textbook PDFs will be posted today.
- Note: Equations of order n generally require n parameters to solve.
- Today, we will consider boundary value problems, which are separate from dynamical systems but not entirely unrelated.
- **Boundary Value Problem**: A problem in which we are solving for a y that has fixed values at the boundaries $x = a, b$. *Also known as* **BVP**.
- The **Brachistochrone problem** is an example of a BVP.
- **Brachistochrone problem**: Suppose you have a frictionless track from $(0, 0)$ to (a, y_0) and release a particle from $(0, 0)$. Which path allows the particle to get to (a, y_0) in the shortest amount of time? *Etymology* brákhistos "shortest" + khrónos "time."

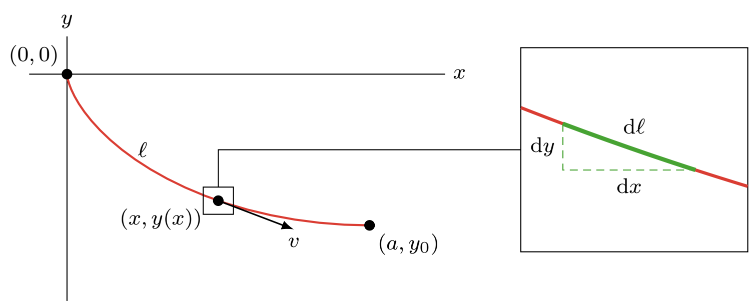


Figure 1.1: Brachistochrone problem.

  - Throughout this derivation, we will make several assumptions. We will do our best to note these assumptions as we go in footnotes. Note that while all of these assumptions are justified in the case of solving this problem, they may not be justified in every related variational problem. Let's begin.
  - Since the track is frictionless, the mechanical energy should be conserved.
  - At a given point along the curve, the particle has a velocity v and is vertical distance y from where it started. We know from physics that

$$\begin{aligned}\frac{1}{2}mv^2 &= mgy\\ v &= \sqrt{2gy}\end{aligned}$$

  - Since $v = \mathrm{d}\ell/\mathrm{d}t$, the time $\mathrm{d}t$ it takes for the particle to traverse an infinitesimal section of track of arc length $\mathrm{d}\ell$ is $\mathrm{d}t = \mathrm{d}\ell/v$.
  - The track should be given by $y = y(x)$[1].
  - Let ℓ denote the arc length of the whole track. Then

$$\mathrm{d}\ell = \sqrt{1 + (y'(x))^2}\,\mathrm{d}x$$

  - Thus, the total time for the particle to traverse the curve is

$$t(y) = \int_0^t \mathrm{d}\tau = \int_0^a \frac{\mathrm{d}\ell}{v} = \int_0^a \frac{\sqrt{1 + (y'(x))^2}\,\mathrm{d}x}{\sqrt{2gy(x)}}$$

[1]There are paths that connect $(0, 0)$ and (a, y_0) that are not functions of x. We are taking those out of consideration.

- We also have $y(0) = 0$ and $y(a) = y_0$.
- We want to find y such that the above integral is minimized. Thus, we define the following **functional**, which is used to solve general fixed-endpoint variational problems (the Brachistochrone problem is a problem of this type).
- Let $J[y] = \int_a^b F(x, y(x), y'(x))\,\mathrm{d}x$.
- The space of functions we're considering is C^1 (the set of all continuously differentiable functions)[2].
- Take a function h, vanishing at a, b.
- Let $f(t) = J[y + th]$. Then

$$f(t) = \int_a^b F(x, \underbrace{y(x) + th(x)}_{z(x,t)}, \underbrace{y'(x) + th'(x)}_{w(x,t)})\,\mathrm{d}x$$

- We know that[3]

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b F\,\mathrm{d}x &= \int_a^b \frac{\mathrm{d}F}{\mathrm{d}t}\,\mathrm{d}x \\ &= \int_a^b \left(\frac{\partial F}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial F}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} + \frac{\partial F}{\partial w}\frac{\mathrm{d}w}{\mathrm{d}t}\right)\mathrm{d}x \\ &= \int_a^b \left(\frac{\partial F}{\partial x}\cdot 0 + \frac{\partial F}{\partial z}\cdot h(x) + \frac{\partial F}{\partial w}\cdot h'(x)\right)\mathrm{d}x \\ &= \int_a^b \left(\frac{\partial F}{\partial z}\cdot h(x) + \frac{\partial F}{\partial w}\cdot h'(x)\right)\mathrm{d}x\end{aligned}$$

  - The last term in the above equation may be integrated by parts as follows. Note that we make use of the hypothesis $h(a) = h(b) = 0$ in eliminating the $[uv]_a^b$ term.

$$\begin{aligned}\int_a^b \frac{\partial F}{\partial w}h'(x)\,\mathrm{d}x &= \left[\frac{\partial F}{\partial w}h(x)\right]_{x=a}^b - \int_a^b h(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)\mathrm{d}x \\ &= \left[\left.\frac{\partial F}{\partial w}\right|_b\cdot 0 - \left.\frac{\partial F}{\partial w}\right|_a\cdot 0\right] - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)h(x)\,\mathrm{d}x \\ &= -\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)h(x)\,\mathrm{d}x\end{aligned}$$

  - Substituting back into the original equation, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b F\,\mathrm{d}x &= \int_a^b \left[\frac{\partial F}{\partial z}\cdot h(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)h(x)\right]\mathrm{d}x \\ &= \int_a^b \left[\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)\right]h(x)\,\mathrm{d}x\end{aligned}$$

  - Therefore,

$$f'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int_a^b F\,\mathrm{d}x = \int_a^b \left\{\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}\right]\right\}h(x)\,\mathrm{d}x$$

- Thus,

$$f'(0) = \int_a^b \left\{\frac{\partial F}{\partial z}(x, y(x), y'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}(x, y(x), y'(x))\right]\right\}h(x)\,\mathrm{d}x = 0$$

  for all h.

[2] This also eliminates some possible paths from consideration.

[3] We must assume sufficient regularity of F here. In particular, we must assume that the derivative of the integral of F is equal to the integral of the derivative of F.

- Now suppose y is the solution. Then y minimizes $J[y]$. But if this is true, then any variation th will cause $J[y+th] > J[y]$. It follows that for every h, $f(t)$ has a minium at $t=0$. But if f has a minimum at 0 for all h, then $f'(0) = 0$ for all h.

– Lemma: Let ϕ be continuous on (a,b). If for every $h \in C^1([a,b])$ vanishing on a,b we have that

$$\int_a^b \phi(x)h(x)\,\mathrm{d}x = 0$$

then $\phi(x) = 0$.

Proof. Suppose for the sake of contradiction that (WLOG) $\phi(x_0) > 0$. Then within some neighborhood $N_\delta(x)$ of x_0, $\phi(x) > 0$ for all $x \in N_\delta(x)$. Now choose h to be a bump function on that interval. Then $\int_a^b \phi(x)h(x)\,\mathrm{d}x > 0$, a contradiction. □

– It follows that

$$\frac{\partial F}{\partial z}(x,y(x),y'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}(x,y(x),y'(x))\right] = 0$$

- This is a second-order differential equation, specifically the **Euler-Lagrange equation**.
- It is a necessary condition for y to be an extrema.
- Euler-Lagrange equations are not easy to solve in general. However, we're lucky here.

– In our example,

$$F(x,z,w) = \sqrt{\frac{1+w^2}{2gz}}$$

- What's nice here is that $F(x,z,w) = F(z,w)$, i.e., there is no dependence on x. This is crucial.

– With this observation in mind, notice that

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}x} &= \frac{\partial F}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}x} + \frac{\partial F}{\partial w}\frac{\mathrm{d}w}{\mathrm{d}x}\\ &= \frac{\partial F}{\partial z}\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\partial F}{\partial w}\frac{\mathrm{d}y'}{\mathrm{d}x}\\ &= \frac{\partial F}{\partial z}\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\partial F}{\partial w}\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\end{aligned}$$

– We now rearrange the E-L equation and multiply through by $\mathrm{d}y/\mathrm{d}x$.

$$\begin{aligned}\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right) &= 0\\ \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial w}\right)\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\partial F}{\partial z}\frac{\mathrm{d}y}{\mathrm{d}x}\end{aligned}$$

– Subtracting the last two results yields

$$\begin{aligned}\frac{\mathrm{d}F}{\mathrm{d}x} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}F}{\mathrm{d}w}\right)\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\partial F}{\partial w}\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\\ \frac{\mathrm{d}F}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}F}{\mathrm{d}w}\right)\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\partial F}{\partial w}\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}F}{\mathrm{d}w}\frac{\mathrm{d}y}{\mathrm{d}x}\right)\\ \frac{\mathrm{d}}{\mathrm{d}x}\left(F - \frac{\mathrm{d}F}{\mathrm{d}w}\frac{\mathrm{d}y}{\mathrm{d}x}\right) &= 0\\ F - \frac{\mathrm{d}F}{\mathrm{d}w}\frac{\mathrm{d}y}{\mathrm{d}x} &= A\end{aligned}$$

where $A \in \mathbb{R}$ depends on the initial conditions.

- From the definition of F, we can calculate

$$\frac{\partial F}{\partial w}=\frac{w}{\sqrt{1+w^2}}\cdot\frac{1}{\sqrt{2gz}}=\frac{y'}{\sqrt{1+(y')^2}}\cdot\frac{1}{\sqrt{2gy}}$$

- It follows that our solution function y satisfies the separable differential equation

$$\begin{aligned}\sqrt{\frac{1+(y')^2}{2gy}}-\frac{y'}{\sqrt{1+(y')^2}}\cdot\frac{1}{\sqrt{2gy}}\cdot y'&=A\\ \frac{1+(y')^2}{\sqrt{1+(y')^2}\sqrt{2gy}}-\frac{(y')^2}{\sqrt{1+(y')^2}\sqrt{2gy}}&=A\\ \frac{1}{\sqrt{2gy(1+(y')^2)}}&=A\\ (y')^2&=\frac{1/2A^2g-y}{y}\end{aligned}$$

- The solution, as we can determine using methods from Calculus I-II, is the **cycloid**

$$\begin{cases}x=a(\theta-\sin\theta)\\ y=a(1-\cos\theta)\end{cases}$$

where the specific parameters come from the boundary values.

- **Functional**: A map from a function space to a set of numbers.
- **Sturm-Liouville problems**: Boundary value problems concerning the integral

$$\int_a^b\left[p(x)(y'(x))^2+q(x)(y(x))^2\right]\mathrm{d}x$$

  - The most basic BVP is a vibrating string. In finding the eigenmode of the vibration, you need to solve the above differential equation.
  - Very important in physics.
  - If time permits at the end of the course, Shao will return to the following topic in detail.
- Next several weeks: *Solvable* differential equations.

1.3 Chapter 1: Introduction

From Teschl (2012).

Section 1.1: Newton's Equations

11/11:
- Before we begin defining abstract terms, let's look at an example in which many of these terms arise.
- Investigation: Describing the motion of a particle using classical mechanics.
  - The location, velocity, and acceleration of a particle are typically given by the related functions[4]

$$x:\mathbb{R}\to\mathbb{R}^3\qquad v=\dot{x}:\mathbb{R}\to\mathbb{R}^3\qquad a=\dot{v}:\mathbb{R}\to\mathbb{R}^3$$

[4]Newton's notation for derivatives uses a number of dots above the dependent variable to indicate the order of the derivative to which we are referring. For example, $\dot{x}$ denotes the first derivative of x with respect to it's independent variable, and $\ddot{x}$ denotes the second derivative of x with respect to it's independent variable.

- The particle does not move randomly, though; its motion is governed by an external force field $F : \mathbb{R}^3 \to \mathbb{R}^3$ which exerts a vector force $F(x)$ on the particle when it is at $x \in \mathbb{R}^3$.
- Additionally, Newton's second law of motion asserts that at every $x \in \mathbb{R}^3$, the force acting on the particle must equal the acceleration of the particle times its mass, that is,

$$m\ddot{x}(t) = F(x(t))$$

for all $t \in \mathbb{R}$.
- Given a force field F, physicists often seek to determine how bodies evolve under F over time. Mathematically, they seek functions $x(t)$ that satisfy $m\ddot{x}(t) = F(x(t))$ for a given F.
- Consider the example of a stone falling toward the Earth under gravity from class.
  - In the vicinity of the surface of Earth, the gravitational force is approximately constant and given by

$$F(x) = -mg\begin{pmatrix}0\\0\\1\end{pmatrix}$$

where g is the positive **gravitational constant** and the $+x_3$ direction is taken to be normal to the Earth's surface.
  - Hence, the system of differential equations reads

$$m\ddot{x}_1 = 0 \qquad m\ddot{x}_2 = 0 \qquad m\ddot{x}_3 = -mg$$

  - The first equation can be integrated with respect to t twice, yielding $x_1(t) = C_2 + C_1t$. Computing $x_1(0)$ and $\dot{x}_1(0)$ shows that $C_2 = x_1(0)$ and $C_1 = v_1(0)$. An analogous result holds for the second equation. For the third equation, we get $x_3(t) = C_2 + C_1t - \frac{1}{2}mgt^2$ where $C_2 = x_3(0)$ and $C_1 = v_1(0)$, again. Thus, the full solution reads

$$x(t) = x(0) + v(0)t - \frac{g}{2}\begin{pmatrix}0\\0\\1\end{pmatrix}t^2$$

- **Differential equation**: A relation between a function $x(t)$ and its derivatives.
  - The equation $m\ddot{x}(t) = F(x(t))$, above, is an example of a differential equation.
- **Second-order** (differential equation): A differential equation in which the highest derivative is of second degree.
  - The equation $m\ddot{x}(t) = F(x(t))$, above, is an example of a second-order differential equation.
- **n^{th}-order** (differential equation): A differential equation in which the highest derivative is of degree n.
- **System** (of differential equations): A finite set of differential equations.
  - Systems of differential equations vary in how related they are, that is, they may or may not share variables.
  - Technically, $m\ddot{x}(t) = F(x(t))$ is a system of differential equations since $x = (x_1, x_2, x_3)$ gives rise to three independent differential equations, one for each Cartesian dimension, as follows.

$$m\ddot{x}_1 = F(x_1(t)) \qquad m\ddot{x}_2 = F(x_2(t)) \qquad m\ddot{x}_3 = F(x_3(t))$$

- **Dependent** (variable): A variable whose value depends on that of another.
  - In our example, x, v, and a are all dependent variables.

- **Independent** (variable): A variable whose value does not depend on that of another.
  - In our example, t is the independent variable.
- **First-order** (system): A system of differential equations in which the differential equation of highest order is first order.
  - One possible rewrite of the system $m\ddot{x}(t) = F(x(t))$ works by increasing the number of dependent variables from $x \in \mathbb{R}^3$ to $(x, v) \in \mathbb{R}^6$. Indeed, we can rewrite the second-order differential equation $m\ddot{x}(t) = F(x(t))$ as the following first-order system.

$$\begin{aligned}\dot{x}(t) &= v(t)\\ \dot{v}(t) &= \frac{1}{m}F(x(t))\end{aligned}$$

  - We will find that the above form is often better suited to theoretical investigations.
- $\boldsymbol{n}^{\textbf{th}}$**-order** (system): A system of differential equations in which the differential equation of highest order is n^{th} order.
- One conclusion of our investigation of gravity's effect on a particle is that the entire fate (past and future) of our particle's position, velocity, and acceleration (under simple gravity) is uniquely determined by its initial location $x(0)$ and velocity $v(0)$.
  - While we could use simple integration to solve this system, we cannot always do this (not even under Newtonian universal gravitation).

Problems

11/24: **1.1.** Consider the case of a stone dropped from a height h above the Earth's surface. Denote by r the distance of the stone from the Earth's surface. The initial condition reads $r(0) = h$, $\dot{r}(0) = 0$. The equation of motion reads

$$\ddot{r} = -\frac{\gamma M}{(R+r)^2}$$

for the exact model and

$$\ddot{r} = -g$$

for the approximate model, where R is the radius of the Earth, M is the mass of the Earth, and $g = \gamma M/R^2$. Note that $\ddot{r}$ is acceleration, not force, and hence we do not see the mass of the stone in the RHS's above.

(i) Transform both equations into a first-order system.

Proof. Define $v = \dot{r}$. Then the first-order system corresponding to the exact model is

$$\boxed{\begin{aligned}\dot{r} &= v\\ \dot{v} &= -\frac{\gamma M}{(R+r)^2}\end{aligned}}$$

and the first-order system corresponding to the approximate model is

$$\boxed{\begin{aligned}\dot{r} &= v\\ \dot{v} &= -g\end{aligned}}$$

□

(ii) Compute the solution to the approximate system corresponding to the given initial condition. Compute the time it takes for the stone to hit the surface ($r = 0$).

Proof. We can integrate $\dot{v} = -g$ to determine that $v(t) = -gt + C_1$ for some $C_1 \in \mathbb{R}$ and then integrate $\dot{r} = v$ to determine that

$$r(t) = -\frac{1}{2}gt^2 + C_1 t + C_2$$

for some additional $C_2 \in \mathbb{R}$. Using the initial conditions, we can determine that

$$h = r(0) = C_2 \qquad\qquad 0 = \dot{r}(0) = v(0) = C_1$$

Thus, the solution to the approximate system corresponding to the given initial conditions is

$$\boxed{r(t) = -\frac{1}{2}gt^2 + h}$$

We can solve the equation $r(t) = 0$ for t as follows.

$$\begin{aligned} 0 &= r(t) \\ &= -\frac{1}{2}gt^2 + h \\ t &= \pm\sqrt{\frac{2h}{g}} \end{aligned}$$

Knowing that $t \geq 0$ by definition, we choose

$$\boxed{t = \sqrt{\frac{2h}{g}}}$$

□

(iii) Assume that the exact equation also has a unique solution corresponding to the given initial condition. What can you say about the time it takes for the stone to hit the surface in comparison to the approximate model? Will it be longer or shorter? Estimate the difference between the solutions in the exact and in the approximate case. *Hints*: You should not compute the solution to the exact equation! Look at the minimum and maximum of the force.

Proof. It will take the stone $\boxed{\text{longer}}$ to hit the surface in the exact model than in the approximate model. How do we *estimate* the difference?? □

(iv) Grab your physics book from high school and give numerical values for the case $h = 10\,\text{m}$.

Proof. For the approximate model,

$$\boxed{t \approx 1.43\,\text{s}}$$

and for the exact model, ... □

1.2. Consider again the exact model from the previous problem and write

$$\ddot{r} = -\frac{\gamma M \varepsilon^2}{(1 + \varepsilon r)^2}$$

where $\varepsilon = 1/R$. It can be shown that the solution $r(t) = r(t, \varepsilon)$ to the above with the given initial conditions is C^∞ with respect to both t, ε. Show that

$$r(t) = h - g\left(1 - \frac{2h}{R}\right)\frac{t^2}{2} + O\left(\frac{1}{R^4}\right)$$

where $g = \gamma M/R^2$. *Hint*: Insert $r(t, \varepsilon) = r_0(t) + r_1(t)\varepsilon + r_2(t)\varepsilon^2 + r_3(t)\varepsilon^3 + O(\varepsilon^4)$ into the differential equation and collect powers of ε. Then solve the corresponding differential equations for $r_0(t), r_1(t), \ldots$ and note that the initial conditions follow from $r(0, \varepsilon) = h$ and $\dot{r}(0, \varepsilon) = 0$. A rigorous justification for this procedure will be given in Section 2.5.

Proof. Taking the hint, we get that

$$r_0''(t) + r_1''(t)\varepsilon + r_2''(t)\varepsilon^2 + r_3''(t)\varepsilon^3 + O(\varepsilon^4) = -\frac{\gamma M\varepsilon^2}{(1+\varepsilon(r_0(t) + r_1(t)\varepsilon + r_2(t)\varepsilon^2 + r_3(t)\varepsilon^3 + O(\varepsilon^4)))^2}$$

We can rewrite the denominator d as follows.

$$\begin{aligned}
d &= (1 + r_0(t)\varepsilon + r_1(t)\varepsilon^2 + r_2(t)\varepsilon^3 + O(\varepsilon^4))^2 \\
&= 1 + r_0(t)\varepsilon + r_1(t)\varepsilon^2 + r_2(t)\varepsilon^3 + r_0(t)\varepsilon + r_0(t)^2\varepsilon^2 + r_0(t)r_1(t)\varepsilon^3 \\
&\quad + r_1(t)\varepsilon^2 + r_0(t)r_1(t)\varepsilon^3 + r_2(t)\varepsilon^3 + O(\varepsilon^4) \\
&= 1 + 2r_0(t)\varepsilon + (r_0(t)^2 + 2r_1(t))\varepsilon^2 + 2(r_0(t)r_1(t) + r_2(t))\varepsilon^3 + O(\varepsilon^4)
\end{aligned}$$

Multiplying both sides of the original equation by d yields

$$\begin{aligned}
-\gamma M\varepsilon^2 &= r_0'' + 2r_0r_0''\varepsilon + (r_0^2 + 2r_1)r_0''\varepsilon^2 + 2(r_0r_1 + r_2)r_0''\varepsilon^3 \\
&\quad + r_1''\varepsilon + 2r_0r_1''\varepsilon^2 + (r_0^2 + 2r_1)r_1''\varepsilon^3 \\
&\quad + r_2''\varepsilon^2 + 2r_0r_2''\varepsilon^3 \\
&\quad + r_3''\varepsilon^3 + O(\varepsilon^4) \\
0 + 0\varepsilon + 0\varepsilon^2 + 0\varepsilon^3 &= r_0'' \\
&\quad + (2r_0r_0'' + r_1'')\varepsilon \\
&\quad + [\gamma M + (r_0^2 + 2r_1)r_0'' + 2r_0r_1'' + r_2'']\varepsilon^2 \\
&\quad + [2(r_0r_1 + r_2)r_0'' + (r_0^2 + 2r_1)r_1'' + 2r_0r_2'' + r_3'']\varepsilon^3
\end{aligned}$$

Thus, by collecting powers of ε and comparing, we have the system of differential equations

$$\begin{aligned}
r_0''(t) &= 0 \\
r_1''(t) &= -2r_0(t)r_0''(t) \\
r_2''(t) &= -\gamma M - (r_0(t)^2 + 2r_1(t))r_0''(t) - 2r_0(t)r_1''(t) \\
r_3''(t) &= -2(r_0(t)r_1(t) + r_2(t))r_0''(t) - (r_0(t)^2 + 2r_1(t))r_1''(t) - 2r_0(t)r_2''(t)
\end{aligned}$$

which we can sequentially solve as follows.

Before we tackle the differential equations, however, a quick note on the initial conditions. We need

$$\begin{aligned}
h &= r(0,\varepsilon) & 0 &= \dot{r}(0,\varepsilon) \\
&= r_0(0) + r_1(0)\varepsilon + r_2(0)\varepsilon^2 + r_3(0)\varepsilon^3 & &= r_0'(0) + r_1'(0)\varepsilon + r_2'(0)\varepsilon^2 + r_3'(0)\varepsilon^3
\end{aligned}$$

But the only way these equations can be satisfied for all ε is if

$$r_0(0) = h \qquad r_1(0) = \cdots = r_3(0) = r_0'(0) = \cdots = r_3'(0) = 0$$

in agreement with the unjustified method from class!

The first differential equation has general solution

$$r_0(t) = C_1t + C_2$$

Applying the initial conditions, we can learn that $C_1 = 0$ and $C_2 = h$. This yields the solution

$$r_0(t) = h$$

We can now move onto the second differential equation, which in light of the formula for r_0 can be rewritten as

$$r_1''(t) = -2 \cdot h \cdot 0 = 0$$

Solving and applying the initial conditions here yields

$$r_1(t) = 0$$

Moving onto the third differential in the same fashion, we have

$$r_2''(t) = -\gamma M - (h^2 + 2 \cdot 0) \cdot 0 - 2 \cdot h \cdot 0 = -\gamma M$$

Thus,

$$r_2(t) = -\frac{\gamma M}{2} t^2$$

For the fourth and final differential equation, we have

$$r_3''(t) = -2\left(h \cdot 0 - \frac{\gamma M}{2} t^2\right) \cdot 0 - (h^2 + 2 \cdot 0) \cdot 0 - 2 \cdot h \cdot -\gamma M = 2h\gamma M$$

Thus,

$$r_3(t) = h\gamma M t^2$$

Combining the last four results, we have that

$$\begin{aligned}
r(t) &= r(t, \varepsilon) \\
&= r_0(t) + r_1(t)\varepsilon + r_2(t)\varepsilon^2 + r_3(t)\varepsilon^3 + O(\varepsilon^4) \\
&= h + 0 \cdot \varepsilon + \left(-\frac{\gamma M}{2} t^2\right) \cdot \varepsilon^2 + h\gamma M t^2 \varepsilon^3 + O(\varepsilon^4) \\
&= h - \gamma M \varepsilon^2 (1 - 2h\varepsilon) \frac{t^2}{2} + O(\varepsilon^4) \\
&= h - \underbrace{\frac{\gamma M}{R^2}}_{g} \left(1 - \frac{2h}{R}\right) \frac{t^2}{2} + O\left(\frac{1}{R^4}\right)
\end{aligned}$$

as desired. ☐

Section 1.2: Classification of Differential Equations

11/11:

- $\boldsymbol{C^k(U,V)}$: The set of functions $f : U \to V$ having continuous derivatives up to order k, where $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$, and $k \in \mathbb{N}_0$.
- $\boldsymbol{C(U,V)}$: The set of continuous functions $f : U \to V$. *Given by*
$$C(U, V) = C^0(U, V)$$
- $\boldsymbol{C^\infty(U,V)}$: The set of smooth functions $f : U \to V$. *Given by*
$$C^\infty(U, V) = \bigcap_{k \in \mathbb{N}} C^k(U, V)$$
- $\boldsymbol{C^k(U)}$: The set of real functions $f : U \to \mathbb{R}$ having continuous derivatives up to order k. *Given by*
$$C^k(U) = C^k(U, \mathbb{R})$$
- **Ordinary differential equation**: An equation of the form
$$F(t, x, x^{(1)}, \dots, x^{(k)}) = 0$$
where $F \in C(U)$ ($U \subseteq \mathbb{R}^{k+2}$ open) relates the unknown function $x \in C^k(J)$ ($J \subseteq \mathbb{R}$), its independent variable t, and its first k derivatives;
$$x^{(j)} := \frac{\mathrm{d}^j x}{\mathrm{d}t^j}$$
for $j \in \mathbb{N}_0$. *Also known as* **ODE**.

- **Order** (of an ODE): The highest derivative appearing in the argument of F.
- **Solution** (of an ODE): A function $\phi \in C^k(I)$ ($I \subseteq J$ an interval) satisfying the equation
$$F(t, \phi(t), \phi^{(1)}(t), \dots, \phi^{(k)}(t)) = 0$$
for all $t \in I$.
- Very little can be said about completely general ODEs.
- Thus, we begin our investigation with the subclass of ODEs that can solved for their highest order derivative, that is, with ODEs of the form
$$x^{(k)} = f(t, x, x^{(1)}, \dots, x^{(k-1)})$$
  - Relation between these ODEs and general ODEs: These ODEs are the ones for which F has nonzero partial derivative with respect to y_k. Indeed, if F satisfies this condition locally near $(x, y_k) \in U$, then the implicit function theorem permits the above rearrangement.
- **System** (of ODEs): A finite set of ODEs. *Given by*
$$\begin{aligned} x_1^{(k)} &= f_1(t, x, x^{(1)}, \dots, x^{(k-1)}) \\ &\vdots \\ x_n^{(k)} &= f_n(t, x, x^{(1)}, \dots, x^{(k-1)}) \end{aligned}$$
  - Note that the use of an unsubscripted $x^{(j)}$ in the argument of each f_i reflects the fact that every f_i is a function of *all* of the components of of x and their derivatives (not just x_i and its derivatives). Symbolically, each f_i is a function of $x_l^{(j)}$ ($l = 1, \dots, n$ and $j = 0, \dots, k-1$), and
$$f_i : \mathbb{R} \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k \text{ times}} \to \mathbb{R}$$
- **Linear** (system): A system of n ordinary differential equations for which each $x_i^{(k)}$ is of the form
$$x_i^{(k)} = g_i(t) + \sum_{l=1}^{n} \sum_{j=0}^{k-1} f_{i,j,l}(t) x_l^{(j)}$$
  - The summations are over the n components of x, and each of their k derivatives up from the function itself ($l = 0$) to $l = k - 1$. In other words, if a (derivative of) a component of x appears in a linear system, the only modification to it from itself should be a functional coefficient in the independent variable.
- **Homogeneous** (linear system): A linear system for which $g_i(t) = 0$.
- **Inhomogeneous** (linear system): A linear system for which $g_i(t) = 0$.
- Teschl (2012) goes over the conversion of a system to a first-order system, as covered in class.
  - Also noted: We can include t as a dependent variables by taking $z = (t, y)$ and making $\dot{z}_1 = 1$, $\dot{z}_i = z_{i+1}$ ($i = 2, \dots, k$), and $\dot{z}_{k+1} = f(z)$.
- **Autonomous** (system): A system in which f does not depend on t.
- We will often limit our studies to autonomous first-order ODEs since, as per the past two results, these encapsulate all higher order, potentially non-autonomous systems as well.
- **Partial differential equation**: A differential equation for which $t \in \mathbb{R}^m$. *Also known as* **PDE**.
  - Name justification: $t \in \mathbb{R}^m$ necessitates the use of partial derivatives.
- Complex values will not be considered until later.

Week 2

Solving Simple ODEs

2.1 Separable ODEs

10/3:

- Do not sit on the left side of the classroom: The sun sucks!
- **Separable** (ODE): An ODE of the form

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t)g(y)$$

  where y is a real[1], unknown, scalar function of t.
- Solving separable ODEs: Formally, evaluate

$$\int \frac{\mathrm{d}y}{g(y)} = \int f(t)\,\mathrm{d}t$$

- Rearrange the initial separable ODE to $\mathrm{d}y/\mathrm{d}t \cdot 1/g = f$ and invoke the law of composite differentiation to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{y_0}^{y(t)} \frac{\mathrm{d}w}{g(w)} - \int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau\right] = 0$$

- It follows that

$$\int_{y_0}^{y(t)} \frac{\mathrm{d}w}{g(w)} = \int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau$$

- Examples:
  1. Exponential growth.
     - We have that

$$\frac{\mathrm{d}y}{\mathrm{d}t} = ky$$

       for $k > 0$ and $y(0) = y_0 > 0$.
     - The solution is

$$\begin{aligned}
\frac{1}{y} \cdot \frac{\mathrm{d}y}{\mathrm{d}t} &= k\\
\log y(t) - \log y_0 &= kt\\
y(t) &= y_0 \mathrm{e}^{kt}
\end{aligned}$$

[1] We'll deal with complex functions later.

2. Logistic growth.
   - We have that
     $$\frac{\mathrm{d}y}{\mathrm{d}t} = ky\left(1 - \frac{y}{M}\right)$$
     for $k, M > 0$ and $y(0) = y_0 > 0$.
   - The solution is
     $$\begin{aligned}
     \frac{M\,\mathrm{d}y}{y(M-y)} &= k\,\mathrm{d}t \\
     \log\frac{y}{M-y} - \log\frac{y_0}{M-y_0} &= kt \\
     \frac{y(M-y_0)}{y_0(M-y)} &= \mathrm{e}^{kt} \\
     y \cdot \frac{M-y_0}{y_0} &= (M-y)\mathrm{e}^{kt} \\
     y \cdot \frac{M-y_0}{y_0} + y\mathrm{e}^{kt} &= M\mathrm{e}^{kt} \\
     y\left(\frac{M-y_0}{y_0} + \mathrm{e}^{kt}\right) &= M\mathrm{e}^{kt} \\
     y\left(\frac{M-y_0+y_0\mathrm{e}^{kt}}{y_0}\right) &= M\mathrm{e}^{kt} \\
     y\left(\frac{M+y_0(\mathrm{e}^{kt}-1)}{y_0}\right) &= M\mathrm{e}^{kt} \\
     y(t) &= \frac{My_0\mathrm{e}^{kt}}{M+y_0(\mathrm{e}^{kt}-1)}
     \end{aligned}$$
   - Sketches the graph of logistic growth and discusses the turning point (for which there is a formula; zero of the second derivative) as well as general trends.
   - If $y_0 < 0$, the solution is not physically meaningful, but it is mathematically insightful.
     - When we integrate, the arguments of our logarithms now have absolute values.
       $$\log\left|\frac{y}{M-y}\right| - \log\left|\frac{y_0}{M-y_0}\right| = kt$$
     - We need to make sure that the denominator of the final logistic form is never equal to zero, but now that y_0 is negative, as t increases, the denominator will approach zero exponentially. It reaches zero when
       $$\begin{aligned}
       M + y_0(\mathrm{e}^{kt}-1) &= 0 \\
       \mathrm{e}^{kt} &= -\frac{M}{y_0} + 1
       \end{aligned}$$
       In other words, $t_{\max} = (1/k)\log(1 - M/y_0)$ because when $t = t_{\max}$, the equation blows up.
     - This is an example of **finite lifespan**.
   - If $y_0 > M$, then you will exponentially decrease to M.
3. Lotka-Volterra predator-prey model.
   - We have that
     $$r' = k_1 r - awr \qquad\qquad w' = -k_2 w + bwr$$
     where r is rabbits and w is wolves.

- We can rename the variables to

$$\begin{cases} x' = Ax - Bxy \\ y' = -Cy + Dxy \end{cases}$$

- Dividing, we get

$$\frac{x'}{y'} = \frac{Ax - Bxy}{-Cy + Dxy}$$
$$\frac{By - A}{y} y' + \frac{Dx - C}{x} x' = 0$$

- Use the fact that x, y are independent variables, so both terms in the above equation are equal to zero??
- Invoke the law of composite differentiation twice and, from the above, know that $0 + 0 = 0$, so we can add the two solutions:

$$\frac{\mathrm{d}}{\mathrm{d}t}(By(t) - A\log y(t)) + \frac{\mathrm{d}}{\mathrm{d}t}(Dx(t) - C\log x(t)) = 0$$
$$By(t) - A\log y(t) + Dx(t) - C\log x(t) = E$$

- Sketches some of the trajectories (they're all closed curves in the xy-plane).

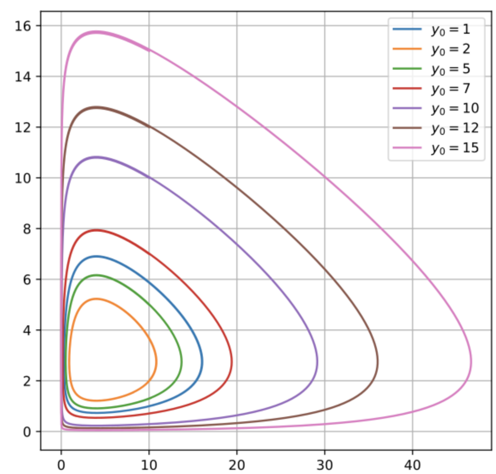


Figure 2.1: Lotka-Volterra solution curves.

- Properties of the curves:
  - The implicit relation which determines them: By the implicit function theorem, the y derivative of the LHS is $B - A/y$ and the x-derivative of the LHS is $D - C/x$. When the partial derivatives are equal to zero, $(C/D, A/B)$ becomes interesting. Turning points happen when the y-coordinate is A/B or the x-coordinate is C/D.

- **Finite lifespan**: Even if the RHS of $\mathrm{d}y/\mathrm{d}t = f(t, y)$ is very regular, the solution can still blow up at some finite time.

- Consider the final ODE from the Brachistochrone problem.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{\frac{B - y}{y}}$$

  - Finding the **primitives**.

    - What are these "primitives" Shao keeps talking about??
  - We should have

$$\int \sqrt{\frac{y}{B-y}}\,\mathrm{d}y = x$$

  - Change of variables: $y = B\sin^2\phi$ and $\mathrm{d}y = 2B\cos\phi\sin\phi\,\mathrm{d}\phi$. Thus,

$$\int \sqrt{\frac{y}{B-y}}\,\mathrm{d}y = \int \frac{\sin\phi}{\cos\phi}\cdot 2B\cos\phi\sin\phi\,\mathrm{d}\phi = 2B\int \sin^2\phi\,\mathrm{d}\phi$$

  - The solution is

$$\begin{cases} x = B\phi - \frac{B}{2}\sin(2\phi) + C \\ y = B\sin^2\phi \end{cases}$$

    - This is a parameterization of a cycloid.

- Later in the week, we will do the SHM, the pendulum, the Kepler 2-body problem, and the Michaelis-Menten equation.
- Separable ODEs are a subset of ODEs of **exact form**.
- ODEs of exact form are of the form

$$g(x,y)\frac{\mathrm{d}y}{\mathrm{d}x} + f(x,y) = 0$$

  where for some $F(x,y)$, $g = \partial F/\partial y$, $f = \partial F/\partial x$, and partials commute. Equivalently,

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}$$

  is our necessary and sufficient condition.
- By the law of composite differentiation,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left[F(x,y(x))\right] &= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\cdot y'(x) \\ &= f(x,y(x)) + g(x,y(x))y'(x) \\ &= 0\end{aligned}$$

  - We solve these with an integrating factor $\mu \neq 0$ such that $(\mu g, \mu f)$ satisfy the constraint.

2.2 Office Hours (Shao)

- **Primitive**: An antiderivative.
- **Law of composite differentiation**: The chain rule.
- Went over how Shao has been applying the law of composite differentiation with respect to separable ODEs:
  - Rearrange the initial separable ODE as follows.

$$\frac{1}{g(y)}\cdot\frac{\mathrm{d}y}{\mathrm{d}t} = f(t)$$

- Define $\mathrm{d}H/\mathrm{d}y = 1/g(y)$. Then, continuing from the above, we have by the law of composite differentiation that

$$
\begin{aligned}
\frac{\mathrm{d}H}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}t} &= f(t) \\
\frac{\mathrm{d}H}{\mathrm{d}t} &= f(t)
\end{aligned}
$$

- From the definition of H, we know that $H(y) = \int_{y_0}^{y} \mathrm{d}w \,/g(w)$. We also have from the FTC that $f(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau$. Thus, continuing from the above, we have that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(H) &= f(t) \\
\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{y_0}^{y} \frac{\mathrm{d}w}{g(w)}\right] &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau \\
\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{y_0}^{y(t)} \frac{\mathrm{d}w}{g(w)} - \int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau\right] &= 0
\end{aligned}
$$

  as desired.
- It follows since $y(t_0) = y_0$ that $C = H(y_0) = 0$, so we can take the above to

$$
\int_{y_0}^{y(t)} \frac{\mathrm{d}w}{g(w)} = \int_{t_0}^{t} f(\tau)\,\mathrm{d}\tau
$$

  knowing that our constant of integration is zero.

- Take away from Brachistochrone problem: Just an example of a BDE; we won't have to answer questions on it.

2.3 ODEs of Exact Form

10/5:

- Last time, we discussed separable ODEs.
- Today, we will study **exact form** equations, as discussed last class.
- **Exact form** (ODE): An ODE of the form

$$
g(x,y)\frac{\mathrm{d}y}{\mathrm{d}x} + f(x,y) = 0
$$

  where

$$
\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}
$$

- For equations of this form, there exists $F(x,y)$ such that

$$
\frac{\partial F}{\partial x} = f \qquad \frac{\partial F}{\partial y} = g \qquad F(x,y(x)) = C
$$

  for some $C \in \mathbb{R}$.
- To solve equations of this form, we need an **integrating factor**.
- **Integrating factor**: A number or function μ such that

$$
\mu g\frac{\mathrm{d}y}{\mathrm{d}x} + \mu f = 0 \qquad \frac{\partial}{\partial x}(\mu g) = \frac{\partial}{\partial y}(\mu f)
$$

- The solution to linear homogeneous equations of the form $\mathrm{d}y/\mathrm{d}t = p(t)y$ is

$$y(t) = y_0 \exp\left[\int_{t_0}^{t} p(\tau)\mathrm{d}\tau\right]$$

- Recall that $\mathrm{e}^{a+ib} = \mathrm{e}^a(\cos b + i\sin b)$, so

$$\mathrm{e}^{ix} = \cos x + i\sin x \qquad \cos x = \frac{1}{2}(\mathrm{e}^{ix} + \mathrm{e}^{-ix}) \qquad \sin x = \frac{1}{2i}(\mathrm{e}^{ix} - \mathrm{e}^{-ix})$$

- Example: If $y' = ky$, then $y' = -\lambda y$.
- If we have an inhomogeneous linear equation $\mathrm{d}y/\mathrm{d}t = p(t)y + f(t)$, then

$$\frac{\mathrm{d}y}{\mathrm{d}t} - py - f = 0$$

  but

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(1) \neq \frac{\mathrm{d}}{\mathrm{d}y}(-p(t)y - f(t))$$

- We wish to find an integrating factor $\mu(t, y)$ such that

$$\mu(t, y)\frac{\mathrm{d}y}{\mathrm{d}t} - \mu(t, y)p(t)y - \mu(t, y)f(t) = 0$$

  and

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mu) = \frac{\mathrm{d}}{\mathrm{d}y}(-\mu py - \mu f)$$

- Solution: Take μ to be a function of t, alone. Then

$$\mu'(t) = \frac{\mathrm{d}}{\mathrm{d}y}(-\mu py - \mu f) = -\mu(t)p(t)$$

  and we now have a homogeneous linear equation with solution

$$\mu(t) = \exp\left[-\int_{t_0}^{t} p(\tau)\mathrm{d}\tau\right]$$

  - If we let $P(t) = \int_{t_0}^{t} p(\tau)\mathrm{d}\tau$, then

$$\begin{aligned} \mathrm{e}^{-P(t)}y'(t) - p(t)y(t)\mathrm{e}^{-P(t)} &= \mathrm{e}^{-P(t)}f(t) \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-P(t)}y(t)\right) &= \mathrm{e}^{-P(t)}f(t) \\ \mathrm{e}^{-P(t)}y(t) &= \int_{t_0}^{t} \mathrm{e}^{-P(\tau)}f(\tau)\mathrm{d}\tau \end{aligned}$$

  - Thus, we finally have the solution to the inhomogeneous problem as follows: The IVP $y' = py + f$, $y(t_0) = y_0$ has solution

$$y(t) = y_0\mathrm{e}^{P(t)-P(t_0)} + \mathrm{e}^{P(t)}\int_{0}^{t} \mathrm{e}^{-P(\tau)}f(\tau)\mathrm{d}\tau$$

    where P is any anti-derivative of p.

- In particular, when $p(t) = a$, we get the **Duhamel formula** (which we should memorize).

- **Duhamel formula**: The following equation, which is the solution to an inhomogeneous linear equation with $p(t)=a$.
$$y(t)=y_0\mathrm{e}^{a(t-t_0)}+\int_{t_0}^{t}\mathrm{e}^{a(t-\tau)}f(\tau)\mathrm{d}\tau$$
    - Important for computing forced oscillation.
- Inspecting the inhomogeneous solution.
    - The first term is the solution to the homogeneous form. The second term deals with the initial value.
- Given an inhomogeneous equation, you can always write its solution as the combination of the solution to the homogeneous problem plus a particular solution, i.e.,
$$y=y_h+y_p$$
    - "The general solution equals the homogeneous solution plus a particular solution."
    - This is related to linear algebra, where the solution to $Ax=b$ is a particular solution x_p plus any vector $x\in\ker A$.
    - Thus, this idea will reappear in the theory of systems of linear ODEs.
- We now look at systems of linear ODEs.
- Consider the harmonic oscillator: A particle of mass m connected to an ideal spring (obeys Hooke's law) with no friction or gravity.
    - Newton's second law: The acceleration is proportional to the restoring force.
    - Hooke's law: The restoring force is of magnitude kx in the opposite direction to the displacement.
    - Thus, the ODE is of the form
$$x''=-\frac{k}{m}x$$
    - However, if there is damping (which will be proportional to the velocity), then the ODE is of the form
$$x''+\frac{b}{m}x'+\frac{k}{m}x=0$$
- Consider an ODE of the form
$$y''+ay'+by=0$$
for $a,b\in\mathbb{C}$.
    - Aim: Find $\mu,\lambda\in\mathbb{C}$ such that
$$(y'-\mu y)'-\lambda(y'-\mu y)=0$$
    - To find the parameters, we expand the above to
$$y''-(\mu+\lambda)y'+\mu\lambda y=0$$
    - Comparing with the original form, we have that $a=-(\mu+\lambda)$ and $b=\mu\lambda$.
    - It follows that μ,λ are the roots of $x^2+ax+b=0$, which we will call the **characteristic polynomial** of the ODE.
    - Substitute $x=y'-\mu y$. Then we can solve
$$x'-\lambda x=0$$
to determine that $x=A\mathrm{e}^{\lambda t}$.

- Returning the substitution, we have that
$$y' - \mu y = A\mathrm{e}^{\lambda t}$$
- Since the above is of the form $y' = ay + f$, we can apply the Duhamel formula. It follows that a particular solution is
$$A\int_0^t \mathrm{e}^{\mu(t-\tau)}\mathrm{e}^{\lambda\tau}\mathrm{d}\tau$$
- Thus, general solutions are of the form
$$y(t) = B\mathrm{e}^{\mu t} + C\mathrm{e}^{\mu t}\int_0^t \mathrm{e}^{(\lambda-\mu)\tau}\mathrm{d}\tau$$
- Evaluating the integral, we get
$$y(t) = B\mathrm{e}^{\mu t} + C\mathrm{e}^{\mu t}\frac{\mathrm{e}^{(\lambda-\mu)t} - 1}{\lambda - \mu}$$
which simplifies (by incorporating constants, etc.) to
$$y(t) = A_1\mathrm{e}^{\mu t} + B_1\mathrm{e}^{\lambda t}$$
for $\mu \neq \lambda$, or
$$y(t) = A_1\mathrm{e}^{\mu t} + B_1 t\mathrm{e}^{\mu t}$$
for $\mu = \lambda$.
- These linearly independent solutions form a basis of the space of solutions; all solutions can be expressed as a linear combination of these two functions.

- If our equation is of the form $y'' + ay' + by = f(t)$, then we just need to apply the Duhamel formula twice.
- Returning to the simple harmonic oscillator problem, we substitute $\omega = \sqrt{k/m}$ to get
$$x'' = \omega^2 x$$
  - The characteristic polynomial is
$$0 = x^2 + \omega^2 = (x + i\omega)(x - i\omega)$$
  - Thus, solutions are of the form
$$x = A_1\mathrm{e}^{i\omega t} + B_1\mathrm{e}^{-i\omega t}$$
  - It follows that the period is $T = 2\pi/\omega$.
  - To get a real (usable) solution, apply Euler's formula to get
$$\begin{aligned} x(t) &= A_1(\cos\omega t + i\sin\omega t) + B_1(\cos\omega t - i\sin\omega t) \\ &= A\cos\omega t + B\sin\omega t \end{aligned}$$
where $A = A_1 + B_1$, $B = iA_1 - iB_1$.
  - To match the initial condition $x(0) = x_0$, $x'(0) = v_0$, we use
$$x(t) = x_0\cos\omega t + \frac{v_0}{\omega}\sin\omega t$$
  - In other words,
$$\begin{cases} A = x_0 \\ \omega B = v_0 \end{cases} \qquad \begin{cases} A_1 + B_1 = x_0 \\ i\omega A_1 - i\omega B_1 = v_0 \end{cases}$$
so
$$\begin{cases} A = x_0 \\ B = \frac{v_0}{\omega} \end{cases} \qquad \begin{cases} A_1 = \frac{1}{2}\left[x_0 - \frac{iv_0}{\omega}\right] \\ B_1 = \frac{1}{2}\left[x_0 + \frac{iv_0}{\omega}\right] \end{cases}$$

2.4 ODE Examples

10/7: - Today, we will investigate a variety of examples of ODEs arising in real life.

- Michaelis-Menten kinetics: If E is an enzyme, S is its substrate, and P is the product, then the mechanism is

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_2} \mathrm{E} + \mathrm{P}$$

- The concentrations that we are concerned with are $[\mathrm{E}], [\mathrm{S}], [\mathrm{ES}], [\mathrm{P}]$.
- From the above mechanism, we can write the four rate laws

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -k_1[\mathrm{E}][\mathrm{S}] + k_{-1}[\mathrm{ES}] \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{E}] = -k_1[\mathrm{E}][\mathrm{S}] + (k_{-1} + k_2)[\mathrm{ES}] \tag{2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ES}] = k_1[\mathrm{E}][\mathrm{S}] - (k_{-1} + k_2)[\mathrm{ES}] \tag{3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{P}] = k_2[\mathrm{ES}] \tag{4}$$

- The initial conditions are $[\mathrm{S}] = [\mathrm{S}]_0$ and $[\mathrm{E}] = [\mathrm{E}]_0$.
- We can reduce these rate laws to the 2D system

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -k_1([\mathrm{E}]_0 - [\mathrm{ES}])[\mathrm{S}] + k_{-1}[\mathrm{ES}] \tag{5}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ES}] = k_1([\mathrm{E}]_0 - [\mathrm{ES}])[\mathrm{S}] - (k_{-1} + k_2)[\mathrm{ES}] \tag{6}$$

  - Note that to do so, we have used two conservation laws: The conservation of the enzyme plus enzyme-substrate complex, and the conservation of the substrate plus enzyme-substrate complex plus products.
- QSSA: Quasi steady-state assumption.
  - Assume that $[\mathrm{E}]_0/[\mathrm{S}]_0 \lll 1$.
  - It follows that $\mathrm{d}[\mathrm{ES}]/\mathrm{d}t \approx 0$ due to saturation of the enzyme and $[\mathrm{S}] \approx [\mathrm{S}]_0$ due to ever-more substrate being available.
- Then

$$[\mathrm{ES}] = \frac{[\mathrm{E}]_0[\mathrm{S}]}{K_M + [\mathrm{S}]}$$

where $k_M = (k_{-1} + k_2)/k_1$ is the **Michaelis-Menten constant**, a usual indication of enzyme activity.
- Substitute the above into Equation 5:

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -\frac{v_{\max}[\mathrm{S}]}{k_M + [\mathrm{S}]}$$

  - Note that $v_{\max} = k_2[\mathrm{E}]_0$.
- The above is now a differential equation of separable form; it's solution is

$$\int_{[\mathrm{S}]_0}^{[\mathrm{S}]} -\frac{(k_M + z)\,\mathrm{d}z}{zv_{\max}} = \int_0^t \mathrm{d}t$$

$$-\frac{k_M}{v_{\max}} \log\frac{[\mathrm{S}]}{[\mathrm{S}]_0} - \frac{1}{v_{\max}}([\mathrm{S}] - [\mathrm{S}]_0) = t$$

$$\begin{aligned}
\log\frac{[\mathrm{S}]}{[\mathrm{S}]_0}+\frac{[\mathrm{S}]}{k_M} &= \frac{[\mathrm{S}]_0-v_{\max}t}{k_M}\\
\frac{[\mathrm{S}]}{[\mathrm{S}]_0}\mathrm{e}^{[\mathrm{S}]/k_M} &= \exp\left(\frac{[\mathrm{S}]_0-v_{\max}t}{k_M}\right)\\
\frac{[\mathrm{S}]}{k_M}\mathrm{e}^{[\mathrm{S}]/k_M} &= \frac{[\mathrm{S}]_0}{k_M}\exp\left(\frac{[\mathrm{S}]_0-v_{\max}t}{k_M}\right)\\
\frac{[\mathrm{S}]}{k_M} &= W\left[\frac{[\mathrm{S}]_0}{k_M}\exp\left(\frac{[\mathrm{S}]_0-v_{\max}t}{k_M}\right)\right]\\
[\mathrm{S}] &= k_M W\left[\frac{[\mathrm{S}]_0}{k_M}\exp\left(\frac{[\mathrm{S}]_0-v_{\max}t}{k_M}\right)\right]
\end{aligned}$$

- Getting from line 5-6 (i.e., the introduction of W): Suppose we have an equation of the form $y\mathrm{e}^y = x$. We cannot express x in terms of y using elementary functions, so we must define W such that $y = W(x)$. Explicitly, W is the unique function of x that satisfies $W(x)\mathrm{e}^{W(x)} = x$.

- Harmonic oscillator.
- Recall that
$$x'' + \frac{k}{m}x = 0$$
- Substituting $\omega = \sqrt{k/m}$, we can solve the above for
$$x(t) = x(0)\cos(\omega t) + \frac{x'(0)}{\omega}\sin(\omega t)$$
- This is an integrable system with n degrees of freedom and $n-1$ scalar conservation laws??
- Conservation of mechanical energy:
$$E = \frac{1}{2}m|x'|^2 + \frac{1}{2}kx^2$$

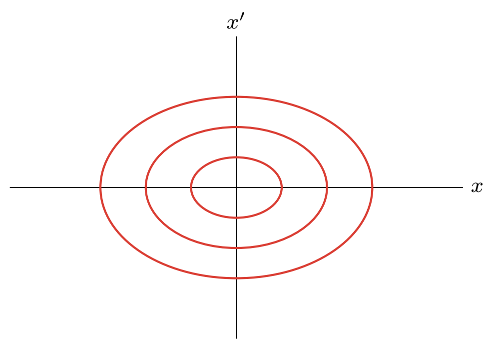


Figure 2.2: Conservation of mechanical energy in the harmonic oscillator.

- Differentiating wrt. x yields
$$\begin{aligned}
0 &= mx'x'' + kxx'\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m(x')^2\right) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}kx^2\right)
\end{aligned}$$
- This means that the solution is an ellipse in the xx'-plane, where each ellipse corresponds to an initial displacement and velocity.

- Mathematical pendulum.
- Equation of motion:

$$\begin{aligned} 0 &= l\theta'' + g\sin\theta \\ &= \ell\theta''\theta' + g\sin\theta\cdot\theta' \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Bigg(\underbrace{\frac{\ell}{2}|\theta'|^2 - g\cos\theta}_{E}\Bigg) \end{aligned}$$

- Initial values:

$$\theta(0) = \theta_0 \qquad\qquad \theta'(0) = 0$$

- It follows from the above that

$$\begin{aligned} \frac{\ell}{2}|\theta'|^2 - g\cos\theta_0 &= -g\cos\theta \\ \frac{\mathrm{d}\theta}{\mathrm{d}t} &= \sqrt{\frac{2g}{\ell}(\cos\theta_0 - \cos\theta)} \\ \int_{\theta_0}^{\theta}\sqrt{\frac{\ell}{2g(\cos\theta_0 - \cos\phi)}}\mathrm{d}\phi &= t \end{aligned}$$

  - This is an elliptical integral (and thus cannot be expressed in terms of the elementary functions).
- Suppose θ_0 is small. Then θ is small, and we can invoke the small angle approximation $\sin\theta \approx \theta$.
  - This yields an approximate equation of motion:

$$\ell\theta'' + g\theta = 0$$

  - From here, we can determine that $\theta(t) \approx \theta_0 \cos\sqrt{g/\ell}\cdot t$ and $T = 2\pi\sqrt{\ell/g}$.
- Kepler problem.
- Two bodies of mass m_1, m_2 are located at positions x_1, x_2 pulling on each other gravitationally.
  - The force of attraction is a conservative central force.
  - The potential between the two masses is a function of their distance, i.e, $U(|x_1 - x_2|)$.
- From Newton's second and third law, we get

$$m_1x_1'' = U'(|x_1 - x_2|)\frac{x_2 - x_1}{|x_2 - x_1|} \qquad\qquad m_2x_2'' = U'(|x_1 - x_2|)\frac{x_1 - x_2}{|x_1 - x_2|}$$

  - The derivative of potential is force.
  - The vector term provides direction.
- Conservation of momentum:

$$\begin{aligned} (m_1x_1 + m_2x_2)'' &= 0 \\ m_1x_1' + m_2x_2' &= C \end{aligned}$$

  - Let $M = m_1 + m_2$. Then the center of mass

$$\frac{m_1}{M}x_1 + \frac{m_2}{M}x_2$$

    moves inertially (i.e., does not accelerate or decelerate; is a stable reference frame) — we'll define it to be the origin.

- Conservation of angular momentum:
$$[m(x_1 - x_2)' \times (x_1 - x_2)]' = 0$$
  - $m = m_1 m_2 / (m_1 + m_2)$.
  - $\times$ indicates the cross product.
  - $L = m(x_1 - x_2)' \times (x_1 - x_2)$.
- It follows that $x_1 - x_2$ is always in a fixed plane, which we may call the **horizon plane**.
- Conservation of mechanical energy:
$$mq'' + U'(|q|)\frac{q}{|q|} = 0$$
$$\frac{m}{2}|q'|^2 + U(|q|) = E$$
  - $q = x_1 - x_2$.
- Introduce polar coordinates (r, ϕ).
  - Then $r^2\phi' = \ell_0$, $r = r(\phi)$, and $\mathrm{d}r/\mathrm{d}\phi = r'(t)/\phi'(t)$.
  - It follows that
$$\frac{m}{2}(|r'|^2 + |\phi'|^2) + U(r) = E$$
  - Since $U(r) = -Gm_1m_2/r$ for Newtonian gravity,
$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 + r^2 = \frac{2GMr^3}{\ell_0^2} + \frac{2Er^4}{m\ell_0^2}$$
  - The substitution $\mu = 1/r$ yields
$$\left(\frac{\mathrm{d}\mu}{\mathrm{d}\phi}\right)^2 + \mu^2 = \frac{2GM}{\ell_0^2}\mu + \frac{2E}{m\ell_0^2}$$
  - Differentiating again gives
$$2\frac{\mathrm{d}\mu}{\mathrm{d}\phi}\frac{\mathrm{d}^2\mu}{\mathrm{d}\phi^2} + 2\mu\frac{\mathrm{d}\mu}{\mathrm{d}\phi} = \frac{2GM}{\ell_0^2}\frac{\mathrm{d}\mu}{\mathrm{d}\phi}$$
  - Substituting $\mu = \cos(t)$ gives
$$\frac{\mathrm{d}^2\mu}{\mathrm{d}\phi^2} + \mu - \frac{GM}{\ell_0^2} = 0$$
    or
$$r = \frac{1}{GM/\ell_0^2 + \varepsilon\cos(\phi - \phi_0)}$$
    - This is a conic section!
- Thus, for example, we can calculate the precession of Mercury.
- Note that while we have determined the trajectory of our 2 bodies in terms of elementary functions, the n-body problem cannot be solved analytically.

2.5 Chapter 1: Introduction

From Teschl (2012).

Section 1.3: First Order Autonomous Equations

11/15:
- We start with the simplest nontrivial case of a first-order autonomous equation:
$$\dot{x} = f(x), \quad x(0) = x_0$$
  - We may let $t_0 = 0$ WLOG: If $\phi(t)$ is a solution to an autonomous equation satisfying $\phi(0) = x_0$, then $\psi(t) = \phi(t - t_0)$ is a solution with $\psi(t_0) = 0$.
- Solving this ODE.
  - Suppose $f(x_0) \neq 0$. Divide both sides by $f(x)$ and integrate from the initial conditions onward to yield
$$\int_0^t \frac{\dot{x}(s)\,\mathrm{d}s}{f(x(s))} = t$$
  - Define
$$F(x) := \int_{x_0}^{x} \frac{\mathrm{d}y}{f(y)}$$
    - Note that this is just the previous equation under the "u-substitution" $y(t) = x(t)$, $\mathrm{d}y = \dot{x}(t)\,\mathrm{d}t$, $y(0) = x(0) = x_0$, $y(t) = x$.
  - Thus, in our new notation, any possible solution x to the ODE must satisfy $F(x(t)) = t$. Since $F(x(t))$ is monotone (near x_0??), it can thus be inverted to yield the unique solution
$$\phi(t) = F^{-1}(t), \quad \phi(0) = F^{-1}(0) = x_0$$
- Teschl (2012) does a deep dive on the maximal integral where ϕ is defined.
- Examples of first-order autonomous systems given.
- Most of this section goes beyond what was covered in class in terms of depth.

Section 1.4: Finding Explicit Solutions

- Solving ODEs for explicit solutions is impossible in general unless the equation is of a particular form.
- This section: Classes of first-order ODEs which are explicitly solvable.
- Strategy: Find a change of variables that transforms the ODE into a solvable form.
- Linear equation.
$$\dot{x} = a(t)x \qquad\qquad \dot{x} = a(t)x + g(t)$$
  - The left equation above is the homogeneous linear equation, and the right equation above is the corresponding inhomogeneous linear equation.
  - The general solution to the homogeneous linear equation is
$$\phi(t) = x_0 A(t, t_0)$$
    where
$$A(t, s) = \mathrm{e}^{\int_s^t a(s)\mathrm{d}s}$$
  - The general solution to the inhomogeneous linear equation is
$$\phi(t) = x_0 A(t, t_0) + \int_{t_0}^{t} A(t, s) g(s)\,\mathrm{d}s$$
  - Teschl (2012) covers the more detailed mathematics of coordinate transformations in depth. This definitely would have been useful to understand for solving the PSet 1 problems, so I should return and understand it before the final.
- Using Mathematica to help solve ODEs and gain an intuition for how they work (e.g., with slope fields).
- Equations of exact form are covered in the problems to this section.

2.6 Chapter 8: Higher Dimensional Dynamical Systems

From Teschl (2012).

Section 8.5: The Kepler Problem

12/6:

- I would need Hamiltonian mechanics (and hence boundary value problems) in order to understand this.

Week 3

Linear Algebra Review

3.1 Elements of Linear Algebra

10/10:

- Today: Review of linear algebra.
- Start with a **vector space** over $\mathbb{R}$ or $\mathbb{C}$ or, more generally, any field K.
- **Vector space** (over K): A set equipped with addition and scalar multiplication such that the following axioms are satisfied.
  1. Commutativity and associativity of addition.
  2. Additive identity and inverse.
  3. Compatibility of scalar multiplication and addition (distributive laws).
  4. The additive identity times any vector is zero.
- In $\mathbb{R}^n, \mathbb{C}^n$, addition is component-wise and scalar multiplication is scaling of the element.
- For a homogeneous equation

$$y' = A(t)y = \begin{pmatrix} a_{11}(t)y^1 + a_{12}(t)y^2 + \cdots \\ \vdots \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{pmatrix}$$

  - If y_1, y_2 are solutions, any linear combination of them is a solution. This is called the **solution space** of the equation.
- **Linearly independent** (set of vectors): A set of vectors $x_1, \dots, x_m \in V$ for which the only coefficients $\lambda_1, \dots, \lambda_m$ such that

$$\lambda_1 x_1 + \cdots + \lambda_m x_m = 0$$

  is $\lambda_1 = \cdots = \lambda_m = 0$.
  - $\lambda_m \neq 0$ implies

$$x_m = -\frac{1}{\lambda_m}(\lambda_1 x_1 + \cdots + \lambda_{m-1} x_{m-1})$$

- **Maximal linear independence group**: A subset $X \subset V$ such that for any $y \in V$, $\{y\} \cup X$ is not linearly independent. *Also known as* **basis**.
- Theorem: Any basis in V has the same cardinality.
- **Dimension** (of V): The cardinality given by the above theorem. *Denoted by* $\boldsymbol{\dim V}$.

- We usually denoted a basis as an ordered n-tuple since the order often matters (for orientation?).
- Notational conventions.
  - For $\mathbb{R}^n, \mathbb{C}^n$, we will always use column vectors.
  - $x_1, x_2, \dots$ denotes vectors.
  - $x^1, x^2, \dots$ denotes the components of a column vector.
  - A vector component squared may be denoted $(x^1)^2$.
- **Standard basis** (for $\mathbb{R}^n$): The set of n vectors of length n which have a 1 as one entry and a zero in all the others and are all distinct.
- **Linear transformation** (of V to V): A mapping $\phi : V \to V$ satisfying

$$\phi(\lambda x + \mu y) = \lambda\phi(x) + \mu\phi(y)$$

- A mapping is completely determined by its action on the basis vectors:

$$\phi\left(\sum_{k=1}^{n} x^k e_k\right) = \sum_{k=1}^{n} x^k \phi(e_k)$$

- **Matrix** (of a linear transformation wrt. the standard basis): The $n \times n$ array

$$\begin{pmatrix} \phi(e_1) & \cdots & \phi(e_n) \end{pmatrix}$$

- If $\phi, \psi : V \to V$ are linear, $\phi \circ \psi$ is also linear.
  - Composition of linear transformations corresponds to matrix multiplication.
- Matrix multiplication: If

$$B = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$$

  then

$$AB = \begin{pmatrix} Ab_1 & \cdots & Ab_n \end{pmatrix}$$

  where

$$Ax = \begin{pmatrix} a_{11}x^1 + \cdots + a_{1n}x^n \\ \vdots \\ a_{n1}x^1 + \cdots + a_{nn}x^n \end{pmatrix}$$

- We can talk about matrix inverses: If it exists, it is unique, and

$$AA^{-1} = A^{-1}A = I_n$$

- Matrix multiplication is not commutative in general. Shao gives a counterexample.
- A is invertible iff the columns of A are a basis for $\mathbb{R}^n$ (resp. $\mathbb{C}^n$).
- **Determinant** (of A): Not explicitly defined.
- Properties of the determinant.
  - Multilinear.

$$\det\begin{pmatrix} a_1 & \cdots & \lambda a_k + \mu\tilde{a}_k & \cdots & a_n \end{pmatrix} = \lambda\det\begin{pmatrix} a_1 & \cdots & a_k & \cdots & a_n \end{pmatrix} + \mu\det\begin{pmatrix} a_1 & \cdots & \tilde{a}_k & \cdots & a_n \end{pmatrix}$$

  - Skew-symmetric.

$$\det\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{pmatrix} = -\det\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{pmatrix}$$

- Theorem: The determinant is uniquely characterized by these two axioms.
- $\det I_n = 1$.
- Shao goes over computing the determinant via minors.
- Special cases:
  - If the matrix is upper- or lower-triangular, the determinant is equal to the product of the diagonal entries.
  - If the matrix is blocked upper- or lower-triangular, e.g.,

$$A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$$

    then $\det A = \det A_1 \cdot \det A_2$.
- $\det(AB) = \det(A)\det(B)$.
- $\det A \neq 0$ iff A is invertible.
- Direct formula to compute the inverse.

$$A^{-1} = \frac{1}{\det A}\left(a_{\ell k}(-1)^{k+\ell}\det A_{k\ell}\right)$$

  - Tedious for higher-dimensional cases, but quite sufficient for $n = 2, 3$.
- Let A be $n \times n$, and let $Ax = b$.
  - If A is invertible, then $x = A^{-1}b$.
  - If A is not invertible and $b \in \operatorname{span}(a_1, \dots, a_n)$, then $x = x_h + x_p$ where $Ax_h = 0$ and $Ax_p = b$.
- **Kernel** (of A): The set of all vectors $y \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$) such that $Ay = 0$.
- **Range** (of A): The set of all linear combinations of $a_1, \dots, a_n$.
- Suppose $\phi : \mathbb{R}^n \to \mathbb{R}^n$ has matrix A under $(e_1, \dots, e_n)$. Let $(q_1, \dots, q_n)$ be another basis.
  - There exists a matrix Q such that $q_k = Qe_k$. Q is called the **connecting matrix** between $(e_1, \dots, e_n)$ and $(q_1, \dots, q_n)$.
  - Claim: Let $x \in \mathbb{R}^n$ have representation $x = (x^1, \dots, x^n)$ under the standard basis. Then under the Q basis, x has representation $x' = Q^{-1}(x^1, \dots, x^n)$. Similarly, $x = Qx'$.
  - Claim: ϕ has matrix $B = Q^{-1}AQ$ with respect to the Q basis.
- Matrix similarity: $A \sim B$ iff there exists Q invertible such that $B = Q^{-1}AQ$.
  - Implies that A and B describe the same matrix under different bases.
  - Matrix product under the old and new bases are related.

$$Q^{-1}ABQ = (Q^{-1}AQ)(Q^{-1}BQ)$$

  - Similarity preserves the determinant:

$$\det\left(Q^{-1}AQ\right) = \det\left(Q^{-1}\right)\det(A)\det(Q) = \det(A)\det\left(Q^{-1}\right)\det(Q) = \det(A)$$

- There is an extra example in Shao's notes (of a linear transformation in two bases).

3.2 Diagonalization and Jordan Normal Form

10/12:

- Similar matrices and Jordan Normal Form (JNF).
- Suppose $A : \mathbb{C}^n \to \mathbb{C}^n$ is linear. We can express A in a different basis with the help of the connecting matrix Q.
- In this lecture, we seek to find the most convenient basis in which to discuss our linear transformation.
- Today we will work in $\mathbb{C}^n$ (but all results hold for $\mathbb{R}^n$, too).
- **Invariant subspace** (of A): A subspace $K \subset \mathbb{C}^n$ such that $A(K) = K$.
- Suppose you have m invariant subspaces $K_1, \dots, K_m \subset \mathbb{C}^n$ whose pairwise intersection is $\{0\}$.
- **Direct sum** (of $K_1, \dots, K_m$): The collection of all vectors which can be represented as sums from each of the subspaces. *Denoted by* $\boldsymbol{K_1 \oplus \cdots \oplus K_m}$. *Given by*

$$K_1 \oplus \cdots \oplus K_m = \left\{ x \in \mathbb{C}^n \mid x = \sum_{j=1}^{m} x_j,\ x_j \in K_j \right\}$$

- Suppose $K_1, K_2 \in \mathbb{C}^n$ are invariant subspaces of A of dimension n_1, n_2, respectively, such that $K_1 \oplus K_2 = \mathbb{C}^n$. Then choosing a basis for K_1 and K_2, the matrix A takes the form

$$\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$$

  where B_1 is an $n_1 \times n_1$ block and B_2 is an $n_2 \times n_2$ block.
- **Eigenvalue** (of A): A complex number $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is not invertible. *Denoted by* $\boldsymbol{\lambda}$.
  - Equivalently, $\det(A - \lambda I) = 0$.
- **Characteristic polynomial**: The polynomial in z defined as follows. *Denoted by* $\boldsymbol{\chi_A(z)}$. *Given by*

$$\chi_A(z) = \det(A - zI)$$

  - Similar matrices have the same characteristic polynomials.
- **Spectrum** (of A): The set of all eigenvalues of A.
- **Eigenvector** (of A): A vector $v \in \mathbb{C}^n$ corresponding to an eigenvalue λ via

$$Av = \lambda v$$

- Claim: The set of all eigenvectors corresponding to λ form an invariant subspace.

  Proof.

$$A(v_1 + v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$$

  □

- **Eigenspace** (of λ): The vector subspace of $\mathbb{C}^n$ equal to the span of the eigenvectors of λ. *Denoted by* $\boldsymbol{V_\lambda}$.
- **Algebraic multiplicity** (of λ): The degree of the $(z - \lambda)$ term in the factorization of the characteristic polynomial. *Denoted by* $\boldsymbol{\alpha_\lambda}$.
- **Geometric multiplicity** (of λ): The dimension of the eigenspace of λ. *Denoted by* $\boldsymbol{\gamma_\lambda}$.

- $\gamma_\lambda \leq \alpha_\lambda$.
- If $\alpha_\lambda = \gamma_\lambda$ for each λ, then each eigenspace V_λ has a basis such that $\oplus_\lambda V_\lambda = \mathbb{C}^n$.
  - Under this basis, the matrix of A is diagonal with all λ's (along the diagonal) repeated according to their algebraic multiplicity.
- **Superdiagonal**: The set of entries in a matrix which are directly above a diagonal entry.
- **Jordan block**: A $d \times d$ matrix corresponding to an eigenvalue λ that has λ as every diagonal entry, 1 as every superdiagonal entry, and zeroes everywhere else. *Denoted by* $\boldsymbol{J_d(\lambda)}$.
  - A Jordan block is an example of a matrix with algebraic multiplicity d and geometric multiplicity 1.
  - The geometric multiplicity γ_j is the number of Jordan blocks with eigenvalue λ_j. Of course, when $\gamma_j = \alpha_j$ (in particular, if $\alpha_j = 1$), there is no Jordan block corresponding to λ_j at all.
- For any linear transformation, we can find a basis such that the matrix is the diagonalized Jordon blocks.
- Theorem: Let A be an $n \times n$ complex matrix. Then there is a **Jordan basis** Q under which

$$Q^{-1}AQ = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$$

  i.e., under which the matrix of $Q^{-1}AQ$ is block-diagonalized Jordan blocks.
  - The proof will not be tested — it is very hard. Shao will sketch it, though.
  - The proof is constructive: It will tell you how to convert a matrix into the Jordan normal form.
- Proof procedure:
  1. Determine the eigenvalues as well as their algebraic and geometric multiplicities.
     (a) Compute $\chi_A(z)$.
     (b) Find $\lambda_1, \dots, \lambda_m$ (factor $\chi_A(z)$).
     (c) Find $\alpha_1, \dots, \alpha_m$ (combine like terms in the factorization of $\chi_A(z)$).
     (d) Find $\gamma_1, \dots, \gamma_m$ ($\gamma_i = n - \operatorname{rank}(A - \lambda_i I)$).
  2. Find the **generalized eigenspaces** of each λ_i. This will allow us to block-diagonalize A.
     (a) For each λ_i, compute the $\ker(A - \lambda_i I) \subset \ker(A - \lambda_i I)^2 \subset \ker(A - \lambda_i I)^3 \subset \cdots$.
     (b) The sequence will stop at some $d_i \in \mathbb{N}$. In particular, it will stop when $\dim \ker(A - \lambda I)^{d_j} = \alpha_i$.
       - Claim: $\mathbb{C}^n = K_1 \oplus \cdots \oplus K_m$.
     (c) Since each K_i is an invariant subspace of A, we know that there is a matrix of the linear transformation corresponding to A of the form

$$\begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix}$$

     We now just need to choose the *best* basis of each K_i, i.e., the one that makes each B_i into a (direct sum of) Jordan block(s).
  3. Find the best basis for each K_i.

(a) Recall that each λ_i corresponds to $\gamma = \gamma_i$ linearly independent eigenvectors, which we will denote $v_{i,1}, \dots, v_{i,\gamma}$. We will block-diagonalize B_i into γ Jordan blocks, each of which corresponds to a $v_{i,j}$ as follows.
Every Jordan block is of the form

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Let the above block be of dimension $k_{i,j} = d$. It follows that this block will be responsible for linearly transforming d vectors in the Jordan basis. Let $v_{i,j,1} = v_{i,j}$ be the first of these d vectors. Then the submatrix of $v_{i,j,1}$ in the Jordan basis corresponding to this Jordan block is

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

which should make sense since we want $Av_{i,j} = \lambda_i v_{i,j}$ and under this definition,

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Now let $v_{i,j,2}$ be the second of the d vectors. Naturally, its submatrix in the Jordan basis should be

$$\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

But this implies that

$$\begin{aligned} \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ \lambda_i \\ \vdots \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_i \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \\ Jv_{i,j,2} &= v_{i,j,1} + \lambda_i v_{i,j,2} \\ (J - \lambda_i I)v_{i,j,2} &= v_{i,j,1} \end{aligned}$$

Naturally, this process will generalize to show that $(J - \lambda_i I)v_{i,j,k} = v_{i,j,k-1}$, i.e., we can recursively determine the $v_{i,j,1}, \dots, v_{i,j,k_{i,j}}$.

(b) However, there is slightly more subtlety than we might guess at first glance. Indeed, of our γ eigenvectors corresponding to λ_i, pick the first γ' to be elements of $\ker(A - \lambda_i I) \cap \operatorname{im}(A - \lambda_i I)$. This is a necessary condition for the existence of $v_{i,j,2}$ such that $(A - \lambda_i I)v_{i,j,2} = v_{i,j,1}$ for $j = 1, \dots, \gamma'$.

(c) Thus, using the above process, we will find $k_{i,j}$ elements of the Jordan basis for each $v_{i,j}$. The full, ordered set of these vectors, listed as follows, constitutes the Jordan basis.

$$\begin{array}{llll} v_{i,1,1}, & v_{i,1,2}, & \cdots, & v_{i,1,k_{i,1}} \\ v_{i,2,1}, & v_{i,2,2}, & \cdots, & v_{i,2,k_{i,2}} \\ \cdots & & & \\ v_{i,\gamma',1}, & v_{i,\gamma',2}, & \cdots, & v_{i,\gamma',k_{i,\gamma'}} \\ v_{i,\gamma'+1}, & v_{i,\gamma'+2}, & \cdots, & v_\gamma \end{array}$$

(d) Note that each of these vectors is naturally an element of the generalized eigenspace K_i since for each $k = 1, \dots, k_{i,j}$, the formula $(A - \lambda_i I)v_{i,j,k} = v_{i,j,k-1}$ implies that

$$(A - \lambda_i I)^k v_{i,j,k} = 0$$

Also note that each $k_{i,j} \leq d_i$ and $k_{i,1} + \cdots + k_{i,\gamma'} + \gamma - \gamma' = \alpha_i$.

(e) Under this basis, the Jordan normal form of A on the generalized eigenspace K_i will be

$$\begin{pmatrix} J_{k_{i,1}}(\lambda) & & & & \\ & J_{k_{i,2}}(\lambda) & & & \\ & & \ddots & & \\ & & & J_{k_{i,\gamma'}}(\lambda) & \\ & & & & \lambda I_{\gamma-\gamma'} \end{pmatrix}$$

- **Generalized eigenspace** (of λ): The kernel of $(A - \lambda I)^{d_\lambda}$. *Denoted by* $\boldsymbol{K_\lambda}$. *Given by*

$$K_\lambda = \ker(A - \lambda I)^{d_\lambda}$$

- $\boldsymbol{d_\lambda}$: The power of $A - \lambda I$ for which the kernel stabilizes.
- The JNF computation can be really heavy; we'll only ever compute 2×2 or 3×3 versions.
- Example[1]:
  - Consider

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$$

  - Then

$$\chi_A(z) = z(z-1)^2$$

  - (1) It follows that

$$\lambda_1 = 0 \qquad\qquad \lambda_2 = 1$$

  - (2) We have that

$$\ker(A - 0I) = \operatorname{span}\left\{\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}\right\} \qquad \ker(A - 1I) = \operatorname{span}\left\{\begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}\right\}$$

    - We call the left vector above q_1 and the right vector above q_2.
  - Thus,

$$A \sim \left(\begin{array}{c|cc} 0 & & \\ \hline & 1 & x \\ & & 1 \end{array}\right)$$

[1] Largely ignore this misguided relic of class that day.

- We find that
$$(A - 1I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix}$$
so
$$\ker(A - I)^2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 10 \end{pmatrix} \right\}$$
- Clearly,
$$\ker(A - I) \subsetneq \ker(A - I)^2$$
so we can stop here because the dimension of the kernel has reached the algebraic multiplicity.
- Since $q_2 \in K_1$, q_3 solves the equation $(A - I)q_3 = q_2$.
- We know that
$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} e_1 = \lambda e_1 \qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} e_2 = e_1 + \lambda e_2$$
- It follows that
$$q_3 = \begin{pmatrix} 0 \\ 3 \\ -5 \end{pmatrix}$$
and hence
$$Q = \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 3 \\ 2 & 5 & -5 \end{pmatrix}$$
and
$$Q^{-1}AQ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

- Simple cases.
- The 2×2 case.
  - $A \in \mathcal{M}^2(\mathbb{C})$ can only have nontrivial Jordan form if it has a single eigenvalue λ with $\alpha_\lambda = 2$ and $\gamma_\lambda = 1$. If both equal 2, then $A = \lambda I_2$. If it has two eigenvalues, then it is regularly diagonalizable.
  - In this particular case, calculate λ from $\chi_Z(z) = (z - \lambda)^2$, find one eigenvector v, and find the other generalized eigenvector u; u will satisfy $(A - \lambda I)u = v$. The connecting matrix will be $Q = (v|u)$[2] and the JNF is
$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$
- The 3×3 case.
  - We divide into three nontrivial cases: $\chi_A(z) = (z - \lambda)^3$ with $\gamma_\lambda = 2$, $\chi_A(z) = (z - \lambda)^3$ with $\gamma_\lambda = 1$, and $\chi_A(z) = (z - \lambda)^2(z - \mu)$ with $\gamma_\lambda = 1$.
  - In the first case, we have two eigenvectors v_1, v_2 (make sure to pick v_1 such that it is also in the column space of $A - \lambda I$!). We can find the third (generalized) eigenvector by solving $(A - \lambda I)u = v_1$. Then $Q = (v_1|u|v_2)$ and the JNF is
$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

[2]Order matters! We need the eigenvector, specifically, to get scaled by λ only.

- In the second case, we have one eigenvector v. We can find the second and third generalized eigenvectors by solving $(A-\lambda I)u_1 = v$ and $(A-\lambda I)u_2 = u_1$. Then $Q = (v|u_1|u_2)$ and

$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

- In the third case, we have two eigenvectors v_λ, v_μ. We can find the third (generalized) eigenvector by solving $(A-\lambda I)u = v_\lambda$. Then $Q = (v_\lambda|u|v_\mu)$ and

$$Q^{-1}AQ = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}$$

3.3 Matrix Calculus

10/14: • Today: Matrix calculus.

- We introduced the Jordan normal form because it is an easy form on which to do matrix calculus.
- **Matrix norm**: A function for $n \times n$ complex matrices such that
  1. $\|A\| \geq 0$, $\|A\| = 0$ iff $A = 0$.
  2. $\|A+B\| \leq \|A\| + \|B\|$.
  3. $\|\lambda A\| = |\lambda|\|A\|$.
  4. $\|AB\| \leq \|A\|\|B\|$.

  Denoted by $||\cdot||$.

  - The first three axioms above are the normal norm axioms; the last one is unique to matrix norms.
- **Operator norm**: The norm defined by

$$\|Ax\| = \sup_{|x|=1} |Ax|$$

- **??**: The norm defined by

$$\|A\| = \sum_{i,j=1}^{n} |a_{i,j}|$$

- Theorem: Any two matrix norms are equivalent.
- **Convergent** (sequence of matrices): A sequence of matrices A_n for which there exists A such that $\|A_n - A\| \to 0$ as $n \to \infty$. *Denoted by* $\boldsymbol{A_n \to A}$.
  - Note that $\|A_n - A\| \to 0$ iff the entries of A_n converge to the entries of A.
- Suppose $A(t) = (a_{ij}(t))_{i,j=1}^n$ is a matrix function. Then

$$A'(t) = (a_{ij}'(t))_{i,j=1}^n \qquad \int_{t_0}^{t} A(t)\,\mathrm{d}t = \left(\int_{t_0}^{t} a_{ij}(\tau)\mathrm{d}\tau \right)_{i,j=1}^{n}$$

- The product rule holds:

$$\frac{\mathrm{d}}{\mathrm{d}t}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t)$$

- However, matrix multiplication is not commutative. This can get us into trouble in the following situation: We might think that
$$\frac{\mathrm{d}}{\mathrm{d}t}[A(t)^2] = 2A'(t)A(t)$$
but, in fact,
$$\frac{\mathrm{d}}{\mathrm{d}t}[A(t)^2] = A'(t)A(t) + A(t)A'(t)$$
  - For example, let
$$A(t) = \begin{pmatrix} 0 & 1 \\ t & 0 \end{pmatrix}$$
  - Then
$$A'(t) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
  - It follows that
$$\frac{\mathrm{d}}{\mathrm{d}t}[A'(t)^2] = \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{A'(t)A(t)} + \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A(t)A'(t)}$$
  - Notice that $A'(t)A(t) \neq A(t)A'(t)$.
- Suppose we have a matrix A and we want to compute A^{100}.
- If A is diagonalizable, then $A^n = Q\Lambda^n Q^{-1}$.
- What if A is not diagonalizable?
  - Then we convert to A to Jordan normal form $A = QBQ^{-1}$. Thus, we just need to compute the powers of the Jordan blocks.
  - Suppose
$$J_d(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$
    - In a given Jordan block, all entries above the diagonal are 1.
  - Decompose
$$J_d(\lambda) = \lambda I_d + N_d$$
  - Note that N_d is nilpotent — every successive power to which you raise it shifts the 1s up one row until it becomes the zero matrix.
  - In computing $[J_d(\lambda)]^m$, invoke the binomial expansion. When $m < d$ invoke the full expansion. When $m \geq d$, neglect all zero terms (terms with N_d^i for $i \geq m$):
$$[J_d(\lambda)]^m = \binom{m}{0}\lambda^m I_d + \binom{m}{1}\lambda^{m-1}N_d + \cdots + \binom{m}{d-1}\lambda^{m-d+1}N_d^{d-1}$$
  - Example: When $d = 3$, then
$$\begin{pmatrix} \lambda & 1 & 0 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & m(m-1)\lambda^{m-2} \\ & \lambda^m & m\lambda^{m-1} \\ & & \lambda^m \end{pmatrix}$$
- We will only compute JNF for 2×2 and 3×3; Shao reviews these cases from last class.

- We now have a formula to compute the powers of matrices with ease, so we can move onto more complicated functions of matrices now.
- Consider the power series
$$f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$$
  - The c_i are complex coefficients.
- **Analytic** (function): A function whose Taylor series (locally) converges and converges to the function in question.
- We can consider an analytic function of matrices:
$$f(A) = c_0 I + c_1 A + c_2 A^2 + \cdots$$
- **Radius of convergence**: The number R such that the series converges absolutely for $\|A\| < R$.
  - We do not talk about the radius of convergence any more in this course.
- **von Neumann series**: The series $I + A + A^2 + \cdots$ converging to $(I_n - A)^{-1}$ for any $\|A\| < 1$.
  - Example: We can check that the von Neumann series for N_d converges.
- Suppose $A = QBQ^{-1}$. Then
$$\begin{aligned} f(A) &= f(QBQ^{-1}) \\ &= c_0 I + c_1(QBQ^{-1}) + c_2(QBQ^{-1})^2 + \cdots \\ &= Q(c_0 I + c_1 B + c_2 B^2 + \cdots)Q^{-1} \\ &= Qf(B)Q^{-1} \end{aligned}$$
  - Going even further,
$$B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \quad \Longrightarrow \quad f(B) = \begin{pmatrix} f(B_1) & 0 \\ 0 & f(B_2) \end{pmatrix}$$
  - In particular, if A is diagonalizable, then
$$f(A) = Q \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} Q^{-1}$$
- Suppose A is not diagonalizable, and f is some analytic function.
  - Then in the viscinity of a, f can be approximated by the Taylor series
$$f(z) = f(a) + f'(a)(z-a) + \frac{1}{2!} f^{(2)}(a)(z-a)^2 + \cdots$$
  - Similarly, we can approximate $f[J_d(\lambda)]$ in the viscinity of λI_d with the Taylor series
$$\begin{aligned} f[J_d(\lambda)] &= f(\lambda I_d + N_d) \\ &= f(\lambda I_d) + f'(\lambda I_d)[(\lambda I_d + N_d) - \lambda I_d] + \frac{1}{2!} f^{(2)}(\lambda I_d)[(\lambda I_d + N_d) - \lambda I_d]^2 + \cdots \\ &= f(\lambda) I_d + f'(\lambda) N_d + \frac{1}{2!} f^{(2)}(\lambda) N_d^2 + \cdots \\ &= \begin{pmatrix} f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(d-1)}(\lambda)}{(d-1)!} \\ & f(\lambda) & \ddots & \vdots \\ & & \ddots & f'(\lambda) \\ & & & f(\lambda) \end{pmatrix} \end{aligned}$$

- **Matrix exponential** (of A): The matrix with identical dimensions to A defined by the following power series. *Denoted by* $\mathbf{e}^{\boldsymbol{A}}$. *Given by*

$$\mathrm{e}^A = I_n + A + \frac{1}{2!}A^2 + \cdots$$

  - This power series is convergent for matrices with $\|A\| < 1$ since $\|A^m\| \leq \|A\|^m \to 0$.
  - Usual rules that you might expect the matrix exponential to obey based on the notation are obeyed.

$$\mathrm{e}^{(t+\tau)A} = \mathrm{e}^{tA}\mathrm{e}^{\tau A} \qquad\qquad \mathrm{e}^{A+B} = \mathrm{e}^A\mathrm{e}^B$$

- An explicit formula for the e^{tA}.
  - We know that $tA = tQBQ^{-1}$, where we may take B be in JNF.
  - Consider $\mathrm{e}^{tJ_3(\lambda)}$, for example.
  - Then from the above, we have that

$$\mathrm{e}^{tJ_3(\lambda)} = \begin{pmatrix} \mathrm{e}^{t\lambda} & t\mathrm{e}^{t\lambda} & \frac{t^2}{2}\mathrm{e}^{t\lambda} \\ & \mathrm{e}^{t\lambda} & t\mathrm{e}^{t\lambda} \\ & & \mathrm{e}^{t\lambda} \end{pmatrix}$$

  - Thus,

$$\mathrm{e}^{tA} = Q\mathrm{e}^{tB}Q^{-1}$$

- Next time: First order linear systems with constant coefficients; will make use of e^{tA}.
- Next Wednesday: Review; next Friday: Midterm.

Week 4

Linear Systems

4.1 Autonomous Linear Systems

10/17:
- Today: General theory for autonomous linear systems.
- Review session Wednesday (no new material).
- First midterm Friday.
  - Test problems will be slight variations of homework problems or examples given in class.
- **Linear autonomous system**: A system of n linear equations written in the following form. *Denoted by* $\boldsymbol{y' = Ay}$. *Given by*

$$\begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}' = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} \qquad\qquad y(0) = 0$$

  - Note that the a_{ij}'s are complex or real.
- The explicit solution is given by $y(t) = \mathrm{e}^{tA} y_0$.
  - Recall that $\mathrm{d}/\mathrm{d}t\left(\mathrm{e}^{tA}\right) = A\mathrm{e}^{tA}$, as we can show via the power series expansion.
- **Picard iteration**: We take

$$\begin{aligned}
y'(t) &= Ay(t) \\
\int_0^t y'(\tau)\,\mathrm{d}\tau &= \int_0^t Ay(\tau)\mathrm{d}\tau \\
y(t) &= y_0 + \int_0^t Ay(\tau_1)\mathrm{d}\tau_1 \\
&= y_0 + \int_0^t A\left[y_0 + \int_0^{\tau_1} Ay(\tau_2)\mathrm{d}\tau_2\right]\mathrm{d}\tau_1 \\
&= y_0 + tAy_0 + \int_0^t\int_0^{\tau_1} A^2 y(\tau_2)\mathrm{d}\tau_2\mathrm{d}\tau_1 \\
&= y_0 + tAy_0 + \int_0^t\int_0^{\tau_1} A^2\left[y_0 + \int_0^{\tau_2} Ay(\tau_3)\mathrm{d}\tau_3\right]\mathrm{d}\tau_2\mathrm{d}\tau_1 \\
&= y_0 + tAy_0 + \frac{t^2A^2}{2} + \int_0^t\int_0^{\tau_1}\int_0^{\tau_2} A^3 y(\tau_3)\mathrm{d}\tau_3\mathrm{d}\tau_2\mathrm{d}\tau_1
\end{aligned}$$

$$= \sum_{k=0}^{m} \frac{t^k A^k}{k!} y_0 + A^{m+1} \underbrace{\int_0^t \cdots \int_0^{\tau_m}}_{m+1} y(\tau_{m+1}) \mathrm{d}\tau_{m+1} \cdots \mathrm{d}\tau_1$$

  - We get from the second to the third line by substituting $y(t)$, as defined into the second line, into where it appears in the integral.
  - This is one form of the Picard iteration. Another that's more consistent with other techniques we'll use later is presented in the reading. The above one substitutes in the first equation each time; the other one substitutes in the new equation each time.

- We want to show that the integral converges to zero.
  - The magnitude of the remainder is less than or equal to

$$\|A\|^{m+1} \left(\sup_{\tau \in [0,t]} |y(\tau)| \right) \frac{t^{m+1}}{(m+1)!}$$

    - Justification of this term: Look at the rightmost term in the last line of the Picard iteration above. Imagine taking the norm of it. Splitting the "scalar" integral from the matrix allows us to take a matrix norm, and the property $\|AB\| \leq \|A\|\|B\|$ tells us that $\left\|A^{m+1}\right\| \leq \|A\|^{m+1}$. Then with respect to the integral, if we evaluate it, we will get the next polynomial term in the sequence — $t^{m+1}/(m+1)!$ — times at most the maximum value of y at every infinitesimal.
  - We can visualize lower-dimensional integrals as the volume of the corresponding unit **simplex**.
    - For example, in $\mathbb{R}^2$,

$$\int_0^1 \int_0^{\tau_1} 1 \mathrm{d}\tau_2 \, \mathrm{d}\tau_1$$

      can be visualized as the area of the unit triangle. This rationalizes why it evaluates to $1/2$, the area of said triangle.
    - In $\mathbb{R}^3$,

$$\int_0^1 \int_0^{\tau_1} \int_0^{\tau_2} 1 \mathrm{d}\tau_3 \mathrm{d}\tau_2 \, \mathrm{d}\tau_1$$

      can be visualized as the area of the unit simplex. This rationalizes why it evaluates to $1/3! = 1/6$, the volume of said simplex.
  - Since $(m+1)! \to \infty$ faster than any other term, the whole thing goes to zero.
  - Thus, since the remainder goes to zero as we add more terms, we eventually reach the limit

$$\begin{aligned} y(t) &= \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k y_0 \\ &= \mathrm{e}^{tA} y_0 \end{aligned}$$

- **Simplex**: A higher-dimensional generalization of a triangle.
- We now consider the following inhomogeneous equation. An appropriate integrating factor still helps.

$$\begin{aligned} y' &= Ay + f(t) \\ y' - Ay &= f(t) \\ \mathrm{e}^{-tA} y' - A\mathrm{e}^{-tA} y &= \mathrm{e}^{-tA} f(t) \\ \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathrm{e}^{-tA} y(t) \right) &= \mathrm{e}^{-tA} f(t) \\ \mathrm{e}^{-tA} y(t) - y_0 &= \int_0^t \mathrm{e}^{-\tau A} f(\tau) \mathrm{d}\tau \\ y(t) &= \mathrm{e}^{tA} y_0 + \int_0^t \mathrm{e}^{(t-\tau)A} f(\tau) \mathrm{d}\tau \end{aligned}$$

- We also call this the Duhamel formula.

- Note that if your time scale starts from t_0, then

$$y(t) = \mathrm{e}^{(t-t_0)A} y(t_0) + \int_{t_0}^{t} \mathrm{e}^{(t-\tau)A} f(\tau)\mathrm{d}\tau$$

- The utility of JNF: If we want to understand $\mathrm{e}^{tA} y_0$, we convert $A = QBQ^{-1}$, allowing us to evaluate e^{tA}.
  - Shao reviews some facts of JNF from previous lectures.

- From last lecture, we have that

$$\mathrm{e}^{tA} y_0 = Q\mathrm{e}^{tB} Q^{-1} y_0$$

- Example: Let

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$$

  - This is the same matrix from a previous lecture. As before, we have that

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 3 \\ 2 & 5 & -5 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

    - Recall that the left two vectors are normal eigenvectors (the leftmost one corresponds to $\lambda_1 = 0$ and the middle one corresponds to $\lambda_2 = 1$) and the rightmost one is a generalized eigenvector.
  - We can compute that

$$\mathrm{e}^{tB} = \begin{pmatrix} \mathrm{e}^{0t} & 0 & 0 \\ 0 & \mathrm{e}^{1t} & 1t\mathrm{e}^{1t} \\ 0 & 0 & \mathrm{e}^{1t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{t} & t\mathrm{e}^{t} \\ 0 & 0 & \mathrm{e}^{t} \end{pmatrix}$$

  - It follows that

$$\begin{aligned} \mathrm{e}^{tA} y_0 &= Q \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{t} & t\mathrm{e}^{t} \\ 0 & 0 & \mathrm{e}^{t} \end{pmatrix} Q^{-1} y_0 \\ &= \begin{pmatrix} -3t\mathrm{e}^t + \mathrm{e}^t & 2t\mathrm{e}^t & t\mathrm{e}^t \\ 3t\mathrm{e}^t - 10\mathrm{e}^t + 10 & -2t\mathrm{e}^t + 6\mathrm{e}^t - 5 & -t\mathrm{e}^t + 3\mathrm{e}^t - 3 \\ -15t\mathrm{e}^t + 20\mathrm{e}^t - 20 & 10t\mathrm{e}^t - 10\mathrm{e}^t + 10 & 5t\mathrm{e}^t - 5\mathrm{e}^t + 6 \end{pmatrix} \begin{pmatrix} y_0^1 \\ y_0^2 \\ y_0^3 \end{pmatrix} \end{aligned}$$

- **Stable** (eigenvalue): An eigenvalue $\lambda_j = \sigma_j + i\beta_j$ for which $\sigma_j < 0$.

- **Unstable** (eigenvalue): An eigenvalue $\lambda_j = \sigma_j + i\beta_j$ for which $\sigma_j > 0$.

- **Stable** (subspace of the system): A generalized eigenspace corresponding a stable eigenvalue.

- **Unstable** (subspace of the system): A generalized eigenspace corresponding an unstable eigenvalue.
  - If λ_j is unstable, then the corresponding entries in e^{tB_j} are exponentially growing functions.
  - If λ_j is stable, then the corresponding entries in e^{tB_j} are exponentially decreasing functions.
  - If $\sigma_j = 0$, then the "stability" depends on the geometric multiplicity??
  - Along the stable subspaces, your points will be attracted to zero.
  - Along the unstable subspaces, your points will be repelled from zero.

  - If $\sigma_h = 0$, then we have rotation around a point, oscillation about zero, or oscillation whose magnitude grows to infinity. We do not talk about its stability.
    - We do not include the eigenvector corresponding to $\lambda_1 = 0$ in the above basis of the stable subspace because the solution oscillates about y_1??
  - The stable subspace of our example is

$$\operatorname{span}\left\{\begin{pmatrix}1\\-1\\5\end{pmatrix},\begin{pmatrix}0\\3\\-5\end{pmatrix}\right\}$$

- Recall that B_j acts on K_j.
  - ... in picture??
  - Recall that $\mathbb{C}^n = K_1 \oplus \cdots \oplus K_m$.
  - P_j is not an *orthogonal* projection, but it is a projection of y_0 onto K_j. It's also a polynomial??
- Consider the order n linear differential equation

$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x = 0$$

  - Then we can make a system out of it:

$$\begin{pmatrix}y^1\\\vdots\\y^n\end{pmatrix}' = \underbrace{\begin{pmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1\\ -a_0 & -a_1 & \cdots & -a_{n-1}\end{pmatrix}}_{F[p]}\begin{pmatrix}y^1\\\vdots\\y^n\end{pmatrix}$$

    - Recall how to do the transformation from Lecture 1.
  - $F[p]$ is the **Frobenius matrix**.
  - The transpose of this matrix is a very special matrix called the **companion matrix** $C[p] = F[p]^T$.
  - Claim: Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$. Then $\chi_{C[p]} = \chi_{F[p]} = p(z)$.

    Proof. Do the Laplace expansion with respect to the last column of $A - zI$ (companion) or last row (Frobenius). □
  - Roots of $p(z)$ are the eigenvalues of $F[p]$ and $C[p]$.
  - Claim: $C[p]$ has **minimal polynomial** $p(z)$.

    Proof. We have that $C[p]e_i = e_{i+1}$ for $i = 1, \dots, n-1$ and

$$C[p]e_n = -a_0e_1 - \cdots - a_{n-1}e_n$$

    which implies that if $r(z)/\deg r < n$ nullifies $C[p]$, then necessarily $r(z) = p(z)$ since $(z-\lambda_j)^{<\alpha_j}$?? □
  - Claim: $C[p], F[p]$ have the same Jordan normal form.
    - More generally, transpose matrices are similar so they have the same JNF.
- **Monic polynomial**: A polynomial whose highest-degree coefficient equals 1.
- **Minimal polynomial** (of A): The unique monic polynomial p of smallest degree such that $p(A) = 0$.
- Theorem: In the Jordan normal form $F[p]$, each λ_j corresponds to only one Jordan block.

- Thus,

$$F[p] \sim \begin{pmatrix} J_{\alpha_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{\alpha_m}(\lambda_m) \end{pmatrix}$$

The implication is that

$$J_d(\lambda) \neq \begin{pmatrix} \lambda & & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}$$

ever??

- Corollary: The solution $y(t)$ is of the form

$$(\cdots) + a_1 \mathrm{e}^{t\lambda_j} + \cdots + c_{\alpha_j - 1} t^{\alpha_j - 1} \mathrm{e}^{t\lambda_j} + \cdots$$

- Example: Solving a second-order ODE.

$$x'' + ax' + bx = 0 \quad \Longleftrightarrow \quad \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}$$

  - The characteristic polynomial of the equation (and this matrix) is $z^2 + az + b = 0$.
  - If $\lambda_1 \neq \lambda_2$, then $x(t) = A\mathrm{e}^{t\lambda_1} + B\mathrm{e}^{t\lambda_2}$. If $\lambda_1 = \lambda_2 = \lambda$, then $x(t) = A\mathrm{e}^{t\lambda} + Bt\mathrm{e}^{t\lambda}$.

4.2 Midterm 1 Review

10/19:

- Notes on Friday's exam.
  - Three problems. All will be calculations for specific equations. They will all be standard examples that appeared in the lectures or homeworks.
  - The materials that you can bring to the exam are the notes on JNF (printed). You will be dealing with the JNF of 2×2 or 3×3 matrices.
- Review session today, no new content.
- Remind Shao to post teaching notes from more recent weeks.
- **Ordinary differential equation**: An equation that involves an unknown function together with its derivatives. *Given by*

$$F(t, y, y', y'', \dots, y^{(n)}) = 0$$

- **Order** (of an ODE): The highest order derivative present in the ODE.
- Two types of ODE problems: IVPs and BVPs.
  - IVPs arise in dynamical systems.
  - BVPs arise in variational problems in physics.
- We are primarily interested in ODEs which can be explicitly solved for $y \in C^1(\mathbb{R}^n)$ (resp. $C^1(\mathbb{C}^n)$).
- Two types of equations:
  - A higher-order scalar equation.
  - The more general form of vector-valued systems of the form $y' = f(t, y)$.
- In order to determine y, the initial value $y(t_0) = y_0$ is needed.
  - If a vector-valued system, you need $y_0^1, \dots, y_0^n$ (all components).

  - If a scalar system, you need $y(t_0), y'(t_0), \dots, y^{(n-1)}(t_0)$.
- The idea of well-posedness is not yet well-defined in the course; we will cover it after the midterm.
- **Well-posed** (IVP): For every initial value, there is only one unique solution, and for a small change in the initial value, there is only a small change in the solution (continuous dependence on initial values).
- The theorem that we've been relying on but haven't proven yet: **Cauchy-Lipschitz / Picard-Lindelöf theorem**.
- **Cauchy-Lipschitz theorem**: If $f(t, y)$ is Lipschitz continuous with respect to y, then the IVP is locally well-posed. *Also known as* **Picard-Lindelöf theorem**.
  - The term **locally well-posed** has not been rigorously defined either.
- Given any ODE, it is usually very easy to verify the Lipschitz condition for the RHS.
- Example of an IVP that is not locally well-posed.
  - $y = \sqrt{y}$, $y(0) = 0$.
  - Note that if we start at any $t_0 > 0$, then this IVP *is* locally well-posed.
- No Cauchy-Lipschitz in the first midterm; just calculations. We will need the precise statement in the second midterm, though.
- We are not going to talk about solutions that require power series because that inevitably involves complex analysis.
- Explicitly solvable equations: Equations of separable form, i.e., the IVP $y'(t) = f(y)g(t)$, $y(t_0) = y_0$.
- From C-L theorem: If $f(y)$ is continuously differentiable in some neighborhood of y_0, then the solution is unique.
- If $f(y_0) = 0$, then $y(t) = y_0$.
  - Because then $y'(t) = f(y_0)g(t) = 0$, so y is a constant function.
- If $f(y) \neq 0$ in some neighborhood of y_0, then the solution should satisfy the implicit equation

$$\int_{y_0}^{y} \frac{\mathrm{d}w}{f(w)} = \int_{t_0}^{t} g(\tau)\mathrm{d}\tau$$

  - We use the chain rule to make separation of variables rigorous: We can differentiate the LHS above wrt. t and get $y'(t)/f(y(t))$.
  - Relating the $f(y_0) = 0$ and $f(y) \neq 0$ cases and not making them overlap: We start integrating from the nonzero value.
- Examples: $y'(t) = p(t)y(t)$ is homogeneous linear. It follows that

$$y(t) = \exp\left[\int_{t_0}^{t} p(\tau)\mathrm{d}\tau\right] y_0$$

- If $p(t) = r \neq 0$, then the solution is exponential growth or decay:

$$y(t) = y_0 \mathrm{e}^{r(t-t_0)}$$

- Logistic growth:

$$y'(t) = ry\left(1 - \frac{y}{M}\right) \quad \Longleftrightarrow \quad y(t) = \frac{My_0\mathrm{e}^{rt}}{M + y_0(\mathrm{e}^{rt} - 1)}$$

  - Shao gives the related implicit integral equation and logarithmic equation as well.
- There exist equations which cannot be solved by separation of variables. One case is equations of the form
$$g(x,y)\frac{\mathrm{d}y}{\mathrm{d}x}+f(x,y)=0$$
where $\partial_x g(x,y)=\partial_y f(x,y)$.
  - In this case, there exists $F(x,y)$ such that $\partial_x F=f$, $\partial_y F=g$, and $F(x,y)=C$ is the relation satisfied by the solution.
  - These are **exact form** equations.
- Not all equations satisfy this relation. However, it is often possible (though potentially quite hard) to find an **integrating factor** by which you can multiply your equation to put it in exact form.
- Special case where it is easy to find the integrating factor: Consider the inhomogeneous linear equation $y'(t)=p(t)y(t)+f(t)$. Then the integrating factor is
$$\mu=\exp\left[-\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]$$
- Multiplying through, we get
$$\begin{aligned}\exp\left[-\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]f(t)&=\exp\left[-\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]y'(t)-\exp\left[-\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]p(t)y(t)\\&=\frac{\mathrm{d}}{\mathrm{d}t}\left\{\exp\left[-\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]y(t)\right\}\\y(t)&=\exp\left[\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]y_0+\exp\left[\int_{t_0}^t p(\tau)\mathrm{d}\tau\right]\cdot\int_{t_0}^t\exp\left[-\int_{t_0}^\tau p(\tau')\mathrm{d}\tau'\right]f(t)\mathrm{d}\tau\end{aligned}$$
- The above formula is complicated, though, so it is probably better to remember the method than to memorize the above.
- When $p(t)=a$ for all t, $y'(t)=ay+f(t)$. The solution is given by the **Duhamel formula**.
- **Duhamel formula**: The following equation, which solves ODEs of the form $y'(t)=ay+f(t)$. *Given by*
$$y(t)=\mathrm{e}^{a(t-t_0)}y_0+\int_{t_0}^t\mathrm{e}^{a(t-\tau)}f(\tau)\mathrm{d}\tau$$
  - We should understand the derivation, but we can apply the Duhamel formula on PSets and exams without further justification.
- Other things (??) are related to this form by some smart transformation.
- Final example of explicitly solvable ODEs: Linear autonomous systems.
- **Linear autonomous system**: A system of equations of the form $y'=Ay$ where A is a constant $n\times n$ matrix and y takes its value in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$).
- The homogeneous solution is
$$y(t)=\mathrm{e}^{tA}y_0$$
where $\mathrm{e}^{tA}=1+\frac{tA}{1!}+\frac{t^2A^2}{2!}+\cdots$.
- In the inhomogeneous case $y'=Ay+f(t)$, our solution is
$$y(t)=\mathrm{e}^{tA}y_0+\int_0^t\mathrm{e}^{(t-\tau)A}f(\tau)\mathrm{d}\tau$$

- We don't want to compute e^{tA} using an infinite power series. Thus, we introduce similarity.
- Let Q be the connecting matrix from the standard basis to the new basis. Then the matrix of Q is the set of new basis vectors q_1, q_2, q_3, i.e., $Q = \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix}$. Then $B = Q^{-1}AQ$ or $A = QBQ^{-1}$.
- We want B to be in the most convenient basis possible. Thus, we take the basis to be the Jordan basis.
- We fortunately have $\mathrm{e}^{tA} = Q\mathrm{e}^{tB}Q^{-1}$.
- Consider $\chi_A(z) = \det(zI_n - A)$ where $n = 2, 3$. If χ_A has distinct roots, then the eigenvalues of A are distinct. At this point, we can find an eigenvector corresponding to each eigenvalue and diagonalize our matrix.
- Alternatively, if χ_A has multiple roots...
  - 2×2 case, A is not diagonal. Then there is only one eigenvector v_λ. In this case, solve $(A-\lambda)u = v_\lambda$. Here, we say that the algebraic multiplicity is 2 and the geometric multiplicity is 1. Then

$$Q = \begin{pmatrix} v_\lambda & u \end{pmatrix} \qquad B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad \mathrm{e}^{tA} = Q\begin{pmatrix} \mathrm{e}^{t\lambda} & t\mathrm{e}^{t\lambda} \\ 0 & \mathrm{e}^{t\lambda} \end{pmatrix} Q^{-1}$$

  - 3×3 case: If we have λ of $\alpha_\lambda = 2$ and μ of $\alpha_\mu = 1$, or if we have λ with $\alpha_\lambda = 3$. First case: Check geometric multiplicity of λ, i.e., how many linearly independent v give $(A - \lambda I)v = 0$. If there is one, solve $(A - \lambda I)u = v_\lambda$. If there are more than one, A is diagonalizable. Second case: Check geometric multiplicity of λ. Divide into two subcases. If $\gamma_\lambda = 1$, then we need to solve $(A - \lambda I)u_1 = v_\lambda$ and $(A - \lambda I)u_2 = u_1$, and we get

$$Q = \begin{pmatrix} v_\lambda & u_1 & u_2 \end{pmatrix} \qquad B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

    If $\gamma_\lambda = 2$, then cleverly choose v_1 such that v_1 is in the column space of $A - \lambda I$. This will allow us to solve $(A - \lambda I)u = v_1$. Then

$$Q = \begin{pmatrix} v_1 & u & v_2 \end{pmatrix} \qquad B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

- For our linear autonomous system $y' = Ay$, λ is an eigenvector of A. Write $\lambda = \sigma + i\beta$. If $\lambda > 0$, then λ is **unstable** and the corresponding generalized eigenspace is said to be an **unstable eigenspace**.
- For example, if the JNF is

$$A = \left(\begin{array}{cc|c} 1 & 1 & \\ & 1 & \\ \hline & & -2 \end{array}\right)$$

  then the eigenspace corresponding to the upper block is said to be unstable, and the other one is said to be stable.
- Consider the vector $\mathrm{e}^{tA}v$. The entries consist of linear combinations of functions of the form $t^k\mathrm{e}^{t\lambda}$. If the real part is greater than zero, the solution grows exponentially fast in the t direction (notice how $t \to \infty$ implies $t^k\mathrm{e}^{t\lambda} \to \infty$). Otherwise, the solution decays exponentially fast (notice how $t \to \infty$ implies $t^k\mathrm{e}^{t\lambda} \to 0$).

4.3 Chapter 3: Linear Equations

From Teschl (2012).

Section 3.1: The Matrix Exponential

12/6:

- Note: This section (in the book's chronology) follows several others that we will only study later in the course. Thus, when terms are bolded but not defined here, they are likely callbacks to prior sections in the book. You can look up their definitions in the portion of these notes treating those sections.
- Herein, we will study the system
$$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$
where A is an $n \times n$ matrix.
- Teschl (2012) reviews how we take the matrix product, the definition of the scalar product, and the definition of the norm.
- Teschl (2012) denotes the identity matrix by $\mathbb{I}$.
- Deriving the solution to the above ODE.
  - We use the **Picard iteration** to derive a Taylor series that converges to the solution to the ODE by the **Picard-Lindelöf theorem**.
  - Start with $x_0(t) = x_0$ as our first approximation. The first few terms are
$$\begin{aligned}
x_0(t) &= x_0\\
x_1(t) &= x_0 + \int_0^t Ax_0(s)\,\mathrm{d}s = x_0 + Ax_0\int_0^t \mathrm{d}s = x_0 + tAx_0\\
x_2(t) &= x_0 + \int_0^t Ax_1(s)\,\mathrm{d}s = x_0 + Ax_0\int_0^t \mathrm{d}s + A^2x_0\int_0^t x\,\mathrm{d}s\\
&= x_0 + tAx_0 + \frac{t^2}{2}A^2x_0
\end{aligned}$$
  - This motivates an induction proof (omitted here and in Teschl (2012)), resulting in the formula
$$x_m(t) = \sum_{j=0}^{m} \frac{t^j}{j!}A^jx_0$$
  - As mentioned above, the Picard-Lindelöf theorem implies that the Taylor series converges and thus
$$x(t) = \lim_{m\to\infty} x_m(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!}A^jx_0$$
  - But this series converges to a variant of the **matrix exponential**, yielding
$$x(t) = \exp(tA)x_0$$
  - Therefore, "to understand our original problem, we have to understand the matrix exponential!" (Teschl, 2012, p. 60).
- **Matrix exponential** (of A): The following matrix. *Denoted by* $\boldsymbol{\exp(A)}$. *Given by*
$$\exp(A) = \sum_{j=0}^{\infty} \frac{1}{j!}A^j$$
- Going forward, we will work in $\mathbb{C}^n$ since $\mathbb{C}$ is algebraically closed (which will be important later on for JCF).
- For use later, we introduce a suitable norm for matrices and give a direct proof for the convergence of the above series in this norm. In particular...

- **Matrix norm** (of A): The following function from the space of $n \times n$ matrices to the nonnegative real numbers, where A is a complex matrix acting on $\mathbb{C}^n$. *Denoted by* $\|\boldsymbol{A}\|$. *Given by*

$$\|A\| = \sup_{x:|x|=1} |Ax|$$

- $\mathbb{C}^{n\times n}$: The vector space of $n \times n$ matrices over $\mathbb{C}$.
- Under the matrix norm, $\mathbb{C}^{n\times n}$ becomes a **Banach space**.
- An interesting formula.

$$\max_{1\leq 1,j\leq n} |A_{i,j}| \leq \|A\| \leq n \max_{1\leq i,j\leq n} |A_{i,j}|$$

  - Implication: A sequence of matrices converges in the matrix norm iff all matrix entries converge.
- Since $\left\|A^j\right\| \leq \|A\|^j$, convergence of the series defining $\exp(A)$ follows from convergence of

$$\sum_{j=0}^{\infty} \frac{\|A\|^j}{j!} = \exp(\|A\|)$$

- **Commutator** (of A, B): The following matrix, where A, B are matrices. *Denoted by* $[\boldsymbol{A}, \boldsymbol{B}]$. *Given by*

$$[A, B] = AB - BA$$

  - The commutator vanishes iff A, B commute.
- Lemma 3.1: Suppose A, B commute, i.e., $[A, B] = 0$. Then

$$\exp(A + B) = \exp(A)\exp(B)$$

- Suppose we perform a linear change of coordinates $y = U^{-1}x$. Then the matrix exponential in the new coordinates is given by

$$U^{-1}\exp(A)U = \exp\left(U^{-1}AU\right)$$

  - This follows from the definition of the matrix exponential, the fact that $U^{-1}A^jU = (U^{-1}AU)^j$ for arbitrary natural number powers j, and the fact that the matrix product is continuous with respect to the matrix norm (i.e., if $A_j \to A$ and $B_j \to B$, then $A_jB_j \to AB$).
  - Thus, to compute $\exp(A)$, we'd prefer a coordinate transform which renders A as simple as possible.
- Theorem 3.2 (Jordan canonical form): Let A be a complex $n \times n$ matrix. Then there exists a linear change of coordinates U such that A transforms into a block matrix

$$U^{-1}AU = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$$

  with each block of the form

$$J = \alpha\mathbb{I} + N = \begin{pmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \alpha & \ddots & \\ & & & \ddots & 1 \\ & & & & \alpha \end{pmatrix}$$

  Here, N is a matrix with ones in the first diagonal above the main diagonal and zeroes elsewhere.

- The numbers α are the eigenvalues of A. The new basis vectors u_j (the columns of U) consist of generalized eigenvectors of A.
- The general details of finding the JCF are quite cumbersome and deferred to Section 3.8. Typically, we use computers to do this nowadays.

- We can now compute the matrix exponential.
  - First, we break into Jordan blocks.

$$\exp\bigl(U^{-1}AU\bigr) = \begin{pmatrix} \exp(J_1) & & \\ & \ddots & \\ & & \exp(J_m) \end{pmatrix}$$

  - Since αI commutes with N, we infer from Lemma 3.1 that we can compute the matrix exponential of a Jordan block as follows.

$$\exp(J) = \exp(\alpha\mathbb{I})\exp(N) = \mathrm{e}^{\alpha}\sum_{j=0}^{k-1}\frac{1}{j!}N^j = \mathrm{e}^{\alpha}\begin{pmatrix} 1 & 1 & \frac{1}{2!} & \cdots & \frac{1}{(k-1)!} \\ & 1 & 1 & \ddots & \vdots \\ & & 1 & \ddots & \frac{1}{2!} \\ & & & \ddots & 1 \\ & & & & 1 \end{pmatrix}$$

    - We assume that J (and hence $\mathbb{I}, N$) is/are $k \times k$.
    - In the last step, we make use of the fact that N^j is a matrix with ones in the j^{th} diagonal above the main diagonal, and thus that N^j vanishes from when j reaches the size of N. Indeed, we could still sum all the way up to ∞; we'd just only be adding on zeroes matrices after the $k-1$ term.

- In two dimensions, the exponential of

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

  is given by

$$\exp(A) = \mathrm{e}^{\delta}\left(\cosh(\Delta)\mathbb{I} + \frac{\sinh(\Delta)}{\Delta}\begin{pmatrix} \gamma & b \\ c & -\gamma \end{pmatrix}\right)$$

  where

$$\delta = \frac{a+d}{2} \qquad \gamma = \frac{a-d}{2} \qquad \Delta = \sqrt{\gamma^2 + bc}$$

  - Special cases:
    - $\Delta = 0$. In this case, define

$$\frac{\sinh(\Delta)}{\Delta} = 1$$

    - Δ is purely imaginary. In this case, we have

$$\cosh(i\Delta) = \cos\Delta \qquad \frac{\sinh(i\Delta)}{i\Delta} = \frac{\sin(\Delta)}{\Delta}$$

  - Derivation: Given as in HW3 Q3.

- If A is in JCF, we can easily see that

$$\det(\exp(A)) = \exp(\operatorname{tr}(A))$$

  - Since the determinant and trace are invariant under coordinate change, this formula holds for arbitrary matrices, too.

- Lemma 3.3: A vector u is an eigenvector of A corresponding to the eigenvalue α iff u is an eigenvector of $\exp(A)$ corresponding to the eigenvalue e^{α}.

  Moreover, the Jordan structure of A and $\exp(A)$ are the same except for the fact that the eigenvalues of A which differ by a multiple of $2\pi i$ (as well as the corresponding Jordan blocks) are mapped to the same eigenvalue of $\exp(A)$. In particular, the geometric and algebraic multiplicity of e^{α} is the sum of the geometric and algebraic multiplicities of the eigenvalues which differ from α by a multiple of $2\pi i$.

  Proof. Given. □

- Teschl (2012) covers a method not described in class in which we can get an alternate **real Jordan canonical form**.
  - This method may explain the complex eigenvectors final step for that subset of planar autonomous systems.

Section 3.8: Appendix – Jordan Canonical Form

- Teschl (2012) has quite a bit to say on JCF!

Week 5

End Quantitative and Intro to Qualitative

5.1 Planar Autonomous Linear Systems

10/24:

- Review of vector fields.
- **Phase diagram**: A diagram that shows the qualitative behavior of an autonomous ordinary differential equation. *Also known as* **phase portrait**.

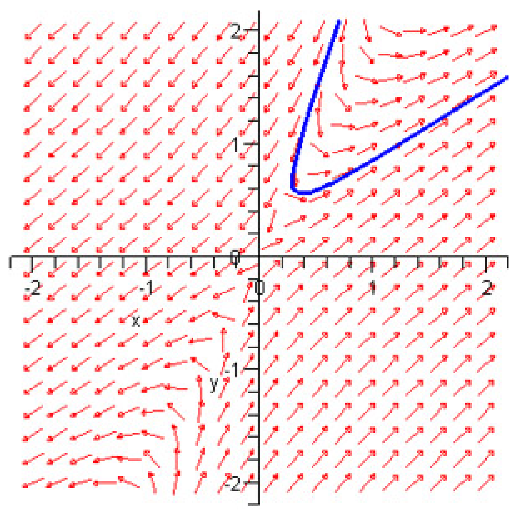
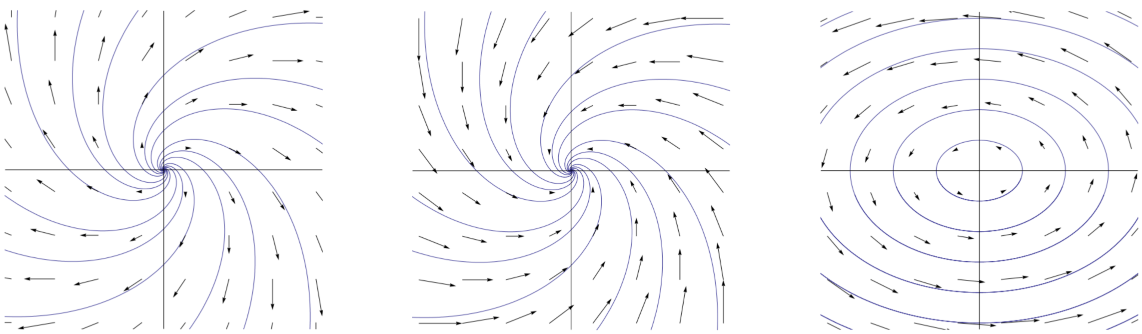


Figure 5.1: Phase diagram example.

  - Consists of a selection of arrows describing, to some extent, a vector field and is often paired with integral curves.
- Suppose $\Omega \subset \mathbb{R}^n$ is open.
- **Vector field** (on Ω): A mapping from $\Omega \to \mathbb{R}^n$. *Denoted by* $\boldsymbol{X}$.
  - Essentially, a vector field assigns to every point of some region a vector; the definition just formalizes this notion.
- **Flow**: A formalization of the idea of the motion of particles in a fluid.
  - The solution to the IVP $\frac{\mathrm{d}y}{\mathrm{d}t} = X(y)$, $y(0) = x$.

- If X is C^1, then for all $x \in \Omega$, there exists a unique solution y to the above IVP.
- **Orbit** (of x under X): The trajectory $y(t,x)$.
  - Recall that the tangent vector to any trajectory at any point coincides with the vector to which X maps that point.
- **Fixed point**: A point $x_0 \in \Omega$ such that $X(x_0) = \bar{0}$.
  - If x_0 is a fixed point, then the trajectory is $y(t) = x_0$.
- Today: We will consider flows on vector fields where the dimension is two and our vector field is linear. In particular...
- Let A be a 2×2 real matrix, and let $X(x) = Ax$.
  - In this case, $x_0 = 0$ is the only fixed point.
  - The flow is given by the linear differential equation $y' = Ay$, $y(0) = x$. The solution is $y(t) = \mathrm{e}^{tA}x$.
- Case 1: A has no real eigenvalues.



(a) $\sigma > 0$. (b) $\sigma < 0$. (c) $\sigma = 0$.

Figure 5.2: Phase diagrams for a planar system with no real eigenvalues.

  - We know that $\chi_A(z)$ is a real polynomial: $\chi_A(z) = z^2 + (\operatorname{trace} A)z + \det A$, and since A is real, both $\operatorname{trace} A$ and $\det A$ are real.
  - Thus, the eigenvalues appear as conjugate pair, i.e., we may write $\lambda = \sigma + i\beta$ and $\bar{\lambda} = \sigma - i\beta$.
    - $\alpha = \gamma = 1$ for both eigenvalues.
    - The eigenvectors must also be complex conjugates.
  - Distinct eigenvalues imply that A is diagonalizable.
  - However, this is not what we want because if we use the complex form, then

$$\mathrm{e}^{tA} = Q \begin{pmatrix} \mathrm{e}^{t\lambda} & 0 \\ 0 & \mathrm{e}^{t\bar{\lambda}} \end{pmatrix} Q^{-1}$$

  - Indeed, we want to get a real matrix out of $Q, \mathrm{e}^{t\Lambda}, Q^{-1}$ all complex. We have

$$\begin{aligned}
\mathrm{e}^{tA}x &= Q \begin{pmatrix} \mathrm{e}^{t(\sigma+i\beta)} & 0 \\ 0 & \mathrm{e}^{t(\sigma-i\beta)} \end{pmatrix} \underbrace{Q^{-1}x}_{z} \\
&= Q \begin{pmatrix} \mathrm{e}^{t(\sigma+i\beta)}z^1 \\ \mathrm{e}^{t(\sigma-i\beta)}z^2 \end{pmatrix} \\
&= z^1\mathrm{e}^{t(\sigma+i\beta)}v + z^2\mathrm{e}^{t(\sigma-i\beta)}\bar{v}
\end{aligned}$$

- Since $y(0) = x = z^1 v + z^2 \bar{v} \in \mathbb{R}^2$ (i.e., $z^1 v + z^2 v$ is *real*), we know that it is equal to its complex conjugate. This tells us that

$$
\begin{aligned}
z^1 v + z^2 \bar{v} &= \bar{z^1}\bar{v} + \bar{z^2} v \\
z^1 &= \bar{z^2}
\end{aligned}
$$

- It follows that

$$
\begin{aligned}
y(t) &= \mathrm{e}^{tA} x \\
&= z^1 \mathrm{e}^{t(\sigma + i\beta)} v + \bar{z^1} \mathrm{e}^{t(\sigma - i\beta)} \bar{v} \\
&= z^1 \mathrm{e}^{t(\sigma + i\beta)} v + \overline{z^1 \mathrm{e}^{t(\sigma + i\beta)} v} \\
&= 2\,\mathrm{Re}(z^1 \mathrm{e}^{t(\sigma + i\beta)} v) \\
&= 2\,\mathrm{Re}(z^1 \mathrm{e}^{\sigma t}(\cos(\beta t) + i\sin(\beta t))(v_1 + iv_2)) \\
&= 2\,\mathrm{Re}(z^1 \mathrm{e}^{\sigma t}(\cos(\beta t) v_1 + i\cos(\beta t) v_2 + i\sin(\beta t) v_1 - \sin(\beta t) v_2)) \\
&= 2\mathrm{e}^{\sigma t}\cos(\beta t)\cdot \mathrm{Re}(z^1 v) - 2\mathrm{e}^{\sigma t}\sin(\beta t)\cdot \mathrm{Im}(z^1 v)
\end{aligned}
$$

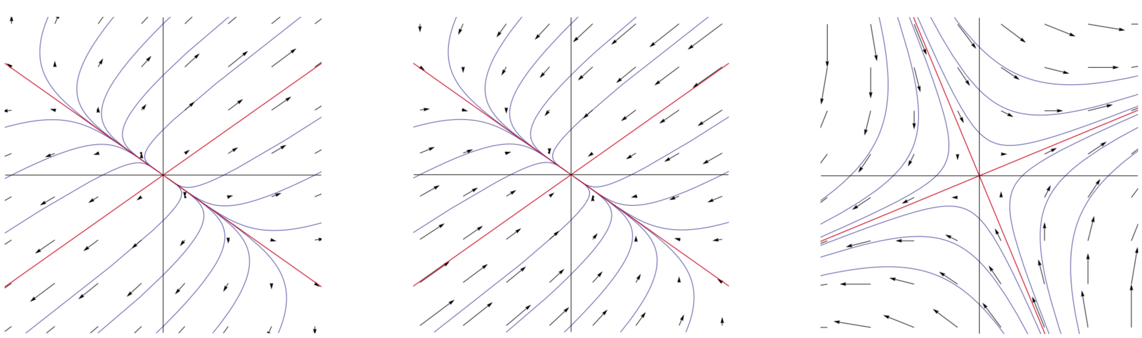
- How do we get from the second-to-last line to the last line above??
- Suppose $\sigma \neq 0$. Then

$$
x \mapsto \begin{pmatrix} \mathrm{Re}(z^1 v) \\ \mathrm{Im}(z^1 v) \end{pmatrix}
$$

  is a real linear transformation on $\mathbb{R}^2$.
- It follows that the trajectories are just spirals in the complex plane.
- If $\sigma > 0$, then the spiral repels from the origin. If $\sigma < 0$, then the spiral attracts to the origin. If $\sigma = 0$, we get an ellipse.
- Therefore, we have completely classified equations of the form

$$
\begin{pmatrix} y^1 \\ y^2 \end{pmatrix}' = \begin{pmatrix} y^2 \\ -\omega^2 y^1 \end{pmatrix}
$$

- Case 2: A has real eigenvalues and *is* diagonalizable.



(a) $\lambda_1, \lambda_2 > 0$.   (b) $\lambda_1, \lambda_2 < 0$.   (c) $\lambda_1 < 0 < \lambda_2$ (WLOG).

Figure 5.3: Phase diagrams for a diagonalizable planar system with real eigenvalues.

- Suppose $\lambda_1, \lambda_2 \in \mathbb{R}$ have corresponding linearly independent eigenvectors v_1, v_2.
- If we choose v_1, v_2 to be our basis, then

$$
\mathrm{e}^{tA} = Q \begin{pmatrix} \mathrm{e}^{t\lambda_1} & 0 \\ 0 & \mathrm{e}^{t\lambda_2} \end{pmatrix} Q^{-1}
$$

  where $Q = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$.

- Thus, as before, the solution may be expressed in the following form, where $z = Q^{-1}x$.

$$y(t) = \mathrm{e}^{tA}x = \mathrm{e}^{\lambda_1 t}z^1 v_1 + \mathrm{e}^{\lambda_2 t}z^2 v_2$$

- Moving forward, it will be convenient to work in the v_1, v_2 basis. We divide into three subcases ($\lambda_1, \lambda_2 > 0$ [Figure 5.3a], $\lambda_1, \lambda_2 < 0$ [Figure 5.3b], and WLOG $\lambda_1 < 0 < \lambda_2$ [Figure 5.3c]).
  1. Notice that

$$\mathrm{e}^{\lambda_2 t} = \mathrm{e}^{(\lambda_2/\lambda_1)(\lambda_1 t)}$$

     i.e., $\mathrm{e}^{\lambda_2 t}$ is a power of $\mathrm{e}^{\lambda_1 t}$. Thus, when the signs are the same, we get a power function $v_2 = v_1^{\lambda_2/\lambda_1}$.
     - Both subspaces v_1, v_2 are unstable here.
  2. If $\lambda_1, \lambda_2 < 0$, then we have the same trajectories, but they're all attracted to the origin instead of repelled.
     - Both subspaces v_1, v_2 are stable here.
  3. When both eigenvalues have different signs, we are considering power functions of a negative power.
     - The stable subspace is v_2 and the unstable subspace is v_1 here.

- Case 3: A has real eigenvalues and *is not* diagonalizable.

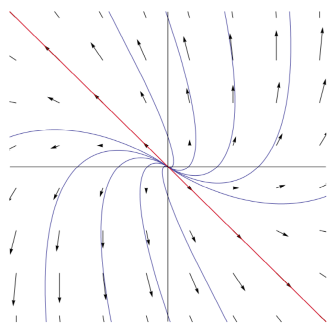


Figure 5.4: Phase diagrams for a nondiagonalizable planar system with real eigenvalues.

- In this case, the matrix exponential is given by

$$\mathrm{e}^{tA} = Q \begin{pmatrix} \mathrm{e}^{t\lambda} & t\mathrm{e}^{t\lambda} \\ 0 & \mathrm{e}^{t\lambda} \end{pmatrix} Q^{-1}$$

- The solution is given by

$$\mathrm{e}^{tA}x = (z^1\mathrm{e}^{t\lambda} + z^2 t\mathrm{e}^{t\lambda})v + z^2\mathrm{e}^{t\lambda}u$$

  where $Q^{-1}x = z$ again.
- In graphing, note that here we have (a distorted version of) the form $y = x \pm x\log x$:

$$y = (z^1\mathrm{e}^{t\lambda} + z^2 t\mathrm{e}^{t\lambda})\hat{\imath} + z^2\mathrm{e}^{t\lambda}\hat{\jmath}$$

  Define $x := \mathrm{e}^{t\lambda}$. Then $t = \lambda^{-1}\ln x$. Substituting, we have

$$\begin{aligned} &= (z^1 x + z^2(\lambda^{-1}\ln x)x)\hat{\imath} + z^2 x\hat{\jmath} \\ &= (z^1 x + z^2\lambda^{-1}x\ln x)\hat{\imath} + z^2 x\hat{\jmath} \end{aligned}$$

- When $\lambda > 0$, the whole space is unstable. Thus, the phase diagram is tangent to the origin.
- When $\lambda < 0$, the trajectories take the same form but this time are attracted to zero. In this case, the whole space is stable.

- We can take x_1 to x_2 iff they are in the same orbit. Conclusion: Orbits never cross.
- Takeaway: You should be able to compute the eigenvalues and eigenvectors and sketch these graphs.
- Shao will post lecture notes after today's lecture!
- Next lecture: The final explicitly solveable case, which is the driven harmonic oscillator.

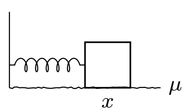
5.2 Driven Harmonic Oscillator and Resonance

10/26: • We are interested in the 2nd order constant coefficient equation

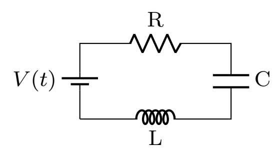
$$x'' + \mu x' + \omega_0^2 x = H_0 \mathrm{e}^{i\omega t}$$

where $\mu \geq 0$ and $\omega_0, \omega > 0$.

- Two cases where this ODE arises:



(a) A driven harmonic oscillator.



(b) An RLC circuit.

Figure 5.5: Origins of the driven harmonic oscillator equation.

1. The driven harmonic oscillator.
   - Consider a mass on a spring.
   - The extent of friction between the mass point and the surface is described by μ.
   - The oscillation is periodically driven by a force of magnitude $H_0 \cos \omega t$.
2. RLC circuit.
   - R is resistance, L is inductance, C is capacitance.
   - We have the laws

     $$LI_L'(t) = V_L \qquad CV_C'(t) = I_C \qquad I_R(t) = V_R(t)/R$$

     - Left: Self-inductance.
     - Right: Ohm's law.
   - Combining them with Kirchhoff's laws

     $$I(t) = I_R = I_C = I_L \qquad V(t) = V_R + V_L + V_C$$

     we get the RLC circuit equation

     $$LI'' + RI' + \frac{I}{C} = V'(t)$$

   - The most interesting cases is when we have a source of alternating current of frequency ω. In this case, $V(t) = V_0 \cos \omega t$ or, in the complex case, $V(t) = V_0 \mathrm{e}^{i\omega t}$. This yields the complex equation

     $$I'' + \frac{R}{L} I' + \frac{1}{LC} I = \frac{i\omega V_0}{L} \mathrm{e}^{i\omega t}$$

   - Here, the friction coefficient $\mu = R/L$ and the frequency is $\omega_0 = \sqrt{1/LC}$.

- Recall that we want to solve the following ODE.

$$x'' + \mu x' + \omega_0^2 x = H_0 \mathrm{e}^{i\omega t}$$

  - The homogeneous linear equation $x'' + \mu x' + \omega_0^2 x = 0$ is well-understood, i.e., we can find all of the *homogeneous* solutions to the above equation.
  - Thus, to solve the above inhomogeneous equation, we just have to find a particular solution.
- WLOG let $\omega > 0$.
- From the homework, a particular solution $x_p(t)$ with initial condition $x_p(0) = x_p'(0) = \mu = 0$ can be obtained from the Duhamel formula as follows.

$$x_p(t) = H_0 \int_0^t \frac{\sin \omega_0 (t - \tau)}{\omega_0} \mathrm{e}^{i\omega\tau} \mathrm{d}\tau$$

  - Substituting

$$\sin\theta = \frac{\mathrm{e}^{i\theta} - \mathrm{e}^{-i\theta}}{2i}$$

    into the above allows us to evaluate it.
  - In particular, it follows that

$$x_p(t) = \begin{cases} \dfrac{H_0}{\omega_0^2 - \omega^2}\left(\mathrm{e}^{i\omega t} - \cos\omega_0 t - \dfrac{i\omega}{\omega_0}\sin\omega_0 t\right) & \omega \neq \omega_0 \\ -\dfrac{iH_0}{2\omega_0}\left(t\mathrm{e}^{i\omega_0 t} - \dfrac{\sin\omega_0 t}{\omega_0}\right) & \omega = \omega_0 \end{cases}$$

    - We compute the $\omega = \omega_0$ case using L'Hôpital's rule to analyze the $\omega \neq \omega_0$ case as $\omega \to \omega_0$.
  - If we pump in energy at the same point that we have deviation ($\omega = \omega_0$), then the amplitude of oscillation goes to ∞.
    - Practically, when $\omega \approx \omega_0$, the long-time behavior of the driven oscillator will be very much like a growing oscillator.
    - Eventually, the amplitude will be approximately $(\omega - \omega_0)^{-1}$.
- **Resonance catastrophe**: Inputing energy into a system at its natural frequency, causing the total energy to grow until a mechanical failure occurs.
  - This is what happened at the Millenium Bridge in London; synchronized footsteps caused the bridge to shake really wildly.
- If $\mu > 0$, there will be a particular solution of the form

$$x_p(t) = A(\omega) H_0 \mathrm{e}^{i\omega t}$$

  - From HW1, we have three cases when $\mu > 0$: $0 < \mu < 2\omega_0$, $\mu = 2\omega_0$, and $\mu > 2\omega_0$. These are just the three cases of the characteristic polynomial??
  - Substituting the proposed form of the particular solution into the differential equation, we get

$$\begin{aligned} x_p'' + \mu x_p' + \omega_0^2 x_p &= H_0 \mathrm{e}^{i\omega t} \\ (-\omega^2 + i\omega\mu + \omega_0^2) H_0 A(\omega) \mathrm{e}^{i\omega t} &= H_0 \mathrm{e}^{i\omega t} \\ (-\omega^2 + i\omega\mu + \omega_0^2) A(\omega) &= 1 \\ A(\omega) &= \frac{1}{\omega_0^2 - \omega^2 + i\mu\omega} \end{aligned}$$

  - In theory, we avoid the resonance catastrophe in this case. In practice, however, when $\omega \to 0$, we still run into issues.

- For mass point:
$$|H_0A(\omega)| = \frac{|H_0|}{\sqrt{(\omega^2-\omega_0^2)^2+\mu^2\omega^2}}$$
  - The norm $|H_0A(\omega)|$ is maximized when $\omega_r = \sqrt{\omega_0^2+\mu^2/2}$.
  - $\omega_r \to \omega_0$ implies $\mu \to 0$??
- As for the argument/angle,
$$\arg(H_0A(\omega)) = \arg H_0 + \arg A(\omega)$$
- We consider $\omega : 0 \to \omega_0 \to +\infty$.
  - When $\omega = 0$, the complex amplitude is $1/\omega_0^2$ so it's a real number in the complex plane.
  - If ω is increased a bit, we get the reciprocal of a complex number. Norm is reciprocal, argument is negative.
  - For $\omega = \omega_0$, we have a purely imaginary number.
  - As $\omega \to \infty$, the argument approaches $-\pi$??
  - Showing the shape of the norm and the argument with respect to ω. This allows us to completely describe the resonance phenomena.

- For the RLC circuit, the discussion is a bit different.
  - The external voltage $V(t) = V_0\mathrm{e}^{i\omega t}$. Thus, $V'(t) = iV_0\omega\mathrm{e}^{i\omega t}$.
  - Here,
$$x_p(t) = \frac{iV_0\omega\mathrm{e}^{i\omega t}}{\omega_0^2-\omega^2+iR\omega/L}$$
  - Look at the complex amplitude.
    - Multiply the numerator and denominator by the inductance L to get
$$x_p(t) = \frac{iV_0\omega L\mathrm{e}^{i\omega t}}{L\omega_0^2-L\omega^2+iR\omega}$$
    - Then,
$$\text{Norm} = \frac{V_0L}{\sqrt{R^2+\left(\frac{1}{C\omega}-\frac{\omega}{L}\right)^2}}$$
    - For an RLC circuit, the resistance does not affect the resonance frequency.
$$\omega_r = \sqrt{\frac{1}{LC}} = \omega_0$$
    - If you have an external source of voltage, then you can vary the capacity of your circuit to ensure that the voltage will be maximized at a given frequency. We can tune our circuit to a very specific resonance frequency (this is used to filter our radio stations). The RLC circuit is only observable when the resonance coincides with the external resonance.

- There will be a bonus problem which is a PDE describing the vibration of a string.
  - Suppose we have a string with fixed endpoints, and suppose it is undergoing a small vibration.
  - Deviation from the equilibrium is described by a function $u(x,t)$.
  - The simplest equation we can derive is the 1D linear wave equation
$$\frac{\partial^2 u}{\partial t^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial x^2} = f(x,t)$$
    - c is the speed of the wave.

    - $f(x,t)$ is the given external force.
  - We can show that when $f(x,t) = 0$, then the vibration of the string is the linear supposition of infinitely many standing waves.

$$u(x,t) = \sum_{k=1}^{\infty} a_k \mathrm{e}^{\frac{\pi k t}{\ell}} \sin \frac{k\pi}{\ell} x$$

    - There are $k-1$ nodes in the string. These are called standing waves.
  - If you drive it with frequency

$$f(x,t) = \cos \omega t \sin \frac{k\pi}{\ell} x$$

    you encounter the resonance catastrophe.

- We are interested in the driven harmonic oscillator because it describes the vibrations, even of PDEs.
- This concludes our discussion of explicitly solvable differential equations.
- Those that are solvable by power series require complex analysis.
- Starting this Friday, we will talk about the qualitative theory of differential equations.
- Cauchy-Lipschitz this Friday.
- Next week: Continuous dependence on initial values and differentiation with respect to the parameter of this equation.
- After this, we will be able to compute classical examples in the theory of perturbations.
- We will be able to solve the procession of Mercury problem (which was the first experimental verification of general relativity).

5.3 Qualitative Theory of ODEs

10/28:

- First issue: Uniqueness — we want to be able to talk about *the* solution to the IVP.
- We will be considering the IVP $y'(t) = f(t,y)$, $y(t_0) = y_0$ for $y(t)$ an $\mathbb{R}^n$-valued function.
- To embed our rough outline of the Cauchy-Lipschitz theorem into analysis, we start with metric spaces.
- **Metric space**: A set and a metric. *Denoted by* $\boldsymbol{(X,d)}$.
- **Metric**: A function from $X \times X \to [0,+\infty)$ that satisfies the following three axioms. *Denoted by* $\boldsymbol{d}$.
  1. $d(x,y) = d(y,x)$.
  2. $d(x,y) \geq 0$ and $d(x,y) = 0$ iff $x = y$.
  3. Triangle inequality: $d(x,z) \leq d(x,y) + d(y,z)$.
- Examples:
  1. $\mathbb{R}^n$ with $d(x,y) = \|x-y\| = \sqrt{\sum_{j=1}^n |x^j - y^j|^2}$.
  2. Continuous functions $y : [a,b] \to \mathbb{R}^n$ with $d(y_1,y_2) = \|y_1 - y_2\| = \sup_{t\in[a,b]} |y_1(t) - y_2(t)|$.
- In Euclidean spaces, we have **completeness**.
- **Cauchy** (sequence): A sequence $\{x_n\} \subset X$ such that for all $\varepsilon > 0$, there exists $N \geq 0$ such that $d(x_m,x_n) < \varepsilon$ for all $m,n > N$.

- **Convergent** (sequence): A sequence $\{x_n\} \subset X$ for which there exists $x \in X$ such that
$$\lim_{n\to\infty} d(x, x_n) = 0$$
- **Complete** (metric space): A metric space (X, d) such that every Cauchy sequence is convergent.
- Theorem (Banach fixed point theorem): Let (X, d) be a complete metric space and let $\Phi : X \to X$ be a function for which there exists $q \in (0, 1)$ such that for all $x, y \in X$,
$$d(\Phi(x), \Phi(y)) \leq q \cdot d(x, y)$$
Then there exists a unique $x \in X$ such that $x = \Phi(x)$.

Proof. We first construct the desired fixed point x.

Pick any $x_0 \in X$. Inductively define $\{x_n\}$ by $x_{n+1} = \Phi(x_n)$, starting from $n = 0$. We will now show that $\{x_n\}$ is a Cauchy sequence. As a lemma, we will prove by induction that
$$d(x_j, x_{j+1}) \leq q^j \cdot d(x_0, x_1)$$
for all $j \in \mathbb{N}_0$. For the base case $j = 0$, equality evidently holds. Now suppose inductively that we have proven that $d(x_j, x_{j+1}) \leq q^j \cdot d(x_0, x_1)$; we want to prove the claim for $j + 1$. But we have that
$$\begin{aligned} d(x_{j+1}, x_{j+2}) &= d(\Phi(x_j), \Phi(x_{j+1})) \\ &\leq q \cdot d(x_j, x_{j+1}) \\ &\leq q \cdot q^j \cdot d(x_0, x_1) \\ &= q^{j+1} \cdot d(x_0, x_1) \end{aligned}$$
as desired.

It follows that
$$\begin{aligned} d(x_n, x_{n+m}) &\leq \sum_{k=0}^{m-1} d(x_{n+k}, x_{n+k+1}) && \text{Triangle inequality} \\ &\leq \sum_{k=0}^{m-1} q^{n+k} \cdot d(x_0, x_1) && \text{Lemma} \\ &= q^n(1 + q + \cdots + q^{m-1}) \cdot d(x_0, x_1) \\ &< q^n(1 + q + \cdots + q^{m-1} + \cdots) \cdot d(x_0, x_1) \\ &= \frac{q^n}{1-q} \cdot d(x_0, x_1) \end{aligned}$$
It follows that the above term will converge to zero as $n \to \infty$, so $\{x_n\}$ is a Cauchy sequence and there exists an x such that $x_n \xrightarrow{d} x$.

We now prove that x is a fixed point of Φ, i.e., that $\Phi(x) = x$. We have that
$$\begin{aligned} d(x, \Phi(x)) &\leq d(x, x_n) + d(x_n, \Phi(x_n)) + d(\Phi(x_n), \Phi(x)) \\ &\leq d(x, x_n) + d(x_n, x_{n+1}) + q \cdot d(x_n, x) \\ &= (1 + q) \cdot d(x, x_n) + d(x_n, x_{n+1}) \end{aligned}$$
where the first term converges since $\{x_n\}$ is convergent and the second term converges since $\{x_n\}$ is Cauchy. Thus, $d(x, \Phi(x)) \to 0$ as $n \to \infty$, so $x = \Phi(x)$, as desired.

Lastly, we prove that x is unique. Suppose that there exists $y \in X$ such that $y = \Phi(y)$. Then
$$d(x, y) = d(\Phi(x), \Phi(y)) \leq q \cdot d(x, y)$$
It follows that $d(x, y) \leq q^n \cdot d(x, y)$, i.e., that $d(x, y) \to 0$ as $n \to \infty$. Therefore, we must have that $d(x, y) = 0$, from which it follows that $x = y$, as desired. $\square$

- Notes on the Banach fixed point theorem.
  - Φ is a **contraction**.
  - Shao gives the example of crumpling a sheet of paper (more specifically, dropping a map of a park in that park; a point coincides).
- Example: Fixed point of the cosine function.
  - Define $\{x_n\}$ by $x_{n+1} = \cos x_n$. If $x_0 \in \mathbb{R}$, then $x_1 \in [-1, 1]$ and $x_2 \in [\cos 1, 1]$.
  - Thus, while cosine is not a contraction on the real numbers ($\cos'(-\pi/2) = 1$, for example), we can show that $\cos : [\cos 1, 1] \to [\cos 1, 1]$ is a contraction: If $x, y \in [\cos 1, 1]$, then

$$
\begin{aligned}
|\cos x - \cos y| &= \left| \int_y^x -\sin t \, \mathrm{d}t \right| \\
&\leq |x - y| \sup_{t \in [\cos 1, 1]} |\sin t| \\
&\leq (\sin 1)|x - y|
\end{aligned}
$$

  - Thus, cosine has a fixed point at the intersection of $y = \cos x$ and $y = x$ of approximate value 0.739...
  - Overall, this is a pretty bad example, though.
- Theorem: Let $y_k : [a, b] \to \mathbb{R}^n$ be a Cauchy sequence of continuous functions under the sup norm. Then the limit exists and is continuous.
  - The proof is based on uniform convergence, which we've encountered before in analysis.
  - It follows that $C[a, b]$ (the metric space of all continuous functions on $[a, b]$) is complete.
  - If $\{y_k\} \subset \bar{B}(0, M)$[1], then the limit y is in $\bar{B}(0, M)$.
- Let's return to our ODE $y'(t) = f(t, y(t))$, $y(t_0) = y_0 \in \mathbb{R}^n$.
- We now have the tools to prove the Cauchy-Lipschitz theorem, and we will presently build up to that.
- Although we do not typically think of it this way, f is still a function with a domain and range. In particular, its domain is the set of ordered pairs where the first entry is a real number and the second entry is an element of the range of y, i.e., an element of $\mathbb{R}^n$. Thus, to begin, we are allowed to impose the following conditions on f.
  - Let $f(t, z)$ be defined on $[t_0, t_0 + a] \times \bar{B}(y_0, b)$ for some $a, b \in \mathbb{R}_+$ (we will put further constraints on the values of a, b later).
  - On this domain, suppose $|f|$ is bounded by some $M \in \mathbb{R}$, i.e., $|f(t, z)| \leq M$ for all t, z in the above set.
  - Let f be Lipschitz continuous in the second argument. In particular, there exists $L > 0$ such that $|f(t, z_1) - f(t, z_2)| \leq L|z_1 - z_2|$ for any $z_1, z_2 \in \mathbb{R}^n$.
- We usually consider a given ODE in differential form. However, there's no reason we can't consider the equivalent integral equation

$$
y(t) = y_0 + \int_{t_0}^t f(\tau, y(\tau)) \mathrm{d}\tau
$$

  - The reason for this change of perspective will become apparent shortly.

[1] $\bar{B}(0, M)$ denotes the set of all functions $y : [a, b] \to \mathbb{R}^n$ with sup norm at most M; topologically, it is the closed ball of radius M centered at the origin in $C[a, b]$.

- Let $\Phi : C[t_0, t_0 + a] \to C[t_0, t_0 + a]$ map functions to functions. Specifically, let it send

$$y(t) \mapsto y_0 + \int_{t_0}^{t} f(\tau, y(\tau))\mathrm{d}\tau$$

  - We denote this by writing $\Phi[y](t) = y_0 + \int_{t_0}^{t} f(\tau, y(\tau))\mathrm{d}\tau$.
  - Notice that the solution of our IVP is exactly the point of $C[t_0, t_0 + a]$ fixed by Φ because

$$y(t) = y_0 + \int_{t_0}^{t} f(\tau, y(\tau))\mathrm{d}\tau \quad \Longleftrightarrow \quad y = \Phi[y]$$

  - *This* motivates all steps taken thus far.
- All that remains is to show that Φ is a contraction on some subset of $C[t_0, t_0 + a]$. Then we can apply the Banach fixed point theorem.
- We first identify this subset. Let

$$X_b = \{y : [t_0, t_0 + a] \to \bar{B}(y_0, b)\}$$

  - By the previous theorem, this is a complete metric space.
  - We now want to relate a and b so that $\Phi(X_b) \subset X_b$ and Φ is a contraction.
  - For $\Phi(X_b) \subset X_b$, we need

$$\|\Phi[y] - y_0\| \leq \int_{t_0}^{t_0+a} |f(\tau, y(\tau))|\mathrm{d}\tau \leq a \cdot M \leq b$$

    so we want $a < b/M$.
  - Moreover, if Φ is to be a contraction, then since

$$\begin{aligned}\|\Phi[y_1] - \Phi[y_2]\| &\leq \int_{t_0}^{t+a} |f(\tau, y_1(\tau)) - f(\tau, y_2(\tau))|\mathrm{d}\tau \\ &\leq La \cdot \|y_1 - y_2\|\end{aligned}$$

    we want $La \in (0, 1)$. We can achieve this by requiring $a < 1/2L$.
  - Thus, choosing

$$a < \min\left(\frac{1}{2L}, \frac{b}{M}\right)$$

    accomplishes all of our goals.
- Therefore, by the Banach fixed point theorem, there exists a unique y such that $y = \Phi[y]$.
  - As we have already remarked, this fixed point is exactly the aforementioned solution to the IVP.
- Conclusion:
- Theorem (Cauchy-Lipschitz theorem): Let $f(t, z)$ be defined on an open subset $\Omega \subset \mathbb{R}_t \times \mathbb{R}_z^n$, $(t_0, y_0) \in \Omega$, such that f is Lipschitz continuous wrt. t, z in some neighborhood of (t_0, y_0). Then the IVP $y'(t) = f(t, y(t))$, $y(t_0) = y_0$ has a unique solution for some $T > 0$ on $[t_0, t_0 + T]$ such that $y(t)$ does not escape that neighborhood.
- If $f \in C_1$, then

$$|f(t, z_1) - f(t, z_2)| \leq \sup_{z \in \bar{B}(y_0, r)} \left\| \frac{\partial f}{\partial z}(t, z) \right\| \cdot |z_1 - z_2|$$

- We use the finite increment theorem of differential calculus to prove that f is Lipschitz continuous if it's continuously differentiable.
- The norm on the RHS above is the matrix norm.
- We have $y(t) = \int_{t_0}^t f(\tau, y(\tau))\mathrm{d}\tau + y_0$ and we use the Banach fixed point theorem (which is proved constructively).
- We have $y_{n+1}(t) = y_0 + \int_{t_0}^t f(\tau, y_n(\tau))\mathrm{d}\tau$.
  - Thus, the Picard iteration is justified by the Banach fixed point theorem.
- We do not use the algorithm from the proof (the Picard iteration) computationally; we use the polygon algorithm. This algorithm is only of theoretical significance.
- **Interval of existence**: The union of intervals containing the interval $[t_0, t_0 + T]$ on which the IVP has a solution.
  - The interval of existence is always open. If $t_0 \in I$ such that $y(t_0) = y$, then $y'(t) = f(t, y(t))$, $y(t_0) = y$.
  - Note that the theorem does not predict when singularity can occur.
  - Example: The interval of existence will always be $x' = 1 + x^2$, $x(t_0) = x_0$. Then $x(t) = \tan(t - t_0 + \arctan(x_0))$. The length of existence is always π.
- Interval of existence: If you consider the IVP $y'(t) = f(t, y(t))$, $y(t_0) = y_0$, then $[t_0, t_0 + T_1], [t_0 + T_1, t_0 + T_2]$. The first is of length T_1, and the second of length $T_2 - T_1$. Continuing on, we get $T_n - T_1$ so that $T_n \to \infty$ or T_n is bounded. This gives us the maximal solution/interval of existence.
- The motherfucker (Shao) made us stay 10 minutes late.

5.4 Chapter 1: Introduction

From Teschl (2012).

Section 1.5: Qualitative Analysis of First-Order Equations

11/15:

- Only a few ODEs are explicitly solvable. However, in many situations, only certain qualitative aspects of the solution (e.g., whether or not it stays within a certain region, what it looks like for large t, etc.) are of interest.
  - "Moreover, even in situations where an exact solution can be obtained, a qualitative analysis can give a better overview of the behavior than the formula for the solution" (Teschl, 2012, p. 20).
- Example: Qualitative analysis of a model of logistic growth.

$$\dot{x}(t) = (1 - x(t))x(t) - h$$

  - Plot the parabola $f(x) = (1 - x)x - h$.
  - The sign of f tells us what direction the solution will move.
  - We divide into three cases ($0 < h < 1/4$, $h = 1/4$, and $h > 1/4$)
  - $0 < h < 1/4$:
    - There are two unique zeroes of the parabola, namely $x_{1,2} = 0.5(1 \pm \sqrt{1 - 4h})$ where we choose $x_1 < x_2$.
    - If $x_0 < x_1$, $f(x_0) < 0$, and the solution will decrease and converge to $-\infty$.
    - If $x_0 = x_1$ (resp., x_2), then the solution will stay fixed for all t, i.e., $x(t) = x_0$ is the solution.

- If $x_1 < x_0 < x_2$, the solution will increase and converge to x_2.
- If $x_0 > x_2$, the solution will decrease and converge to x_2.
- Note how this mirrors our understanding of the logistic growth model, as well as the discussion from Lecture 2.1.

- What we have seen in the above example motivates the following lemma.
- Lemma 1.1: Consider the first-order autonomous initial value problem

$$\dot{x} = f(x), \quad x(0) = x_0$$

  where $f \in C(\mathbb{R})$ is such that the solutions are unique.

  1. If $f(x_0) = 0$, then $x(t) = x_0$ for all t.
  2. If $f(x_0) < 0$, then $x(t)$ converges to the first zero left of x_0. If there is no such zero, the solution converges to $-\infty$.
  3. If $f(x_0) > 0$, then $x(t)$ converges to the first zero right of x_0. If there is no such zero, the solution converges to $+\infty$.

- Teschl (2012) qualitatively analyzes a Riccati type ODE; I would need to study the previous subsection to understand this in full depth.
- Observations from a Mathematica-facilitated numerical analysis:
  - Mathematica complains about the step size getting too small on one side of the interval, so the solution only exists for finite time.
  - There is symmetry with respect to the transformation $(t, x) \mapsto (-t, -x)$; thus, we only need to consider $t \geq 0$.
  - For different sets of initial conditions, solutions never cross; this is a consequence of uniqueness (see Lecture 5.1).
  - There seem to be two cases: either the solution escapes to $+\infty$ in finite time or it converges to the line $x = -t$.
  - We need further proof, however, to verify that these observations are true and not just numerical glitches. In other words, Mathematica can guide our intuition, but we must still do the hard math ourselves.
- A more in-depth qualitative sign analysis of the same equation follows to verify the above claims.
- The extent to which this qualitative analysis is covered, though, goes well-beyond anything mentioned in class.
- The final theorem could be useful, though again, it is never covered.

Section 1.6: Qualitative Analysis of First-Order Periodic Equations

- Again, largely beyond the scope of class.
- **Periodic** (ODE): An ODE for which $f(t+1, x) = f(t, x)$, where we take the period to be 1 WLOG.

5.5 Chapter 2: Initial Value Problems

From Teschl (2012).

Section 2.1: Fixed Point Theorems

- Starting now, we are going to follow a brief tangent into linear algebra and mathematical analysis. The purpose of this is to facilitate "an easy and transparent proof of our basic existence and uniqueness theorem" (Teschl, 2012, p. 35).
- **Vector space**: *Definition assumed. Denoted by* $\boldsymbol{X}$.
- **Norm** (on X): A map from $X \to [0,\infty)$ satisfying the following requirements. *Denoted by* $||\cdot||$.
  1. $\|0\| = 0$, $\|x\| > 0$ for $x \in X \setminus \{0\}$.
  2. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$ and $x \in X$.
  3. $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (**triangle inequality**).
- **Inverse triangle inequality**: The following inequality, which holds for all $x, y \in X$. *Given by*

$$|\|x\| - \|y\|| \leq \|x - y\|$$

- **Normed vector space**: A vector space X along with a norm on X. *Also known as* **normed space**. *Denoted by* $\boldsymbol{(X, ||\cdot||)}$.
- **Convergent** (sequence of vectors): A sequence of vectors $\{f_n\}$ such that

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

  for some $f \in X$. *Denoted by* $\boldsymbol{f_n \to f}$.
- **Limit** (of a convergent sequence of vectors): The vector f in the above definition.
- **Continuous** (mapping between two normed spaces): A function $F : X \to Y$, where X, Y are normed spaces, such that $f_n \to f$ implies $F(f_n) \to F(f)$.
- Example: The norm, vector addition, and scalar multiplication are all continuous under the above definition.
  - Problem 2.2 walks us through this.
- **Cauchy** (sequence): Defined as in class.
- **Complete** (vector space): A vector space for which every Cauchy sequence has a limit.
- **Banach space**: A complete normed space.
- Example: $\mathbb{R}^n$ and $\mathbb{C}^n$ are Banach spaces under the usual **Euclidean norm**.
- **Euclidean norm**: The following norm, applicable to $\mathbb{R}^n$ and $\mathbb{C}^n$. *Given by*

$$|x| = \sqrt{\sum_{j=1}^{n} |x_j|^2}$$

- We will mainly be interested in the following example of a Banach space.
- Example: The set of continuous function $C(I)$ on $I \subset \mathbb{R}$ a compact interval.
  - A vector space if addition and scalar multiplication are defined pointwise on the functions.
  - A normed space under

$$\|x\| = \sup_{t\in I} |x(t)|$$

    - Problem 2.3 walks us through verifying that the above norm satisfies the three axioms.

- A complete vector space because...
  - $\{x_n\}$ Cauchy implies that $x_n \to x$ pointwise. But is $x \in C(I)$?
  - Well, $x_n \to x$ implies

$$\lim_{n\to\infty} \sup_{t\in I} |x_n(t) - x(t)| = \lim_{n\to\infty} \|x_n - x\| = 0$$

  - The above equation implies that x_n converges **uniformly** to x.
  - Since $x_n \to x$ uniformly and each x_n is continuous, we have by the Uniform Limit Theorem[2] that x is continuous, i.e., $x \in C(I)$ as desired.

- **Uniformly convergent** (sequence of functions): A convergent sequence of functions $\{x_n\}$ such that for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then

$$|x_n(t) - x(t)| < \varepsilon$$

  for all $t \in I$.

- **Endofunction**: A function from some set to the same set. *Denoted by* $\boldsymbol{K : C \to C}$.
  - Note that we do not require that K be bijective, only that it map a set to (some part of) itself.

- **Fixed point** (of an endofunction): An element $x \in C$ such that $K(x) = x$.

- **Contraction**: An endofunction K for which there exists a constant $\theta \in [0, 1)$ such that

$$\|K(x) - K(y)\| \leq \theta \|x - y\|$$

  for all $x, y \in C$.

- Denoting function composition.
  - We inductively define $K^0(x) = \mathrm{id}(x) = x$, $K^n(x) = K(K^{n-1}(x))$.

- Theorem 2.1 (Contraction principle): Let C be a (nonempty) closed subset of a Banach space X and let $K : C \to C$ be a contraction. Then K has a unique fixed point $\bar{x} \in C$ such that

$$\|K^n(x) - \bar{x}\| \leq \frac{\theta^n}{1-\theta} \|K(x) - x\|$$

  for all $x \in C$.

  Proof. As in class.

  Note that we use the closed-ness hypothesis to guarantee that C is still complete, i.e., that $\bar{x} \in C$ (we could very well define K on $C \setminus \{\bar{x}\}$, and it would still be a contraction). ☐

- Note that the contraction principle and the Banach fixed point theorem from class are the same statement; the contraction principle just states the inequality used in the proof of the Banach fixed point theorem as an additional result.

[2] Theorem 17.6 from Honors Calculus IBL.

Section 2.2: The Basic Existence and Uniqueness Result

- We now prove the basic existence and uniqueness result, using the theory from the previous section.
- **Initial value problem**: A mathematical question for which the desired solution is a function that satisfies both a differential equation and maps a specific point of its domain to a specific point of its codomain. *Also known as* **IVP**. *Given by*

$$\dot{x} = f(t,x), \quad x(t_0) = x_0$$

- For the purposes of our study, we will suppose that $f \in C(U, \mathbb{R}^n)$, where $U \subset \mathbb{R}^{n+1}$ is open, and that $(t_0, x_0) \in U$.
- To begin, integrate the ODE as follows.

$$\begin{aligned}
\int_{x_0}^{x} \mathrm{d}\chi &= \int_{t_0}^{t} f(s, x(s)) \, \mathrm{d}s \\
x(t) - x_0 &= \int_{t_0}^{t} f(s, x(s)) \, \mathrm{d}s \\
x(t) &= x_0 + \int_{t_0}^{t} f(s, x(s)) \, \mathrm{d}s
\end{aligned}$$

- **Integral equation**: A form of an ODE that reexpresses the equation in terms of (the) integral(s) of a function instead of its deriavative(s).
  - The last line above is the integral equation corresponding to the IVP ODE.
- How changing to the integral equation helps.
  - Notice that $x_0(t) = x_0$ is an approximate solution for t close to t_0.
  - Plugging $x_0(t)$ into the integral gives another approximate solution

    $$x_1(t) = x_0 + \int_{t_0}^{t} f(s, x_0(s)) \, \mathrm{d}s$$

  - In fact, as we will soon show, $x_1(t)$ is a "better" approximation than $x_0(t)$. To this end, iterating the procedure infinitely many times will eventually get you to $x(t)$. Indeed, if we recursively define a sequence of functions

    $$x_m(t) = K^m(x_0)(t)$$

    where

    $$K(x)(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) \, \mathrm{d}s$$

    we can show that the limit is the desired solution.
  - Since the limit will necessarily satisfy $K(x) = x$, the tack we take is proving that K is a contraction on some subset of $C(I)$ (which we have proven is a Banach space) and invoking the contraction principle.
  - At this point, the previous section should be fully motivated.
- Going forward, we will take $t_0 = 0$ and only consider the case where $t \geq 0$; the other cases can be handled analogously (but require considerably more absolute value symbols).
- Take

  $$X = C([0,T], \mathbb{R}^n)$$

  to be our Banach space for some suitable T (we will put constraints on the value of T as we build the rest of our argument).

- We need some $C \subset X$ on which to define K as an endofunction. Let's try
$$C = \overline{B_\delta(x_0)}$$
  - In words, we let C be the closed ball of radius δ surrounding the constant function $x_0(t)$ in $X = C([0,T],\mathbb{R}^n)$.
    - Geometrically, C is the set of all functions $f : [0,T] \to \mathbb{R}^n$ such that the vector $f(t) - x_0(t) \in \mathbb{R}^n$ is within δ of the origin for all $t \in [0,T]$.
    - Notice how the definition of the sup norm on $C(I)$ motivates the above geometric picture.
  - This will end up working, as long as we have a suitable relation (to be derived) between T and δ.
- Constraints on f.
  - By hypothesis, $(0,x_0) \in U$, where we recall that U is the domain of f.
  - We additionally require that U contains $V = [0,T] \times \overline{B_\delta(x_0)}$.
  - Since V is compact and f is continuous, we have by the Extreme Value Theorem in $\mathbb{R}^n$[3] that f has a maximum M on V. In particular, there exists
$$M = \max_{(t,x)\in V} |f(t,x)|$$
  - When we later seek to prove that K is a contraction, it will be useful to know that f is **locally Lipschitz continuous** in the second argument and **uniformly Lipschitz continuous** in the first argument. In particular, for every compact $V_0 \subset U$, the **Lipschitz constant** (which depends on V_0) is finite.
- **Lipschitz continuous** (function): A function $f : X \to Y$, where X,Y are normed spaces, such that the following number is finite.
$$L = \sup_{x_1 \neq x_2 \in X} \frac{|f(x_1) - f(x_2)}{|x_1 - x_2|}$$
Also known as **uniformly Lipschitz continuous**.
  - Lipschitz continuity is a **strong** form of **uniform continuity**.
- **Lipschitz constant** (of a Lipschitz continuous function): The following value, where f is Lipschitz continuous. *Denoted by* $\boldsymbol{L}$. *Given by*
$$L = \sup_{x_1 \neq x_2 \in X} \frac{|f(x_1) - f(x_2)}{|x_1 - x_2|}$$
  - For functions in $C(I)$, L corresponds to the smallest bound on the slope of f.
- **Uniformly continuous** (function): A function $f : X \to Y$, where X,Y are normed spaces, such that for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x_1,x_2 \in X$, $|x_1 - x_2| < \delta$ implies
$$|f(y) - f(x)| < \varepsilon$$
- **Strong** (condition): A condition that implies some other condition that does not, in turn, imply the original condition.
  - For example, all Lipschitz continuous functions are uniformly continuous, but not all uniformly continuous functions are Lipschitz continuous.
- **Locally Lipschitz continuous** (function): A function $f : X \to Y$, where X,Y are normed spaces, such that for every compact $X_0 \subset X$, $f : X_0 \to Y$ is Lipschitz continuous.

[3]Theorem 18.44 from Honors Calculus IBL.

- **Graph** (of $x \in C$): The set defined as follows. *Denoted by* $\boldsymbol{G(x)}$. *Given by*

$$G(x) = \{(t, x(t)) \mid t \in [0, T]\}$$

- Proving that $K : C \to C$.
  - At this point, we can compute

$$\begin{aligned} |K(x)(t) - x_0| &= \left| \int_0^t f(s, x(s)) \, \mathrm{d}s \right| \\ &\leq \int_0^t |f(s, x(s))| \, \mathrm{d}s && \text{Theorem 13.26}^{[4]} \\ &\leq tM && \text{Theorem 13.27}^{[4]} \end{aligned}$$

    for any x satisfying $G(x) \subset V$.
  - Thus, if we take

$$T \leq \frac{\delta}{M}$$

    then $|K(x)(t) - x_0| \leq TM \leq \delta$ for all $t \in [0, T]$.
  - It follows that under this definition of T, $\|K(x) - x_0\| \leq \delta$, so $K(x) \in \overline{B_\delta(x_0)}$ for all x with graph in V.
  - In the special case $M = 0$ (which would imply $T = \infty$), we may take T to be some arbitrary positive real number.

- Proving that K is a contraction.
  - To estimate $|K(x)(t) - K(y)(t)|$, we finally invoke the Lipschitz continuity constraint on f. In particular, we have

$$\begin{aligned} |K(x)(t) - K(y)(t)| &\leq \int_0^t |f(s, x(s)) - f(s, y(s))| \, \mathrm{d}s \\ &\leq L \int_0^t |x(s) - y(s)| \, \mathrm{d}s \\ &\leq Lt \sup_{0 \leq s \leq t} |x(s) - y(s)| \\ &\leq LT\|x - y\| \end{aligned}$$

  - Thus, if we take

$$T < \frac{1}{L}$$

    we have that K is a contraction.

- Having two constraints on the value of T, we may now formally take

$$T = \min\left(\frac{\delta}{M}, \frac{1}{2L}\right)$$

  - Note that this definition satisfies $T \leq \delta/M$ and $T < L^{-1}$.
  - If either of $M, L = 0$, we understand that that constraint no longer matters and we need only consider the other one. If $M = L = 0$, then we may take T to be any positive real number.

- Having defined a contraction $K : C \to C$, the existence and uniqueness of our ODE follows from the contraction principle:

[4] From Honors Calculus IBL.

- Theorem 2.2 (Picard-Lindelöf): Suppose $f \in C(U, \mathbb{R}^n)$, where $U \subset \mathbb{R}^{n+1}$ is open and $(t_0, x_0) \in U$. If f is locally Lipschitz continuous in the second argument, uniformly with respect to the first, then there exists a unique local solution $\bar{x}(t) \in C^1(I)$ of the IVP

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

  where I is some interval around t_0.

  More specifically, if $V = [t_0, t_0 + T] \times \overline{B_\delta(x_0)} \subset U$ and M denotes the maximum of $|f|$ on V, then the solution exists for at least $t \in [t_0, t_0 + T]$ and remains in $\overline{B_\delta(x_0)}$, where

$$T = \frac{\delta}{M}$$[5]

  The analogous result holds for the interval $[t_0 - T, t_0]$.

- **Picard iteration**: The procedure for finding the solution, involving calculating successive elements of the sequence $\{x_m\}$ as defined above by $x_m(t) = K^m(x_0)(t)$.
  - The Picard iteration is useful for proofs, such as the above one, but it is not suitable for actually finding the solution since the integrals are not computable in general except by tedious numerical methods.
  - If $f(t, x)$ is **analytic**, however, then $x_m(t)$ equals the Taylor expansion of the solution $\bar{x}(t)$ about t_0 up to order m. This can be used for numerical computations.
    - Problem 4.4 walks us through this.
- **Analytic** (function): A function that is locally given by a convergent power series.
- $f \in C^1(U, \mathbb{R}^n)$ implies that f is locally Lipschitz continuous in the second argument, uniformly with respect to the first.
  - Problem 2.5 walks us through this.
  - It follows from this statement that if $f \in C^1(U, \mathbb{R}^n)$, then the corresponding IVP $\dot{x} = f(t, x)$, $x(t_0) = x_0$ has a unique solution as per the Picard-Lindelöf theorem.
- This observation can be slightly extended via the following lemma.
- Lemma 2.3: Suppose $f \in C^k(U, \mathbb{R}^n)$ for some $k \geq 1$, where $U \subset \mathbb{R}^{n+1}$ is open and $(t_0, x_0) \in U$. Then the local solution $\bar{x}$ of the corresponding IVP is $C^{k+1}(I)$.

  Proof. We induct on k. For the base case $k = 1$, $\bar{x}(t) \in C^1$ by the Picard-Lindelöf theorem. Additionally, since $f \in C^1$, we have that $\dot{\bar{x}}(t) = f(t, \bar{x}(t)) \in C^1$. Thus, since the derivative of $\bar{x}$ is continuously differentiable, $\bar{x}(t)$ must be *twice* continuously differentiable, or C^2. The inductive step is very similar. □

Section 2.3: Some Extensions

- Extensions of the Picard-Lindelöf theorem that weren't covered in class.

5.6 Chapter 3: Linear Equations

From Teschl (2012).

[5] We are not missing the Lipschitz constraint here; indeed, it is superfluous, as will be shown in the next section.

Section 3.2: Linear Autonomous First-Order Systems

12/6:

- The solutions of an autonomous linear first-order system are given by $x(t) = \exp(tA)x_0$.
  - Rigorous perspective: $\exp(tA)$ is an isomorphism between all initial conditions x_0 and all solutions.
  - Implication: The set of solutions is a vector space isomorphic to $\mathbb{R}^n$ (resp. $\mathbb{C}^n$).
- To understand the dynamics of the solution, we need to understand $\exp(tA)$.
- We do so starting with the case of two real dimensions (which covers all prototypical cases).
- Teschl (2012) does exactly what we did in class in Lecture 5.1.
- **Source**: A descriptor applied to the origin of the real plane when both eigenvalues of the planar linear autonomous first-order system in question have *positive* real part, and thus all solutions grow exponentially as $t \to +\infty$ and decay exponentially as $t \to -\infty$.
- **Sink**: A descriptor applied to the origin of the real plane when both eigenvalues of the planar linear autonomous first-order system in question have *negative* real part, and thus all solutions decay exponentially as $t \to +\infty$ and grow exponentially as $t \to -\infty$.
- **Saddle**: A descriptor applied to the origin of the real plane when one eigenvalues of the planar linear autonomous first-order system in question has positive real part and the other has negative real part.
- **Center**: A descriptor applied to the origin of the real plane when both eigenvaluess of the planar linear autonomous first-order system in question have *purely imaginary* value.
- Theorem 3.4: A solution of the linear system

  $$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

  converges to 0 as $t \to +\infty$ iff the initial condition x_0 lies in the subspace spanned by the generalized eigenspaces corresponding to eigenvalues with negative real part.

  It will remain bounded as $t \to +\infty$ iff x_0 lies in the subspace spanned by the generalized eigenspaces corresponding to eigenvalues with negative real part plus the eigenspaces corresponding to eigenvalues with vanishing real part.

  The behavior as $t \to -\infty$ mimics the above description, except that we replace "negative" with "positive."
- **Stable** (linear system): A linear system for which all solutions remain bounded as $t \to +\infty$.
- **Asymptotically stable** (linear system): A linear system for which all solutions converge to 0 as $t \to +\infty$.
- Corollary 3.5: The linear system

  $$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

  is stable iff all eigenvalues α_j of A satisfy $\operatorname{Re}(\alpha_j) \leq 0$, and for all eigenvalues with $\operatorname{Re}(\alpha_j) = 0$, the corresponding algebraic and geometric multiplicities are equal. Moreover, in this case, there is a constant C such that

  $$\|\exp(tA)\| \leq C$$

  for all $t \geq 0$.
- Corollary 3.6: The linear system

  $$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

  is asymptotically stable iff all eigenvalues α_j of A satisfy $\operatorname{Re}(\alpha_j) < 0$. Moreover, in this case, there is a constant $C = C(\alpha)$ for every $\alpha < \min\{-\operatorname{Re}(\alpha_j)\}_{j=1}^m$ (i.e., for every number of lesser value than the absolute value of the real component of the eigenvalue with smallest real part) such that

  $$\|\exp(tA)\| \leq C\mathrm{e}^{-t\alpha}$$

  for all $t \geq 0$.

Proof. It remains to prove the second claim.

Since $\left\|U\exp(tJ)U^{-1}\right\| \leq \|U\|\|\exp(tJ)\|\left\|U^{-1}\right\|$, we can assume that A is in JCF WLOG[6]. Now observe that

$$\begin{aligned}
\|\exp(tA)\| &= \|\exp(-t\alpha\mathbb{I} + tA + t\alpha\mathbb{I})\| \\
&= \|\exp(-t\alpha\mathbb{I})\exp(t(A+\alpha\mathbb{I}))\| \\
&= \left\|\mathrm{e}^{-t\alpha}\exp(\mathbb{I})\right\|\|\exp(t(A+\alpha\mathbb{I}))\| \\
&= \mathrm{e}^{-t\alpha}\|\exp(t(A+\alpha\mathbb{I}))\|
\end{aligned}$$

Since $\mathrm{Re}(\alpha_j + \alpha) < 0$ (sum of two negative components will be negative), all entries of the matrix $\exp(t(A+\alpha\mathbb{I}))$ are bounded as $t \to +\infty$ and consequently $\|\exp(t(A+\alpha\mathbb{I}))\|$ is bounded — say by C — so we may write

$$\begin{aligned}
\|\exp(t(A+\alpha\mathbb{I}))\| &\leq C \\
\mathrm{e}^{t\alpha}\|\exp(tA)\| &\leq C \\
\|\exp(tA)\| &\leq C\mathrm{e}^{-t\alpha}
\end{aligned}$$

as desired. $\square$

- **Hurwitz matrix**: A matrix, the eigenvalues of which all satisfy $\mathrm{Re}(\alpha_j) < 0$.
- **Routh-Hurwitz criterion**: What does this mean??
- Teschl (2012) reiterates **Duhamel's formula** for linear autonomous first-order systems.

Section 3.3: Linear Autonomous Equations of Order n

- Consider the n^{th} order equation
$$x^{(n)} + c_{n-1}x^{(n-1)} + \cdots + c_1\dot{x} + c_0 x = 0$$
with initial conditions
$$x(0) = x_0, \quad \ldots, \quad x^{(n-1)}(0) = x_{n-1}$$
- As described in depth in Lecture 4.1 and 1.1, we can transform this equation to a linear system with matrix
$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -c_0 & -c_1 & \cdots & \cdots & -c_{n-1} \end{pmatrix}$$
- Compute the characteristic polynomial by performing the Laplace expansion on the last row of
$$z\mathbb{I} - A = \begin{pmatrix} z & -1 & & & \\ & z & -1 & & \\ & & \ddots & \ddots & \\ & & & z & -1 \\ c_0 & c_1 & \cdots & \cdots & z + c_{n-1} \end{pmatrix}$$
to yield
$$\chi_A(z) = \det(z\mathbb{I} - A) = z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0$$

[6] Note that it also makes intuitive sense that the matrix norm should be the same for A, J since they are the same linear transformation in different bases and will thus distort the unit n-sphere symmetrically.

  - Note that it is probably for the best to just accept that the characteristic polynomial has this nice form. We can prove this claim more rigorously using induction on n.

- Claim: The geometric multiplicity of every eigenvalue is one.
- Theorem 3.7: Let α_j, $1 \leq j \leq m$, be the zeros of the characteristic polynomial

$$z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0 = \prod_{j=1}^{m} (z - \alpha_j)^{a_j}$$

associated with A, and let a_j be the corresponding multiplicities. Then the functions

$$x_{j,k}(t) = t^k \exp(a_j t)$$

for $0 \leq k < a_j$ and $1 \leq j \leq m$ are n linearly independent solutions of

$$x^{(n)} + c_{n-1}x^{(n-1)} + \cdots + c_1 \dot{x} + c_0 x = 0$$

In particular, any other solution can be written as a linear combination of these solutions.

Proof. Given. □

- Notes on Theorem 3.7.
  - This theorem could be incredibly helpful to me!
  - Note that if $\alpha_j = \lambda_j + i\omega_j$, we take

$$t^k \mathrm{e}^{\lambda_j t} \cos(\omega_j t) \qquad\qquad t^k \mathrm{e}^{\lambda_j t} \sin(\omega_j t)$$

- Example illustrating the power of Theorem 3.7.
  - Consider the ODE

$$\ddot{x} + \omega_0^2 x = 0, \quad \omega_0 > 0$$

  - The characteristic polynomial is $\alpha^2 + \omega_0^2 = 0$ and the zeroes are $\alpha_1 = i\omega_0$ and $\alpha_2 = -i\omega_0$.
  - Hence, for $\omega_0 > 0$, a basis of solutions is

$$x_1(t) = \mathrm{e}^{i\omega_0 t} \qquad\qquad x_2(t) = \mathrm{e}^{-i\omega_0 t}$$

  - If we want real solutions, we can get

$$x_1(t) = \cos(\omega_0 t) \qquad\qquad x_2(t) = \sin(\omega_0 t)$$

  - If $\omega_0 = 0$, we only have one zero $\alpha_1 = 0$ of multiplicity $a_1 = 2$, so a basis of solutions is given by

$$x_{1,0}(t) = 1 \qquad\qquad x_{1,1}(t) = t$$

- Consider the inhomogeneous equation

$$x^{(n)} + c_{n-1}x^{(n-1)} + \cdots + c_1 \dot{x} + c_0 x = g(t)$$

  - The general solution is given by Duhamel's formula as follows, where $x_h(t)$ is an arbitrary solution of the homogeneous equation (as provided by the above) and $u(t)$ is the solution of the homogeneous equation corresponding to the initial condition $u(0) = \dot{u}(0) = \cdots = u^{(n-2)}(0) = 0$ and $u^{(n-1)} = 1$ (see Problem 3.21).

$$x(t) = x_h(t) + \int_0^t u(t-s)g(s)\,\mathrm{d}s$$

- Algorithm for solving a linear n^{th} order equation with constant coefficients.
  1. Start with the homogeneous equation.
  2. Compute the zeroes of the characteristic polynomial and write down the general solution as a linear combination of the fundamental solutions.
  3. Find a particular solution of the inhomogeneous equation and determine the unknown constants of the homogeneous equation from the initial conditions.
- Lemma 3.8: Consider the inhomogeneous equation

$$x^{(n)} + c_{n-1}x^{(n-1)} + \cdots + c_1\dot{x} + c_0x = p(t)\mathrm{e}^{\beta t}$$

  where $p(t)$ is a polynomial. Then there is a particular solution of the same form $x_p(t) = q(t)\mathrm{e}^{\beta t}$, where $q(t)$ is a polynomial which satisfies $\deg(q) = \deg(p)$ if $\beta \notin \{\alpha_j\}_{j=1}^m$ is not equal to any of the characteristic eigenvalues and $\det(q) = \deg(p) + a_j$ if $\beta = \alpha_j$ is equal to one of the characteristic eigenvalues whose algebraic multiplicity is a_j.
  - In the case $\beta = \alpha_j$, we assume the first a_j coefficients to be zero since they correspond to a homogeneous solution.
  - If you allow complex values $\beta = \lambda + i\omega$, then we get real $g(t) = p(t)\mathrm{e}^{\lambda t}\cos(\omega t)$ or $g(t) = p(t)\mathrm{e}^{\lambda t}\sin(\omega t)$.
  - Linear combinations of such terms comes for free by linearity.
  - Of particular importance: The case of second order, which appears in a vast number of applications.
- **RLC circuit**: An inductor L, capacitor C, and resistor R arranged in a loop together with an external power source V.
- Teschl (2012) does the RLC circuit derivation associated with Figure 5.5b. Particularly interesting notes:
  - Since a circuit is a set of elements, each of which has two connectors (in and out) and every connector is connected to one or more connectors of the other elements, a circuit is mathematically a **directed graph**.
  - At each time t, there will be a certain current $I(t)$ flowing through each element and a certain voltage difference $V(t)$ between its connectors. The state space of the system is given by the pairs (I, V) of all elements in the circuit.
  - The pairs (I, V) must satisfy **Kirchoff's first law** and **Kirchoff's second law**. In this case, we have respectively that

$$I_R = I_L = I_C \qquad\qquad V_R + V_L + V_C = V$$

  - We can write a further three equations

$$L\dot{I}_L = V_L \qquad\qquad C\dot{V}_C = I_C \qquad\qquad V_R = RI_R$$

    where $L, C, R > 0$ are the inductance, capacitance, and resistance, respectively; $I_L(t), I_C(t), I_R(t)$ is the current through the inductor, capacitor, and resistor, respectively; and $V_L(t), V_C(t), V_R(t)$ is the voltage difference across the inductor, capacitor, and resistor, respectively.
  - Combining these five equations into one.
    - We first differentiate (wrt. t) and rearrange the latter three to put them in terms of $\dot{V}_i$ $(i = L, C, R)$.

$$\dot{V}_L = L\ddot{I}_L \qquad\qquad \dot{V}_C = \frac{I_C}{C} \qquad\qquad \dot{V}_R = R\dot{I}_R$$

- We then differentiate Kirchoff's second law and substitute in the above.

$$\begin{aligned}\dot{V}(t) &= \dot{V}_R + \dot{V}_L + \dot{V}_C \\ &= R\dot{I}_R + L\ddot{I}_L + \frac{I_C}{C}\end{aligned}$$

- Lastly, we take advantage of Kirchoff's first law and drop the subscript from all of the currents.

$$L\ddot{I}(t) + R\dot{I}(t) + \frac{1}{C}I(t) = \dot{V}(t)$$

– To get our final equation, use the complex voltage $V(t) = V_0 e^{i\omega t}$ and divide through by L.

$$\ddot{I} + \frac{R}{L}\dot{I} + \frac{1}{LC}I = \frac{i\omega V_0}{L}e^{i\omega t}$$

- **Kirchoff's first law**: Charge is conserved, i.e., the sum over all currents in a circuit element (graph vertex) must vanish.
- **Kirchoff's second law**: The voltage corresponds to a potential, i.e., the sum over all voltage differences in a closed loop must vanish.
- Analysis of the RLC circuit equation. *picture*
  – The eigenvalues are

$$\alpha_{1,2} = -\eta \pm \sqrt{\eta^2 - \omega_0^2}$$

  where we have defined

$$\eta = \frac{R}{2L} \qquad\qquad \omega_0 = \frac{1}{\sqrt{LC}}$$

  – Three cases to consider based on the discriminant: $\eta > \omega_0$ (**over damping**), $\eta = \omega_0$ (**critical damping**), and $\eta < \omega_0$ (**under damping**).
  – It follows from Theorem 3.7 that the respective general homogeneous solutions for the current are

$$I_h(t) = k_1 e^{\alpha_1 t} + k_2 e^{\alpha_2 t} \qquad I_h(t) = (k_1 + k_2 t)e^{-\eta t} \qquad I_h(t) = k_1 e^{-\eta t}\cos(\beta t) + k_2 e^{-\eta t}\sin(\beta t)$$

  for $k_1, k_2 \in \mathbb{C}$ and where $\beta = \sqrt{\omega_0^2 - \eta^2} > 0$ in the last equation (under damping).
  – In every case, the real part of both eigenvalues is negative, so the homogeneous solution decays exponentially as $t \to +\infty$.
  – Observation: For fixed $\eta > 0$, the choice $\omega_0 = \eta$ gives the fastest decay without an oscillatory component.
  – To get a particular solution for the inhomogeneous equation, use $I_i(t) = ke^{i\omega t}$ as the ansatz:

$$\begin{aligned}-L\omega^2 ke^{i\omega t} + Ri\omega ke^{i\omega t} + \frac{1}{C}ke^{i\omega t} &= \frac{i\omega V_0}{L}e^{i\omega t} \\ -L\omega^2 k + Ri\omega k + \frac{1}{C}k &= \frac{i\omega V_0}{L} \\ k &= \frac{V_0}{R + i\left(L\omega - \frac{1}{\omega C}\right)}\end{aligned}$$

  – Since the homogeneous solution decays exponentially, we have after a short time

$$I(t) \approx I_i(t) = \frac{V_0}{Z}e^{i\omega t} = \frac{1}{Z}V(t)$$

  where

$$Z = R + Z_L + Z_C, \quad Z_L = iL\omega, \quad Z_C = -\frac{i}{\omega C}$$

  is the **complex impedence**.

  - It follows from the above that the current obtains its maximum when $|Z|^2$ (an easily computed measure of complex radius) is minimal. If we want to minimize

$$|Z|^2 = R^2 + \left(L\omega - \frac{1}{\omega C}\right)^2$$

then we should minimize R and let

$$L\omega - \frac{1}{\omega C} = 0$$
$$\omega = \frac{1}{\sqrt{LC}} = \omega_0$$

The frequency $\omega_0/2\pi$ is called the **resonance frequency** of the circuit.

  - By changing one of the parameters — say C — you can thus tune the circuit to a specific resonance frequency.
    - "This idea is for example used ot filter your favorite radio station out of many other available ones. In this case, the external power source corresponds to the signal picked up by your antenna and the RLC circuit starts only oscillating if the carrying frequency of your radio station matches its resonance frequency" (Teschl, 2012, p. 79).

- Many other systems beyond RLC circuits can be described (at least for small amplitudes) by the differential equation

$$\ddot{x} + 2\eta\dot{x} + \omega_0^2 x = 0, \quad \eta, \omega_0 > 0$$

  - η is called the **damping factor**.
  - Making the equation inhomogeneous with a periodic **forcing** term $\cos(\omega t)$ produces maximal effect if the forcing is resonant, that is, if ω coincides with ω_0.
  - If $\eta = 0$, the solution corresponds to a free (undamped) oscillation $x(t) = k_1\cos(\omega_0 t) + k_2\sin(\omega_0 t)$ and a resonant forcing will result in a solution whose amplitude tends to ∞.

Week 6

Qualitative Theory of ODEs

6.1 More Cauchy-Lipschitz and Intro to Continuous Dependence

10/31:

- Last time, we built up a proof to the Cauchy-Lipschitz theorem intuitively.
  - We begin today with a direct proof that is very similar, but slightly different.
- Theorem (Cauchy-Lipschitz theorem): Let $f(t,z)$ be defined on an open subset $\Omega \subset \mathbb{R} \times \mathbb{R}^n$, let $(t_0, y_0) \in \Omega$, let $|f|$ be bounded on Ω, and let f be Lipschitz continuous in z and continuous wrt. t in some neighborhood of (t_0, y_0). Then the IVP $y'(t) = f(t, y(t))$, $y(t_0) = y_0$ has a unique solution on $[t_0, t_0 + T]$ for some $T > 0$ such that $y(t)$ does not escape Ω.

  Proof. Let $f(t,z)$ be defined for $(t,z) \in [t_0, t_0 + a] \times \bar{B}(y_0, b) \subset \Omega$. Let $|f(t,z)| \leq M$. Let $|f(t,z_1) - f(t,z_2)| \leq L|z_1 - z_2|$ for all $z_1, z_2 \in \bar{B}(y_0, b)$.

  Define $\{y_n\}$ recursively, starting from $y_0(t) = y_0$, by

  $$y_{k+1}(t) = y_0 + \int_{t_0}^{t} f(\tau, y_k(\tau))\mathrm{d}\tau$$

  Since f is continuous with respect to t, it is integrable with respect to t, so the above sequence is well-defined on $[t_0, t_0 + T]$. Choose $T = \min(a, b/M, 1/2L)$. Then

  $$\|y_k - y_0\| \leq T \cdot M \leq \frac{b}{M} \cdot M = b$$

  so no y_k escapes $\bar{B}(y_0, b)$. Additionally,

  $$\begin{aligned}
  \|y_{k+1} - y_k\| &\leq \int_{t_0}^{t} \|f(\tau, y_k(\tau)) - f(\tau, y_{k-1}(\tau))\|\mathrm{d}\tau \\
  &\leq TL\|y_k - y_{k-1}\| \\
  &\leq \frac{1}{2}\|y_k - y_{k-1}\| \\
  &\leq \left(\frac{1}{2}\right)^k \|y_1 - y_0\|
  \end{aligned}$$

  Thus, the difference between successive terms in the sequence is controlled by a geometric progression, so $\{y_n\}$ is a Cauchy sequence in the function space. It follows that $\{y_k\}$ is uniformly convergent to some continuous $y : [t_0, t_0 + T] \to \mathbb{R}^n$. $\square$

- This completes the proof. Although it's more concrete than the contraction mapping one, they are virtually the same: In both cases, we obtain an approximate sequence controlled by a geometric progression.

- Examples of the Picard iteration:
  1. Consider an linear autonomous systems $y' = Ay$, A an $n \times n$ matrix, and $y(0) = y_0$.
     - We know that the solution is $y(t) = \mathrm{e}^{tA} y_0$. However, we can derive this using the Picard iteration.
     - Indeed, via this procedure, let's determine the first couple of Picard iterates.

$$y_0(t) = y_0 \qquad \begin{aligned} y_1(t) &= y_0 + \int_0^t Ay_0(\tau)\mathrm{d}\tau \\ &= y_0 + tAy_0 \end{aligned} \qquad \begin{aligned} y_2(t) &= y_0 + \int_0^t Ay_1(\tau)\mathrm{d}\tau \\ &= y_0 + tAy_0 + \frac{1}{2}t^2A^2y_0 \end{aligned}$$

     - It follows inductively that

$$y_k(t) = \sum_{j=0}^{k} \frac{t^j A^j}{j!} y_0$$

     - Since the term above is exactly the power series definition of e^{tA}, we have that $y_k(t) \to \mathrm{e}^{tA} y_0$ with local uniformity in t, as desired.
  2. Consider the ODE $y' = y^2$, $y(0) = 1$.
     - We know that the solution is $y(t) = 1/(1-t)$. We will now also derive this via the Picard iteration.
     - Choose $b = 1$, so that

$$\bar{B}(y_0, b) = \{y \mid |y - y(0)| \leq 1\} = \{y \mid |y - 1| \leq 1\} = [0, 2]$$

     - On this interval, $f(t, y) = y^2$ has maximum slope $L = 4$. Thus, we should take $T \leq 1/2L = 1/8$.
     - It follows that $|y_1^2 - y_2^2| \leq 4|y_1 - y_2|$ for all $y_1, y_2 \in \bar{B}(y_0, b)$.
     - Calculate the first few Picard iterates.

$$\begin{aligned} y_1(t) &= 1 + \int_0^t (y_0(\tau))^2 \mathrm{d}\tau = 1 + t \\ y_2(t) &= 1 + \int_0^t (1+\tau)^2 \mathrm{d}\tau = 1 + t + t^2 + \frac{t^3}{3} \\ y_3(t) &= 1 + \int_0^t \left(1 + \tau + \tau^2 + \frac{\tau^3}{3}\right)^2 \mathrm{d}\tau = 1 + t + t^2 + t^3 + \frac{2t^4}{3} + \frac{t^5}{3} + \frac{t^6}{9} + \frac{t^7}{63} \end{aligned}$$

     - It follows by induction that

$$\begin{aligned} |y_k(t) - (1 + t + \cdots + t^k)| &\leq t^{k+1} \\ \left| y_k(t) - \frac{1 - t^{k+1}}{1 - t} \right| &\leq t^{k+1} \end{aligned}$$

       It follows that $|t| < 1/8$.
     - For $|t| < 1/8$, $y(t) = 1/(1-t)$. Blows up as $t \to 1$.
     - Some more details on the bounding of the error term are presented in the lecture notes document.
- Lemma (Grönwall's inequality): Let $\varphi(t)$ be a real function defined for $t \in [t_0, t_0 + T]$ such that

$$\varphi(t) \leq f(t) + a \int_{t_0}^t \varphi(\tau)\mathrm{d}\tau$$

  Then

$$\varphi(t) \leq f(t) + a \int_{t_0}^t \mathrm{e}^{a(t-\tau)} f(\tau)\mathrm{d}\tau$$

Proof. Multiply both sides by e^{-at}:

$$\begin{aligned}
\mathrm{e}^{-at}\varphi(t) - a\mathrm{e}^{-at}\int_{t_0}^{t}\varphi(\tau)\mathrm{d}\tau &\leq \mathrm{e}^{-at}f(t)\\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-at}\int_{t_0}^{t}\varphi(\tau)\mathrm{d}\tau\right) &\leq \mathrm{e}^{-at}f(t)\\
\mathrm{e}^{-at}\int_{t_0}^{t}\varphi(\tau)\mathrm{d}\tau &\leq \int_{t_0}^{t}\mathrm{e}^{-a\tau}f(\tau)\mathrm{d}\tau\\
\int_{t_0}^{t}\varphi(\tau)\mathrm{d}\tau &\leq \int_{t_0}^{t}\mathrm{e}^{a(t-\tau)}f(\tau)\mathrm{d}\tau
\end{aligned}$$

Substituting back into the original equality yields the result at this point. □

- Note that there is no sign condition on $f(t)$ or a.
- Grönwall's inequality is very important and we should remember it.
- It is also exactly what we need to prove continuous dependence.
- Theorem: Let $f(t,z), g(t,z)$ be defined on $\Omega \subset \mathbb{R}^1_t \times \mathbb{R}^n_z$, an open and bounded a region containing (t_0, y_0) and (t_0, w_0). Let the functions be L - Lipschitz wrt. z. Consider two initial value problems $y' = f(t,y)$, $y(t_0) = y_0$ and $w' = g(t,w)$, $w(t_0) = w_0$. If $|f(t,z) - g(t,z)| < M$, then for $t \in [t_0, t_0 + T]$,

$$|y(t) - w(t)| \leq \mathrm{e}^{LT}|y_0 - w_0| + \frac{M}{L}(\mathrm{e}^{LT} - 1)$$

Proof. We have that

$$\begin{aligned}
|y(t) - w(t)| &= \left|\left[y_0 + \int_{t_0}^{t} f(\tau, y(\tau))\mathrm{d}\tau\right] - \left[w_0 + \int_{t_0}^{t} g(\tau, y(\tau))\mathrm{d}\tau\right]\right|\\
&= \left|[y_0 - w_0] + \int_{t_0}^{t}[f(\tau, y(\tau)) - g(\tau, y(\tau))]\mathrm{d}\tau\right|\\
&\leq |y_0 - w_0| + \left|\int_{t_0}^{t}[f(\tau, y(\tau)) - g(\tau, w(\tau))]\mathrm{d}\tau\right|\\
&\leq |y_0 - w_0| + \int_{t_0}^{t}|f(\tau, y(\tau)) - g(\tau, w(\tau))|\mathrm{d}\tau
\end{aligned}$$

where we get from the second to the third line using the triangle inequality, and the third to the fourth line using Theorem 13.26 of Honors Calculus IBL. We also know that

$$\begin{aligned}
|f(\tau, y(\tau)) - g(\tau, w(\tau))| &\leq |f(\tau, y(\tau)) - f(\tau, w(\tau))| + |f(\tau, w(\tau)) - g(\tau, w(\tau))|\\
&\leq L|y(\tau) - w(\tau)| + M
\end{aligned}$$

Combining what we've obtained, we have

$$\begin{aligned}
\underbrace{|y(t) - w(t)|}_{\psi(t)} &\leq \underbrace{|y_0 - w_0| + M(t - t_0)}_{f(t)} + \underbrace{L}_{a}\int_{t_0}^{t}\underbrace{|y(\tau) - w(\tau)|}_{\psi(t)}\mathrm{d}\tau\\
&\leq MT + |y_0 - w_0| + L\int_{t_0}^{t}\mathrm{e}^{L(t-\tau)}[|y_0 - w_0| + M(t - \tau)]\mathrm{d}\tau && \text{Grönwall}\\
&\leq \mathrm{e}^{LT}|y_0 - w_0| + \frac{M}{L}(\mathrm{e}^{TL} - 1)
\end{aligned}$$

as desired. □

- Note: Getting from directly from Grönwall's inequality in the second line above to the last line above is quite messy. A consequence of Grönwall's inequality explored in the book makes this much easier. *Prove Equation 2.38 via Problem 2.12.*
- Implication: The IVP is not just solvable itself, but is solvable wrt. perturbation of the initial conditions and RHS within a small, finite interval in time.
- Suppose $y' = 0$, $y(0) = 1$ and $w' = \varepsilon w$, $w(0) = 1$. Then $y(t) = 1$ and $w(t) = \mathrm{e}^{\varepsilon t}$ and solutions are only close when t is small.
  - $t \leq 1/\varepsilon$??
- This is important in physics. In most physical scenarios, the RHS is C^1. This is called determinism.

6.2 Differentiability With Respect To Parameters

11/2:

- Review: Implicit Function Theorem.
  - Gives you a sufficient condition for which an implicit relation defines a function.
  - Does not give you the function, but tells you that it must exist and that it is unique.
- Theorem (Implicit Function Theorem): Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be C^k in some neighborhood of $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$ a point satisfying $F(x_0, y_0) = 0$. If the truncated Jacobian matrix $\frac{\partial F}{\partial y}(x_0, y_0)$, which is $m \times m$, is invertible, then there is a neighborhood U of x_0 such that there is a unique function $f : U \to \mathbb{R}^m$ with $y_0 = f(x_0)$ and $F(x, f(x)) = 0$ and

$$f'(x) = -\left(\frac{\partial F}{\partial y}(x, y)\right)^{-1} \cdot \frac{\partial F}{\partial x}(x, f(x))$$

  - The proof is based on the Banach fixed point theorem (this may be false?? I think Shao is confusing the proof of this theorem with the proof of the Inverse Function Theorem).
  - The motivation for the last equality (the line above) is that if $F(x, f(x)) = 0$, then by the chain rule for partial derivatives,

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}x}(F(x, f(x))) \\
&= \frac{\partial F}{\partial x}(x, f(x)) \cdot \frac{\mathrm{d}x}{\mathrm{d}x} + \left[\frac{\partial F}{\partial y}(x, y)\right] \cdot \frac{\mathrm{d}f}{\mathrm{d}x} \\
&= \frac{\partial F}{\partial x}(x, f(x)) + \left[\frac{\partial F}{\partial y}(x, y)\right] \cdot f'(x) \\
f'(x) &= -\left(\frac{\partial F}{\partial y}(x, y)\right)^{-1} \cdot \frac{\partial F}{\partial x}(x, f(x))
\end{aligned}$$

  - Recall that we know that the matrix bracketed in line 2 is invertible by hypothesis.
  - Additionally, since $\partial F/\partial x = A$ is $n \times m$ and $\partial F/\partial y = B$ is $m \times m$, $f' = -A^{-1}B$ is $n \times m$, as it should be for a function $f : \mathbb{R}^n \to \mathbb{R}^m$.
- Consider the IVP

$$y' = f(t, y; \mu), \quad y(t_0) = x(\mu)$$

  - This ODE and its initial condition both depend on a parameter $\mu \in B(0, r) \subset \mathbb{R}^m$ (usually we take $m = 1$ so μ is just real).
  - We denote the solution by $y(t; \mu)$.

- Suppose $|x(\mu)| < C$ for $\mu \in B(0,r)$ and $x(\mu) \in C^1$. Suppose the RHS $f(t,z;\mu)$ of the ODE is defined on $[t_0, t_0+a] \times \bar{B}(x(0), b+C) \times B(0,r)$, is C^1 in all variables, is bounded by M on its domain, and is L-Lipschitz in z.

- By Cauchy-Lipschitz, for small
$$T \leq \min\left(a, \frac{b}{M}, \frac{1}{2L}\right)$$
and $\mu \in B(0,r)$ (r small), the solution *exists* on $[t_0, t_0+T]$ and its value does not escape $\bar{B}(x(0), b+C)$.
  - We now aim to show that the solution is *differentiable* wrt. μ on this interval.

- If $y(t;\mu)$ satisfies $y'(t;\mu) = f(t, y(t;\mu);\mu)$ and if the Jacobian matrix $J = \partial y/\partial\mu$ exists, then J satisfies the **first variation equation**.

- **First variation equation**: The following linear differential equation. *Given by*
$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\frac{\partial y}{\partial \mu}(t;\mu)}_{J(t;\mu)} = \underbrace{\frac{\partial f}{\partial z}(t, y(t;\mu);\mu)}_{A(t;\mu)} \cdot \underbrace{\frac{\partial y}{\partial \mu}(t;\mu)}_{J(t;\mu)} + \frac{\partial f}{\partial \mu}\,(t, y(t;\mu);\mu), \quad \frac{\partial y}{\partial \mu}\,(t_0,\mu) = \frac{\partial x}{\partial \mu}\,(\mu)$$

- The first variation equation has a unique solution, but we do not yet know that $y(t;\mu)$ is even differentiable with respect to μ. We presently verify this claim.

- Theorem[1]: $y(t;\mu)$ is C^1 in μ and $\partial y/\partial\mu\,(t;\mu)$ satisfies the first variation equation.

Proof. Let $\Theta(t;\mu) = y(t;\mu+h) - y(t;\mu) - J(t;\mu)h$ for h small. Aim, show that $\Theta(t;\mu) = o(h)$ as $h \to 0$.

We compute
$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Theta(t;\mu) &= y'(t;\mu+h) - y'(t;\mu) - J'(t;\mu)h \\ &= \underbrace{f(t, y(t;\mu+h);\mu+h) - f(t,y(t;\mu);\mu)}_{I} - \underbrace{\frac{\partial f}{\partial z}(t,y(t;\mu);\mu)J(t;\mu) + \frac{\partial f}{\partial \mu}(t,y(t;\mu);\mu)}_{II}\end{aligned}$$

I denotes the first term; II denotes the second term.

We have that
$$I = \frac{\partial f}{\partial z}(t,y(t;\mu);\mu)[y(t;\mu+h) - y(t;\mu)] + \frac{\partial f}{\partial \mu}(t,y(t;\mu);\mu)h + \underbrace{R(t;\mu,h)}_{o(h)}$$

color coding

$$\begin{aligned} I - II &= \underbrace{\text{green} - \text{blue}}_{\Theta(t;\mu)} + R(t;\mu,h) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Theta(t;\mu) = \Theta(t;\mu) + \underbrace{R(t;\mu,h)}_{o(h)} \\ \Theta(t_0;\mu) &= o(h) \\ |\Theta(t;\mu)| &\leq C\int_{t_0}^{t} |R(\tau;\mu,h)|\mathrm{d}\tau && \text{Grönwall} \\ &= o(h)\end{aligned}$$

circle terms cancel. □

[1]See the proof from the book, transcribed below.

- Example: First order derivatives must satisfy the first variational equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial y}{\partial \mu}(t;\mu) = \frac{\partial f}{\partial z}(t, y(t;\mu);\mu) \cdot \frac{\partial y}{\partial \mu}(t;\mu)$$

  and the second order derivative must satisfy the second variational equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial^2 y}{\partial \mu^2} = \frac{\partial^2 f}{\partial z^2}\left(\frac{\partial y}{\partial \mu}\frac{\partial^2 y}{\partial \mu^2}\right) + \frac{\partial^2 f}{\partial z \partial \mu}\frac{\partial y}{\partial \mu} + \frac{\partial^2 f}{\partial \mu \partial z}(-)\frac{\partial y}{\partial \mu} + \frac{\partial^2 f}{\partial \mu^2}(-)$$

- Corollary: If $f(t,z;\mu)$ is C^k in (t,z,μ), $y(t_0) = x(\mu)$ is C^k, then $y(t;\mu)$ is C^k in μ.
- The Taylor expansion

$$y(t;\mu) = y(t;0) + y_1\mu + y_2\mu^2 + \cdots + y_k\mu^k + O(\mu^{k+1})$$

  of $y(t;\mu)$ about 0 gives an approximation of said function up to order k in μ.
  - Misc notes: but you can cut off the expansion at k?? $y(t;0)$ being solvable implies inductively that the rest are solvable??
  - We can take this Taylor expansion because we assume that y is continuously differentiable k times with respect to μ.
  - The coefficients y_j are given as follows.

$$y_j = \frac{1}{j!}\frac{\partial^j y}{\partial \mu^j}\ (t;0)$$

- Application of the Taylor expansion: It can be substituted into the ODE as follows.

$$\begin{aligned}
\dot{y} &= f(t,y;\mu) \\
\frac{\mathrm{d}}{\mathrm{d}t}(y(t;\mu)) &= f(t, y(t;\mu);\mu) \\
\frac{\mathrm{d}}{\mathrm{d}t}(y(t;0)) + \frac{\mathrm{d}y_1}{\mathrm{d}t}\mu + \cdots + \frac{\mathrm{d}y_k}{\mathrm{d}t}\mu^k + O(\mu^{k+1}) &= f(t, y(t;0) + y_1\mu + \cdots + y_k\mu^k + O(\mu^{k+1});\mu)
\end{aligned}$$

  - Then you can match coefficients of the various μ terms on the LHS and RHS and solve for $y_0, \dots, y_k$.
  - When to use this method: Sometimes, you can view equations that aren't explicitly solvable as perturbations of an easily solvable system.
- Simple example (more complex ones next lecture):

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \mu y, \quad y(0) = 1$$

  - First off, we know that there is an explicit solution ($y(t) = \mathrm{e}^{\mu t}$). Thus, we will be able to check our final answer.
  - Suppose $y \in C^2$ with respect to μ. Then

$$y(t;\mu) = y_0 + y_1\mu + y_2\mu^2 + O(\mu^3)$$

  - It follows by substituting into the above differential equation that

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}t} &= \mu y \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(y_0 + y_1\mu + y_2\mu^2\right) &= \mu(y_0 + y_1\mu + y_2\mu^2) \\
\frac{\mathrm{d}y_0}{\mathrm{d}t} + \frac{\mathrm{d}y_1}{\mathrm{d}t}\mu + \frac{\mathrm{d}y_2}{\mathrm{d}t}\mu^2 &= 0 + y_0\mu + y_1\mu^2 + y_2\mu^3
\end{aligned}$$

- By comparing coefficients, this yields the sequentially solvable differential equations

$$\frac{\mathrm{d}y_0}{\mathrm{d}t} = 0 \qquad \frac{\mathrm{d}y_1}{\mathrm{d}t} = y_0 \qquad \frac{\mathrm{d}y_2}{\mathrm{d}t} = y_1$$

where we apply the initial condition $y_0(0) = 1$ to solve the left ODE above.

- Solving, we get

$$y_0(t) = 1 \qquad y_1(t) = t \qquad y_2(t) = \frac{t^2}{2}$$

  - Where do the other initial conditions (all zero) come from??

- Therefore, our approximate solution is

$$y(t) = 1 + t\mu + \frac{1}{2}t^2\mu^2 + O(\mu^3)$$

which does indeed give the first three terms in the Taylor series expansion of the solution $\mathrm{e}^{\mu t}$.

- The perturbative solution fails in large time intervals — polynomials inevitably grow slower than exponential functions.
- Next time: Several examples applying what we've learned today.
- This week's homework: Some basic Lipschitz definitions and also computations with the perturbative series.

6.3 Variational Examples

11/4:

- We begin today with a more direct and less involved proof of the variation of parameters theorem.

Proof. Let $y'(t;\mu) = f(t, y(t;\mu);\mu)$ with $y(t_0;\mu) = x(\mu)$. Assume Lipschitz continuity and C^1-ness of the ODE and the initial condition on μ. Then differentiation with respect to μ must satisfy the first variational equation. In particular, let $J(t;\mu)$ be the solution of

$$J'(t;\mu) = \underbrace{\frac{\partial f}{\partial z}(t, y(t;\mu);\mu)}_{A(t;\mu)} J(t,\mu) + \underbrace{\frac{\partial f}{\partial \mu}(t, y(t;\mu);\mu)}_{F(t;\mu)}, \quad J(t_0;\mu) = \frac{\partial x}{\partial \mu}$$

Consider the Picard iteration sequence defined by

$$y_{n+1}(t;\mu) = \underbrace{f(t; y_n(t;\mu);\mu)}_{A_n(t;\mu)}, \quad y_n(t_0,\mu) = x(\mu)$$

Differentiating we get

$$\frac{\partial y_n}{\partial \mu}(t;\mu)$$

which we may call $J_n(t;\mu)$. We want to prove that the sequence of functions J_n converges uniformly to J. This makes sense since A and F uniformly converge. Moreover, under this definition of J_n, we have that

$$J_{n+1}'(t;\mu) = \underbrace{\frac{\partial f}{\partial z}(t, y_n(t;\mu);\mu)}_{A_n(t;\mu)} J_n(t;\mu) + \underbrace{\frac{\partial f}{\partial \mu}(t, y_n(t;\mu);\mu)}_{F_n(t;\mu)}, \quad J_n(t_0;\mu) = \frac{\partial x}{\partial \mu}(\mu)$$

Thus, Step 1 is to show that $\{\|J_n\|\}$ is bounded on $[t_0, t_0 + T]$. To do so, we note that

$$\|J_{n+1}\| \leq \frac{1}{2}\|J_n\| + \sup\left|\frac{\partial f}{\partial \mu}\right|$$

so that $\|J_n\|$ forms a bounded sequence. By induction,

$$\|J_n\| \leq 2C$$

We now embark on Step 2: Proving $J_n \to J$ uniformly. First off, we have that

$$\begin{aligned}
(J - J_{n+1})'(t;\mu) &= \frac{\mathrm{d}}{\mathrm{d}t}(J(t;\mu) - J_{n+1}(t;\mu)) \\
&= A(t;\mu)J(t;\mu) + F(t;\mu) - A_n(t;\mu)J_n(t;\mu) - F_n(t;\mu) \\
&= A(t;\mu)J(t;\mu) + A_n(t;\mu)J(t;\mu) - A_n(t;\mu)J(t;\mu) \\
&\quad - A_n(t;\mu)J_n(t;\mu) + F(t;\mu) - F_n(t;\mu) \\
&= A_n(t;\mu)(J - J_n)(t;\mu) + (A - A_n)(t;\mu)J(t;\mu) + (F - F_n)(t;\mu)
\end{aligned}$$

and

$$J(t_0;\mu) - J_{n+1}(t_0;\mu) = 0$$

Integrating once again on $[t_0, t_0 + T]$, we get

$$\|J - J_{n+1}\| \leq \frac{1}{2}\|J - J_n\| + \delta_n$$

where $\delta_n \to 0$ since we "obviously" have that $A_n \to A$ and $F_n \to F$ uniformly.

We now proceed via a standard analysis argument. Fix $\delta > 0$, choose N such that $\delta_n < \delta$ for $n \geq N$. Then we can control it by $\frac{1}{2}\|J - J_n\| + \delta$ for $n \geq N$. Then

$$\|J - J_{n+1}\| - 2\delta \leq \frac{1}{2}\|J - J_n\| - 2\delta$$

for all $n \geq N$, so we have by iteration that $\|J - J_{n+1}\| \leq 2\delta + \frac{1}{2^{n-N}}\|J - J_N\|$, so $\lim_{n\to\infty}\|J - J_n\| < 2\delta$ for arbitrary $\delta > 0$. Therefore, $\|J - J_n\| \to 0$, so $J_n \to J$ uniformly.

So in conclusion, $J_n \to J$ uniformly and we recall that $J_n = \partial y_n/\partial\mu$ where $y_n \to y$ uniformly. □

- We now look at examples. The ones in the HW will be no more difficult than these.
- Example (same one as last time):
  - Consider $y' = \mu y$ with $y(0) = 1$.
  - In order to find asymptotic expansion wrt. μ, we use the **ansatz** $y(t;\mu) = y_0 + y_1\mu + y_2\mu^2 + \cdots + y_n\mu^n + O(\mu^{n+1})$.
    - The differentiation theorem asserts that $y(t;\mu)$ can be differentiated wrt. μ so many times.
  - We can compute
    $$\mu y(t;\mu) = 0 + y_0\mu + y_1\mu^2 + \cdots + y_{n-1}\mu^n + O(\mu^{n+1})$$
    and
    $$y'(t;\mu) = y_0' + y_1'\mu + y_2'\mu^2 + \cdots + y_n'\mu^n + O(\mu^{n+1})$$
    and set them equal to yield a system of differential equations.
  - The initial conditions are $y_0(0) = 1$ and then $y_1(0) = \cdots = y_n(0) = 0$.
  - $y_0' = 0$ with $y_0(0) = 1$ implies that $y_0(t) = 1$.
  - Then the first order approximation is $y_1' = y_0 = 1$, so solving and applying the initial conditions, we get $y_1(t) = t$.
  - Continuing on, the second order approximation is $y_2(t) = t^2/2$.
  - Inductively, $y_m(t) = t^m/m!$.
  - In conclusion, we obtain the desired approximate solution.

- **Ansatz**: The form of the solution that you guess.
- In general, this shows the technique well: Use a polynomial ansatz and compare terms to yield an inductive sequence of explicitly solvable equations up to a certain point.
- Example: Mathematical pendulum.
  - Suppose that the length of the rope is ℓ and the gravitational acceleration is g. Then
$$\theta''(t;\mu) = -\frac{g}{\ell}\sin[\theta(t;\mu)]$$
  - Assume a small angle, $\theta(0) = \mu$ and $\theta'(0) = 0$.
  - Substitute $\omega_0^2 = g/\ell$.
  - In HS, we learned that the harmonic oscillator approximation of the mathematical pendulum is justified for small θ. We now justify this.
  - Ansatz: $\theta_0 + \theta_1\mu + \theta_2\mu^2 + \theta_3\mu^3 + O(\mu^4)$.
  - Recall that
$$\sin\theta = \theta - \frac{\theta^3}{6} + O(\theta^5)$$
  - First step, solve to determine $\theta_0 = 0$.
  - Then we only have a term of order $O(\mu)$ and $O(\mu^3)$ to worry about.
  - Substitute the expansion in:
$$\begin{aligned}\sin\theta &= \theta - \frac{\theta^3}{6} + O(\theta^5)\\ &= \left(\theta_0 + \theta_1\mu + \theta_2\mu^2 + \theta_3\mu^3\right) - \frac{1}{6}\left(\theta_0 + \theta_1\mu + \theta_2\mu^2 + \theta_3\mu^3\right)^3\\ &= 0 + \theta_1\mu + \theta_2\mu^2 + \left(\theta_3 - \frac{\theta_1^3}{6}\right)\mu^3 + O(\mu^4)\end{aligned}$$
  - We also have that
$$\theta''(t;\mu) = \theta_1''\mu + \theta_2''\mu^2 + \theta_3''\mu^3 + O(\theta^4)$$
and
$$-\omega_0^2\sin\big(\theta_1\mu + \theta_2\mu^2 + \theta_3\mu^3 + O(\mu^4)\big) = -\omega_0^2\theta_1\mu - \omega_0^2\theta_2\mu^2 - \omega_0^2\left(\theta_3 - \frac{\theta_1^3}{6}\right)\mu^3 + O(\mu^4)$$
  - Initial conditions: $\theta_0 = 0$, $\theta_1(0) = 1$, and $\theta_2(0) = \theta_3(0) = \theta_1'(0) = \cdots = \theta_3'(0) = 0$.
  - First order: $\theta_1'' = -\omega_0^2\theta_1$, $\theta_1(0) = 1$, $\theta_1'(0) = 0$. Implies $\theta_1(t) = \cos\omega_0 t$. This is why we can use the harmonic oscillator approximation.
  - Second order: $\theta_2 = -\omega_0^2\theta_2$. Initial conditions imply $\theta_2(t) = 0$.
  - Third order: $\theta_3'' = -\omega_0^2\theta_3 + \frac{\omega_0^2\theta_1^3}{6}$. Implies that
$$\theta_3(t) = \frac{\omega_0 t}{16}\sin\omega_0 t + \frac{1}{192}(\cos\omega_0 t - \cos 3\omega_0 t)$$
    - We have to apply some trigonometric identities to verify this??
  - In conclusion, we have the approximation of our solution up to order $O(\mu^3)$ as
$$\theta(t;\mu) = \mu\cos\omega_0 t + \mu^3\left[\frac{\omega_0 t}{16}\sin\omega_0 t + \frac{1}{192}(\cos\omega_0 t - \cos 3\omega_0 t)\right] + O(\mu^4)$$
    - This approximation is only good for T in a fixed, small time interval because the second term is not periodic.

- We now investigate the period of the mathematical pendulum.
  - The first order approximation (harmonic oscillator) gives the period as $T \approx 2\pi/\omega_0 = 2\pi\sqrt{\ell/g}$.
  - Let $T(\mu)$ denote the period of the mathematical pendulum as a function of the starting angle μ.
  - $T(\mu)$ should be approximately equal to the period of $\theta(t;\mu)$. Additionally, thinking about the mathematical pendulum intuitively, the period $T(\mu)$ should be about four times the first positive zero of $\theta(t;\mu)$.
    - Indeed, in a full cycle, the pendulum must go from the positive extreme, to zero, to the negative extreme, back to zero, and back to the original position, so there are our four parts.
    - Example: In the harmonic oscillator approximation, the first zero is at $\pi/2\omega_0$, and the period is $2\pi/\omega_0 = 4 \cdot \pi/2\omega_0$.
  - Thus, determining the period $T(\mu)$ becomes a problem of finding t such that $\theta(t;\mu) = 0$.
  - The zeroes of $\theta(t;\mu)$ will be equal to the zeroes of $\theta(t;\mu)/\mu$, so we seek t such that the implicit function

$$F(t;\mu) = \frac{\theta(t;\mu)}{\mu} = \cos\omega_0 t + \mu^2\left[\frac{\omega_0 t}{16}\sin\omega_0 t + \frac{1}{192}(\cos\omega_0 t - \cos 3\omega_0 t)\right] = 0$$

  - When $\mu = 0$, the mathematical pendulum is stationary, but this does technically mean that it has a zero at $(\pi/2\omega_0; 0)$. This point is important because for μ small enough that the harmonic oscillator approximation is good, the first zero should be very close to $\pi/2\omega_0$. Thus, we choose to solve $F(t;\mu) = 0$ around $(t_0;\mu_0) = (\pi/2\omega_0; 0)$.
  - The requirement for the Implicit Function Theorem is met since

$$\begin{aligned}
\frac{\partial F}{\partial t}(t_0;\mu_0) &= -\omega_0\sin\omega_0 t_0 + \mu_0^2\left(\frac{\omega_0}{16}\sin\omega_0 t_0 + \frac{\omega_0^2 t_0}{16}\cos\omega_0 t_0 + \frac{1}{192}(-\omega_0\sin\omega_0 t_0 + 3\omega_0\sin 3\omega_0 t_0)\right)\\
&= -\omega_0\sin\frac{\pi}{2} + 0^2(\dots)\\
&= -\omega_0\\
&\neq 0
\end{aligned}$$

  - Thus, there exists $t_1(\mu)$ smooth defined on some neighborhood of $\mu_0 = 0$ satisfying $t_1(0) = \pi/2\omega_0$ and $F(t_1(\mu);\mu) = 0$.
  - We cannot (easily??) obtain $t_1(\mu)$ directly, so we will look for its second-order Taylor expansion

$$t_1(\mu) = \frac{\pi}{2\omega_0} + b_1\mu + b_2\mu^2 + O(\mu^3)$$

  - We need not compute a bunch of derivatives to find b_1, b_2, though. Indeed, we can just substitute into $F(t_1(\mu);\mu) = 0$ and compare different powers of μ. Doing so, we obtain

$$\begin{aligned}
0 &= F(t_1(\mu);\mu)\\
&= \cos\left(\frac{\pi}{2} + \omega_0 b_1\mu + \omega_0 b_2\mu^2 + O(\mu^3)\right)\\
&\quad + \mu^2\left[\frac{1}{16}\left(\frac{\pi}{2} + \omega_0 b_1\mu + \omega_0 b_2\mu^2 + O(\mu^3)\right)\sin\left(\frac{\pi}{2} + \omega_0 b_1\mu + \omega_0 b_2\mu^2 + O(\mu^3)\right)\right.\\
&\quad \left. + \frac{1}{192}\left(\cos\left(\frac{\pi}{2} + \omega_0 b_1\mu + \omega_0 b_2\mu^2 + O(\mu^3)\right) - \cos 3\left(\frac{\pi}{2} + \omega_0 b_1\mu + \omega_0 b_2\mu^2 + O(\mu^3)\right)\right)\right]\\
&= -\omega_0 b_1\mu + \left(\frac{\pi}{32} - \omega_0 b_2\right)\mu^2 + O(\mu^3)
\end{aligned}$$

  from which we can determine that

$$\begin{aligned}
0 &= -\omega_0 b_1 & 0 &= \frac{\pi}{32} - \omega_0 b_2\\
b_1 &= 0 & b_2 &= \frac{\pi}{32\omega_0}
\end{aligned}$$

– Thus,

$$\begin{aligned}T(\mu) &= 4 \cdot t_1(\mu) \\ &= \frac{2\pi}{\omega_0} + \frac{\pi}{8\omega_0}\mu^2 + O(\mu^3) \\ &= 2\pi\sqrt{\frac{\ell}{g}}\left(1 + \frac{1}{16}\mu^2 + O(\mu^3)\right)\end{aligned}$$

- We calculate an accumulation that is a perturbation of an ODE in the bonus this week, reproducing Einstein's work.

6.4 Chapter 2: Initial Value Problems

From Teschl (2012).

Section 2.4: Dependence on the Initial Condition

11/15:

- In applications from which ODEs are derived, we usually only know several data approximately. In other words, we're primarily concerned with **well-posed** IVPs.
- **Well-posed** (IVP): An IVP for which, from an intuitive standpoint, small changes in the data result in small changes of the solution.
- That an IVP (under certain conditions) is well-posed will be proven by our next theorem.
- To prove this theorem, we will need the following lemma.
- Lemma 2.7 (Generalized Grönwall's inequality): Suppose $\psi(t)$ satisfies

$$\psi(t) \leq \alpha(t) + \int_0^t \beta(s)\psi(s)\,\mathrm{d}s$$

for all $t \in [0,T]$. Suppose also that $\alpha(t) \in \mathbb{R}$ and $\beta(t) \geq 0$ for all $t \in [0,T]$. Then

$$\psi(t) \leq \alpha(t) + \int_0^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s$$

for all $t \in [0,T]$.
If, in addition, $\alpha(s) \leq \alpha(t)$ for $s \leq t$, then

$$\psi(t) \leq \alpha(t)\exp\left(\int_0^t \beta(s)\,\mathrm{d}s\right)$$

for all $t \in [0,T]$.

Proof. Let

$$\phi(t) := \exp\left(-\int_0^t \beta(s)\,\mathrm{d}s\right)$$

Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\phi(t)\int_0^t \beta(s)\psi(s)\,\mathrm{d}s\right) &= -\beta(t)\phi(t)\cdot\int_0^t \beta(s)\psi(s)\,\mathrm{d}s + \phi(t)\cdot\beta(t)\psi(t) \\ &= \beta(t)\phi(t)\left(\phi(t) - \int_0^t \beta(s)\phi(s)\,\mathrm{d}s\right) \\ &\leq \alpha(t)\beta(t)\phi(t)\end{aligned}$$

where the first equality holds by the product rule and the FTC, and the last inequality above holds by the first assumption in the statement of the lemma. Integrating the above inequality with respect to t and dividing the result by $\phi(t)$ shows that

$$\int_0^t \beta(s)\psi(s)\,\mathrm{d}s \leq \int_0^t \alpha(s)\beta(s)\frac{\phi(s)}{\phi(t)}\,\mathrm{d}s\,\alpha(t) + \int_0^t \beta(s)\psi(s)\,\mathrm{d}s \quad \leq \alpha(t) + \int_0^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s$$

It follows that

$$\psi(t) \leq \alpha(t) + \int_0^t \beta(s)\psi(s)\,\mathrm{d}s \leq \alpha(t) + \int_0^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s$$

as desired.

The proof of the second claim is covered in Problem 2.11 (and is not applicable to course content). □

- A simple consequence of the generalized Grönwall's inequality.
  - If
    $$\psi(t) \leq \alpha + \int_0^t (\beta\psi(s) + \gamma)\,\mathrm{d}s$$
    for all $t \in [0, T]$, where $\alpha, \gamma \in \mathbb{R}$ and $\beta \geq 0$, then
    $$\psi(t) \leq \alpha \mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)$$
    for all $t \in [0, T]$.
  - See Problem 2.12 for the proof.
- We can now show that the IVP is well-posed.
- Theorem 2.8: Suppose $f, g \in C(U, \mathbb{R}^n)$ and let f be locally Lipschitz continuous in the second argument, uniformly with respect to the first. If $x(t), y(t)$ are respective solutions of the IVPs

$$\begin{aligned} \dot{x} &= f(t,x) \\ x(t_0) &= x_0 \end{aligned} \qquad\qquad \begin{aligned} \dot{y} &= g(t,y) \\ y(t_0) &= y_0 \end{aligned}$$

then

$$|x(t) - y(t)| \leq |x_0 - y_0|\mathrm{e}^{L|t-t_0|} + \frac{M}{L}(\mathrm{e}^{L|t-t_0|} - 1)$$

where L is the Lipschitz constant of $f : V \to \mathbb{R}^n$, $M = \|f - g\|$ for $f, g : V \to \mathbb{R}^n$, and $V \subset U$ contains $G(x), G(y)$.

Proof. WLOG let $t_0 = 0$. Then

$$\begin{aligned} |x(t) - y(t)| &\leq |x_0 - y_0| + \int_0^t |f(s, x(s)) - g(s, y(s))|\,\mathrm{d}s \\ &\leq |x_0 - y_0| + \int_0^t (L|x(s) - y(s)| + M)\,\mathrm{d}s \end{aligned}$$

Thus, taking

$$\underbrace{|x(t) - y(t)|}_{\psi(t)} \leq \underbrace{|x_0 - y_0|}_{\alpha} + \int_0^t (\underbrace{L}_{\beta}\underbrace{|x(s) - y(s)|}_{\phi(s)} + \underbrace{M}_{\gamma})\,\mathrm{d}s$$

we have by the above consequence of the generalized Grönwall's inequality that

$$|x(t) - y(t)| \leq |x_0 - y_0|\mathrm{e}^{Lt} + \frac{M}{L}(\mathrm{e}^{Lt} - 1)$$

as desired. □

- Establishing continuous dependence on the initial condition.
  - Denote the solution of the IVP by $\phi(t,t_0,x_0)$ to emphasize the dependence on the initial condition.
  - Then in the special case $f=g$ (i.e., where $M=0$), Theorem 2.8 implies that
$$|\phi(t,t_0,x_0)-\phi(t,t_0,y_0)| \leq |x_0-y_0|\mathrm{e}^{L|t-t_0|}$$
  - In other words, ϕ depends continuously on the initial value.
  - Of course, this bound blows up exponentially as t increases, but the linear equation $\dot{x}=x$ shows that we cannot define a better bound in general.
- We now formalize the above notion.
- Theorem 2.9: Suppose $f\in C(U,\mathbb{R}^n)$ is locally Lipschitz continuous in the second argument, uniformly with respect to the first. Around each point $(t_0,x_0)\in U$, we can find a compact set $I\times B\subset U$ such that $\phi(t,s,x)\in C(I\times I\times B,\mathbb{R}^n)$. Moreover, $\phi(t,t_0,x_0)$ is Lipschitz continuous with
$$|\phi(t,t_0,x_0)-\phi(s,s_0,y_0)| \leq |x_0-y_0|\mathrm{e}^{L|t-t_0|}+(|t-s|+|t_0-s_0|\mathrm{e}^{L|t-s_0|})M$$
where L is the Lipschitz constant of $f:V\to\mathbb{R}^n$, $M=\|f\|$ for $f:V\to\mathbb{R}^n$, and $V\subset U$ compact contains $I\times\phi(I\times I\times B)$.

Proof. By the Picard-Lindelöf theorem, there exists $V=[t_0-\varepsilon,t_0+\varepsilon]\times\overline{B_\delta(x_0)}$ such that $\phi(t,t_0,x_0)$ exists and is continuous for $|t-t_0|<\varepsilon$. It can be shown that $\phi(t,t_1,x_1)$ exists for $|t-t_1|\leq\varepsilon/2$, provided that $|t_1-t_0|\leq\varepsilon/2$ and $|x_1-x_0|\leq\delta/2$. Thus, choose $I=[t_0-\varepsilon/2,t_0+\varepsilon/2]$ and $B=\overline{B_{\delta/2}(x_0)}$.

Moreover, we have that
$$\begin{aligned}
|\phi(t,t_0,x_0)-\phi(s,s_0,y_0)| &\leq |\phi(t,t_0,x_0)-\phi(t,t_0,y_0)|\\
&\quad+|\phi(t,t_0,y_0)-\phi(t,s_0,y_0)|\\
&\quad+|\phi(t,s_0,y_0)-\phi(s,s_0,y_0)|\\
&\leq |x_0-y_0|\mathrm{e}^{L|t-t_0|}\\
&\quad+\left|\int_{t_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\
&\quad+\left|\int_s^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|
\end{aligned}$$
We estimated the first term using the note on continuous dependence directly preceding this theorem and proof. We can estimate the third term to be $M|t-s|$. We can estimate the second term as follows: Abbreviating $\Delta(t):=|\phi(t,t_0,y_0)-\phi(t,s_0,y_0)|$ and assuming WLOG that $t_0\leq s_0\leq t$, we have that
$$\begin{aligned}
\Delta(t) &= \left|\int_{t_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\
&= \left|\int_{t_0}^{s_0} f(r,\phi(r,t_0,y_0))\,\mathrm{d}r+\int_{s_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\
&\leq \int_{t_0}^{s_0}|f(r,\phi(r,t_0,y_0))|\,\mathrm{d}r+\int_{s_0}^t|f(r,\phi(r,t_0,y_0))-f(r,\phi(r,s_0,y_0))|\,\mathrm{d}r\\
&\leq |t_0-s_0|M+L\int_{s_0}^t\Delta(r)\,\mathrm{d}r
\end{aligned}$$
Therefore, by Grönwall's inequality (as in Problem 2.12; note that $\gamma=0$ here),
$$\Delta(t)\leq|t_0-s_0|M\mathrm{e}^{L|t-s_0|}$$
as desired. □

- By Problem 1.8, we have $\phi(t,t_0,x_0)=\phi(t-t_0,0,x_0)$ for an autonomous system, so we may consider $\phi(t,x_0)=\phi(t,0,x_0)$.
- The previous result of *continuous* dependence on initial conditions is not good enough in every situation; sometimes, we need to be able to *differentiate* with respect to the initial condition.

11/23:

- Before we prove that a certain set of conditions will imply the existence of the derivative of a solution ϕ with respect to x, we will assume such a derivative exists and investigate some of its properties (so as to motivate the following theorem and proof).
  - Suppose $\phi(t,t_0,x)$ is differentiable with respect to x.
  - If we assume $f\in C^k(U,\mathbb{R}^n)$ for some $k\geq 1$ and ϕ is sufficiently regular that its partial derivatives commute as well, then we may write[2]

$$\begin{aligned}\dot{\phi}(t,t_0,x)&=f(t,\phi(t,t_0,x))\\ \frac{\partial\phi}{\partial t}&=f(t,\phi)\\ \frac{\partial^2\phi}{\partial x\partial t}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\\ \frac{\partial^2\phi}{\partial t\partial x}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\\ \frac{\partial}{\partial t}\frac{\partial\phi}{\partial x}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\end{aligned}$$

  - Thus, if $\partial\phi/\partial x$ exists, then it satisfies the **first variational equation** with $A(t,x)=\partial f/\partial\phi$.
  - It also necessarily satisfies the corresponding integral equation

$$\underbrace{\frac{\partial\phi}{\partial x}(t,t_0,x)}_{y(t)}=\underbrace{\frac{\partial\phi}{\partial x}(t_0,t_0,x)}_{\mathbb{I}}+\int_{t_0}^{t}\underbrace{\frac{\partial f}{\partial\phi}(s,\phi(s,t_0,x))}_{A(s,x)}\cdot\underbrace{\frac{\partial\phi}{\partial x}(s,t_0,x)}_{y(s)}\,\mathrm{d}s$$

- **First variational equation**: The following differential equation. *Given by*

$$\dot{y}=A(t,x)y$$

  - Which is the correct first variational equation?? This one, or the one with $F(t;\mu)$ term?
- Indeed, we have shown that $\partial\phi/\partial x$ necessarily satisfies the above integral equation if it exists. But using fixed point techniques, we can further show that said equation *has* a solution (which is $\partial\phi/\partial x$) under certain conditions. We will not go into this argument in depth, but it will be used in our next theorem.
- Theorem 2.10: Suppose $f\in C^k(U,\mathbb{R}^n)$, $k\geq 1$. Around each point $(t_0,x_0)\in U$, we can find an open set $I\times B\subset U$ such that $\phi(t,s,x)\in C^k(I\times I\times B,\mathbb{R}^n)$. Moreover, $\partial\phi/\partial t\,(t,s,x)\in C^k(I\times I\times B,\mathbb{R}^n)$ and if D_k is a partial derivative of order k, then $D_k\phi$ satisfies the higher order variational equation obtained from

$$\frac{\partial}{\partial t}D_k\phi(t,s,x)=D_k\frac{\partial}{\partial t}\phi(t,s,x)=D_kf(t,\phi(t,s,x))$$

  by applying the chain rule repeatedly. In particular, this equation is linear in $D_k\phi$ and it also follows that the corresponding higher-order derivatives commute.

  Proof. To prove the claim, we induct on k. For the base case $k=1$, we need only prove that $\phi(t,x)$ is *differentiable* at every $x_1\in B$. We now define some terms that will be useful in our argument.

[2] Note that dot-denoted derivatives refer to derivatives with respect to t.

Using the trick on Teschl (2012, p. 7), add t to the dependent variables to make the ODE autonomous. This allows us to consider $\phi(t,x)=\phi(t,0,x)$ WLOG. Thus, we have by Theorem 2.9 that there exists a set $I\times B\subset U$ such that $\phi(t,x_0)\in C(I\times B,\mathbb{R}^n)$. We may take $I=(-T,T)$ and $B=B_\delta(x_0)$ for some $T,\delta>0$ such that $\overline{I\times B}\subset U$. Additionally, let $x_1=0$ WLOG. Furthermore, let $\phi(t):=\phi(t,x_1)$, $A(t):=A(t,x_1)$, and $\psi(t)$ denote the solution to the first variational equation $\dot{\psi}(t)=A(t)\psi(t)$ corresponding to the initial condition $\psi(t_0)=\mathbb{I}$ ($\psi(t)$ is guaranteed to exist by the aforementioned fixed point argument). Lastly, let

$$\theta(t,x):=\frac{\phi(t,x)-\phi(t)-\psi(t)x}{|x|}$$

The proof strategy is thus: If we can show that $\lim_{x\to x_1=0}\theta(t,x)=0$, then we will have proven that $\partial\phi/\partial x$ exists and is equal to ψ. To do so, we will derive a bound on $|\theta(t,x)|$ that we can more easily control. Let's begin.

Since $f\in C^1$, we have that

$$\begin{aligned}
f(y)-f(x)&=\int_x^y\frac{\partial f}{\partial x}(t)\,\mathrm{d}t\\
\frac{f(y)-f(x)}{y-x}&=\frac{1}{y-x}\int_x^y\frac{\partial f}{\partial x}(t)\,\mathrm{d}t\\
&=\int_0^1\frac{\partial f}{\partial x}(x+t(y-x))\,\mathrm{d}t\\
&=\frac{\partial f}{\partial x}(x)+\int_0^1\left(\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right)\mathrm{d}t\\
f(y)-f(x)&=\frac{\partial f}{\partial x}(x)(y-x)+|y-x|R(y,x)
\end{aligned}$$

where

$$|R(y-x)|\leq\max_{t\in[0,1]}\left\|\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right\|$$

Let's take another minute to justify the above algebraic manipulations from a geometric perspective. The first line should be a fairly straightforward application of the FTC. The second line rephrases the equality as two different ways of calculating the average slope of f on $[x,y]$; specifically, we may either do rise over run (LHS) or find the area under the derivative and divide by the "length" to get the average "height." The third line shrinks the domain of integration to an interval of unit length, compressing all of the information in $\partial f/\partial x$ along with it, and making it so that we no longer have to divide through by the "length." The fourth line adds and subtracts the starting point so that geometrically, we take the area in two parts: All of the area beneath the first point, and then all of the area that "changes." The last line sees us multiply both sides by $y-x$ and then take the integral to be a sort of "remainder." Alternatively, we may justify

$$f(y)=f(x)+\frac{\partial f}{\partial x}(x)(y-x)+\left(\int_0^1\left(\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right)\mathrm{d}t\right)(y-x)$$

by analogy: If $f(x)=x^2$ and thus $\mathrm{d}f/\mathrm{d}x=2x$, then 4^2 equals 2^2, plus the area under the curve $2x$ from $x=2$ to $x=4$ partitioned into the rectangle with base length $y-x=2$ and height $\mathrm{d}f/\mathrm{d}x\,(2)=4$ and the triangle sitting on top of the aforementioned rectangle and below the graph of $\mathrm{d}f/\mathrm{d}x$. Note that all of these justifications are taken from the perspective of f being a single-variable function; to justify the transformations in the multivariable case would likely be much more complicated and is beyond my grasp at the moment (recall that $\partial f/\partial x$ is actually a matrix!). Also note that as such, the norm in the bound on $|R(y,x)|$ given above is the matrix norm.

Since $f\in C^1$, its partial derivatives are uniformly continuous in a neighborhood of x_1. It follows that $\lim_{y\to x}|R(y,x)|=0$ where the argument converges uniformly in x in some neighborhood of $x_1=0$.

Using the above expression for $f(y)-f(x)$, we have that

$$
\begin{aligned}
\dot{\theta}(t,x) &= \frac{1}{|x|}[\dot{\phi}(t,x)-\dot{\phi}(t)-\dot{\psi}(t)x] \\
&= \frac{1}{|x|}[f(\phi(t,x))-f(\phi(t))-A(t)\psi(t)x] \\
&= \frac{1}{|x|}\left[\frac{\partial f}{\partial x}(\phi(t))(\phi(t,x)-\phi(t))+|\phi(t,x)-\phi(t)|R(\phi(t,x),\phi(t))-A(t)\psi(t)x\right] \\
&= A(t)\cdot\frac{\phi(t,x)-\phi(t)-\psi(t)x}{|x|}+\frac{|\phi(t,x)-\phi(t)|}{|x|}R(\phi(t,x),\phi(t)) \\
&= A(t)\theta(t,x)+\frac{|\phi(t,x)-\phi(t)|}{|x|}R(\phi(t,x),\phi(t))
\end{aligned}
$$

Integrating and taking absolute values yields

$$
\begin{aligned}
|\theta(t,x)| &= |\theta(t,x)-\theta(0,x)| \\
&= \left|\int_0^t\left(A(s)\theta(s,x)+\frac{|\phi(s,x)-\phi(s)|}{|x|}R(\phi(s,x),\phi(s))\right)\mathrm{d}s\right| \\
&\leq \int_0^t\frac{1}{|x|}|\phi(s,x)-\phi(s,0)|\cdot|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s \\
&\leq \int_0^t\frac{1}{|x|}|x-0|\mathrm{e}^{L|s-0|}\cdot|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s \\
&\leq \mathrm{e}^{LT}\int_0^T|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s
\end{aligned}
$$

Note that $\theta(0,x)=0$ since we have taken $t_0=x_1=0$ WLOG. Also note that we use the continuous dependence on initial conditions equation to get from line 2 to line 3. Lastly, note that from the next to the last to the last line, we use the inequality $\mathrm{e}^{Ls}\leq\mathrm{e}^{LT}$ to transform the exponential function into a constant that bounds it (recall that $s\leq T$ by definition).

Define

$$\tilde{R}(x)=\mathrm{e}^{LT}\int_0^T|R(\phi(s,x),\phi(s))|\,\mathrm{d}s$$

so that

$$|\theta(t,x)|\leq\tilde{R}(x)+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s$$

It follows by the Generalized Grönwall's inequality that

$$|\theta(t,x)|\leq\tilde{R}(x)\exp\left(\int_0^T\|A(s)\|\,\mathrm{d}s\right)$$

This combined with the fact from earlier that $\lim_{y\to x}|R(y,x)|=0$ and hence $\lim_{x\to 0}\tilde{R}(x)=0$ implies that $\lim_{x\to 0}\theta(t,x)=0$, as desired.

Additionally, $\partial\phi/\partial x$ is continuous as the solution to the first variational equation. This completes the base case.

Now suppose via strong induction that the claim holds for $1,\dots,k$, and let $f\in C^{k+1}$. Then $\phi(t,x)\in C^1$ and $\partial\phi/\partial x$ solves the first variational equation, as per the base case. But since $A(t,x)\in C^k$ and hence $\partial\phi/\partial x\in C^k$, we have by Lemma 2.3 that $\phi(t,x)\in C^{k+1}$. $\square$

- This theorem also allows us to handle dependence on parameters.

- In particular, if f depends on some paramenters $\lambda \in \Lambda \subset \mathbb{R}^p$ such that we are now solving the IVP

$$\dot{x}(t) = f(t, x, \lambda), \quad x(t_0) = x_0$$

for a solution $\phi(t, t_0, x_0, \lambda)$, then we have the following result.

- Theorem 2.11: Suppose $f \in C^k(U \times \Lambda, \mathbb{R}^n)$, $k \geq 1$. Around each point $(t_0, x_0, \lambda_0) \in U \times \Lambda$ we can find an open set $I \times B \times \Lambda_0 \subset U \times \Lambda$ such that $\phi(t, s, x, \lambda) \in C^k(I \times I \times B \times \Lambda_0, \mathbb{R}^n)$.

Proof. Largely follows from Theorem 2.10. Noteworthy modificatinos: We add the parameters λ to the dependent variables and require $\dot{\lambda} = 0$ (i.e., that the parameters do not change with time and therefore uniquely determine a solution over time). ☐

Problems

11/15: **2.12.** Show that if

$$\psi(t) \leq \alpha + \int_0^t (\beta\psi(s) + \gamma)\,\mathrm{d}s$$

for all $t \in [0, T]$, where $\alpha, \gamma \in \mathbb{R}$ and $\beta \geq 0$, then

$$\psi(t) \leq \alpha \mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)$$

for all $t \in [0, T]$. *Hint*: Introduce $\tilde{\psi}(t) = \psi(t) + \gamma/\beta$.

Proof. Taking the hint and substituting $\psi(t) = \tilde{\psi}(t) - \gamma/\beta$, we get

$$\begin{aligned}
\tilde{\psi}(t) - \frac{\gamma}{\beta} &\leq \alpha + \int_0^t \left(\beta\left(\tilde{\psi}(t) - \frac{\gamma}{\beta}\right) + \gamma\right)\mathrm{d}s \\
\tilde{\psi}(t) &\leq \alpha + \frac{\gamma}{\beta} + \int_0^t \beta\tilde{\psi}(t)\,\mathrm{d}s
\end{aligned}$$

It follows by the generalized Grönwall's inequality that

$$\begin{aligned}
\tilde{\psi}(t) &\leq \alpha + \frac{\gamma}{\beta} + \int_0^t \left(\alpha + \frac{\gamma}{\beta}\right)\beta\exp\left(\int_s^t \beta\,\mathrm{d}r\right)\mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + \int_0^t (\alpha\beta + \gamma)\mathrm{e}^{\beta(t-s)}\,\mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + (\alpha\beta + \gamma)\mathrm{e}^{\beta t}\int_0^t \mathrm{e}^{-\beta s}\,\mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + (\alpha\beta + \gamma)\mathrm{e}^{\beta t}\left(\frac{1}{-\beta}\mathrm{e}^{-\beta t} - \frac{1}{-\beta}\mathrm{e}^0\right) \\
&= \alpha + \frac{\gamma}{\beta} + \left(\alpha + \frac{\gamma}{\beta}\right)\mathrm{e}^{\beta t}(1 - \mathrm{e}^{-\beta t}) \\
&= \alpha + \frac{\gamma}{\beta} + \left(\alpha + \frac{\gamma}{\beta}\right)(\mathrm{e}^{\beta t} - 1) \\
&= \alpha + \frac{\gamma}{\beta} + \alpha\mathrm{e}^{\beta t} - \alpha + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1) \\
&= \frac{\gamma}{\beta} + \alpha\mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)
\end{aligned}$$

Subtracting γ/β from both sides and returning the substitution yields the desired result. ☐

Section 2.5: Regular Perturbation Theory

12/6:
- Goal of this section: Justify the perturbation method proposed in Problem 1.2 using Theorem 2.11.
- **Regular perturbation problem**: An IVP of the following form. *Given by*

$$\dot{x} = f(t, x, \varepsilon), \quad x(t_0) = x_0$$

- Let's try to solve such a problem in the general case first.
  - Theorem 2.11: $f \in C^1$ implies $\phi(t, \varepsilon) = \phi(t, \varepsilon, t_0, x_0) \in C^1$.
  - If $\phi \in C^1$, then we have the following Taylor expansion.

$$\phi(t, \varepsilon) = \phi_0(t) + \phi_1(t)\varepsilon + o(\varepsilon)$$

  - If $\varepsilon = 0$, $\phi(t, \varepsilon) = \phi_0(t)$. Thus, $\phi_0(t)$ is the solution to the unperturbed equation

$$\dot{\phi}_0 = f_0(t, \phi_0), \quad \phi_0(t_0) = x_0$$

    where $f_0(t, x) = f(t, x, 0)$.
  - Additionally, we can deduce from the Taylor series expansion that

$$\begin{aligned} \phi &= \phi_0 + \phi_1 \varepsilon \\ \frac{\partial \phi}{\partial \varepsilon} &= \phi_1 \end{aligned}$$

  - Thus, we have from the original ODE a new kind of first variational equation satisfied by $\phi_1 = \partial \phi / \partial \varepsilon$, which we can use to solve for said quantity.

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= f(t, \phi(t, \varepsilon), \varepsilon) \\ \frac{\partial^2 \phi}{\partial \varepsilon \partial t} &= \frac{\partial f}{\partial t}\frac{\partial t}{\partial \varepsilon} + \frac{\partial f}{\partial x}\frac{\partial \phi}{\partial \varepsilon} + \frac{\partial f}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \varepsilon} \\ \frac{\partial}{\partial t}\frac{\partial \phi}{\partial \varepsilon} &= \frac{\partial f}{\partial x}\frac{\partial \phi}{\partial \varepsilon} + \frac{\partial f}{\partial \varepsilon} \\ \frac{\partial}{\partial t} \underbrace{\frac{\partial}{\partial \varepsilon}\phi(t, \varepsilon)\bigg|_{\varepsilon = 0}}_{\phi_1} &= \underbrace{\frac{\partial}{\partial x} f(t, \phi_0(t), 0)\bigg|_{\varepsilon = 0}}_{f_{10}(t, \phi_0(t))} \phi_1 + \underbrace{\frac{\partial}{\partial \varepsilon} f(t, \phi_0(t), \varepsilon)\bigg|_{\varepsilon = 0}}_{f_{11}(t, \phi_0(t))} \end{aligned}$$

    Note that the corresponding initial condition is

$$\begin{aligned} x(t_0) &= x_0 \\ \phi(t_0, \varepsilon) &= x_0 \\ \frac{\partial}{\partial \varepsilon}\phi(t_0, \varepsilon)\bigg|_{\varepsilon = 0} &= 0 \\ \phi_1(t_0) &= 0 \end{aligned}$$

  - "Hence, once we have the solution of the unperturbed problem $\phi_0(t)$, we can then compute the correction term $\phi_1(t)$ by solving another linear equation" (Teschl, 2012, p. 49).
  - Therefore, the "plug in the ansatz and compare coefficients" procedure is justified.
- Example: Consider the equation

$$\dot{v} = -\varepsilon v - g, \quad v(0) = 0, \quad \varepsilon \geq 0$$

  which models the velocity of a falling object with air resistance.

- We can calculate the explicit solution and use it to check our answer.
$$\phi(t,\varepsilon) = \frac{g}{\varepsilon}(\mathrm{e}^{-\varepsilon t} - 1)$$
- Applying perturbation techniques, we start with the unperturbed problem
$$\dot{\phi}_0 = -g, \quad \phi_0(0) = 0$$
and solve to get
$$\phi_0(t) = -gt$$
- Now we use the first variational equation. We have
$$\begin{aligned} f_{10}(t,\phi_0(t)) &= \frac{\partial}{\partial x} f(t,\phi_0(t),0)\bigg|_{\varepsilon=0} \\ &= \frac{\partial}{\partial x}(-0\cdot\phi_0(t) - g)\bigg|_{\varepsilon=0} \\ &= \frac{\partial}{\partial x}(-g)\bigg|_{\varepsilon=0} \\ &= 0|_{\varepsilon=0} \\ &= 0 \end{aligned} \qquad \begin{aligned} f_{11}(t,\phi_0(t)) &= \frac{\partial}{\partial \varepsilon} f(t,\phi_0(t),\varepsilon)\bigg|_{\varepsilon=0} \\ &= \frac{\partial}{\partial \varepsilon}(-\varepsilon\phi_0(t) - g)\bigg|_{\varepsilon=0} \\ &= -\phi_0(t)|_{\varepsilon=0} \\ &= gt|_{\varepsilon=0} \\ &= gt \end{aligned}$$
so that
$$\begin{aligned} \dot{\phi}_1 &= 0\cdot\phi_1 + gt \\ \dot{\phi}_1 &= gt, \quad \phi_1(0) = 0 \end{aligned}$$
which we can solve to get
$$\phi_1(t) = \frac{g}{2}t^2$$
- Therefore, our overall approximation is
$$v(t) = -g\left(t - \varepsilon\cdot\frac{t^2}{2} + o(\varepsilon)\right)$$
which does indeed correspond to the Taylor expansion of the exact solution.
- Note that the approximation is only valid for fixed time and will get worse as t increases; indeed, the appoximation diverges in the long run while the actual solution converges to g/ε.

- Extending this procedure.
- Theorem 2.12: Let Λ be some open interval. Suppose $f \in C^k(U\times\Lambda,\mathbb{R}^n)$, $k \geq 1$ and fix some values $(t_0, x_0, \varepsilon_0) \in U \times \Lambda$. Let $\phi(t,\varepsilon) \in C^k(I\times\Lambda_0,\mathbb{R}^n)$ be the solution of the initial value problem
$$\dot{x} = f(t,x,\varepsilon), \quad x(t_0) = x_0$$
guaranteed to exist by Theorem 2.11. Then
$$\phi(t,\varepsilon) = \sum_{j=0}^{k} \frac{\phi_j(t)}{j!}(\varepsilon-\varepsilon_0)^j + o((\varepsilon-\varepsilon_0)^k)$$
where the coefficients can be obtained by recursively solving
$$\dot{\phi}_j = f_j(t,\phi_0,\dots,\phi_j,\varepsilon_0), \quad \phi_j(t_0) = \begin{cases} x_0 & j = 0 \\ 0 & j \geq 1 \end{cases}$$

where the function f_j is recursively defined via

$$f_{j+1}(t, x_0, \ldots, x_{j+1}, \varepsilon) = \frac{\partial f_j}{\partial \varepsilon}(t, x_0, \ldots, x_j, \varepsilon) + \sum_{k=0}^{j} \frac{\partial f_j}{\partial x_k}(t, x_0, \ldots, x_j, \varepsilon) x_{k+1}, \quad f_0(t, x_0, \varepsilon) = f(t, x_0, \varepsilon)$$

If we assume $f \in C^{k+1}$, the error term will be $O((\varepsilon - \varepsilon_0)^{k+1})$ uniformly for $t \in I$.

Proof. We plug the power series definition of $\phi(t, \varepsilon)$ given above into the differential equation and compare powers of ε.

Estimating the remainder in the Taylor expansion: Since $f \in C^{k+1}$, we know that $\partial^{k+1}\phi / \partial\varepsilon^{k+1}$ is continuous and hence bounded on $I \times \Lambda_0$, as desired. □

- If we're willing to deal with the Taylor series in more than one variable, we can definitely admit more than one parameter.
- Include the case where the initial condition depends on ε by simply replacing the initial conditions for $\phi_j(t_0)$ by the corresponding expansion coefficients of $x_0(\varepsilon)$.
- The Taylor series will converge if f is analytic with respect to all variables (see Theorem 4.2).

Problems

2.16. Compute the next term ϕ_2 in the above example.

Week 7

Solution Existence and Stability

7.1 Peano Existence Theorem

11/7:

- Today: Peano Existence Theorem.
- For an IVP of a first-order differential system, as long as the RHS is continuous, we get at least one solution.
- The proof provides an algorithm that can be really useful in computing the solution provided that uniqueness exists.
- We will need a theorem from analysis to start.
- Theorem (Arzelà-Ascoli[1]): Let $h_k : [a,b] \to \mathbb{R}^n$ be a sequence of functions that is uniformly bounded and uniformly Lipschitz continuous wrt. L. Then $\{h_k\}$ contains a uniformly convergent subsequence and the limit has the same bound and Lipschitz constant.

  Proof. Recall the property of sequential compactness[2], i.e., that every bounded sequence of numbers contains a convergent subsequence. We want to prove this for a sequence of functions. To do so, we will need the Cantor diagonalization technique.

  $\mathbb{Q}$ is countable. Thus, we can enumerate the rationals in $[a,b]$ by $r_1, r_2, r_3, \dots$. Since $\{h_k(r_1)\}$ is a bounded sequence of numbers, we have by the above that there is a subsequence C_1 — say $h_1^{(1)}, h_2^{(1)}, h_3^{(1)}, \dots$ — such that $C_1 = \{h_k^{(1)}(r_1)\}$ is a convergent subsequence in $\mathbb{R}^n$ of the original sequence. Now C_1 is still a bounded sequence, so we can obtain a subsequence C_2 of *it* — say $h_1^{(2)}, h_2^{(2)}, h_3^{(2)}, \dots$ — such that $C_2 = \{h_k^{(2)}(r_2)\}$ is a convergent subsequence in $\mathbb{R}^n$ at r_2 (and, by inductive hypothesis, at r_1!). Inductively, we can obtain $C_\ell = \{h_k^{(\ell)}\}_{\ell,k=1}^\infty$ convergent at $r_1, r_2, \dots, r_\ell$. We then write down the elements of the sequences as a table. (For example, the k^{th} row of the table is a sequence that converges at $r_1, \dots, r_k$.)

$$\begin{matrix}
h_1^{(1)} & h_2^{(1)} & h_3^{(1)} & \cdots \\
h_1^{(2)} & h_2^{(2)} & h_3^{(2)} & \cdots \\
h_1^{(3)} & h_2^{(3)} & h_3^{(3)} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{matrix}$$

  Consider the diagonal sequence $\{f_\ell\}_{\ell=1}^\infty$ where $f_\ell = h_\ell^{(\ell)}$. By definition, it converges at all rational points. We now seek to prove that it converges uniformly at *all* points.

[1] This is not the full Arzelà-Ascoli theorem, but a special case. The proof is similar, regardless, though. See Honors Analysis in $\mathbb{R}^n$ I Notes.

[2] The Bolzano-Weierstrass Theorem/Theorem 15.18 from Honors Calculus IBL.

To prove that $\{f_\ell\}$ is a uniformly convergent sequence of functions, it will suffice to show that for all $\varepsilon > 0$, there exists N such that if $k, \ell > N$, then $|f_k(t) - f_\ell(t)| < \varepsilon$ for all $t \in [a,b]$. Let $\varepsilon > 0$ be arbitrary. Divide $[a,b]$ into m congruent subintervals I_α ($\alpha = 1, \dots, m$) such that $|I_\alpha| \leq \varepsilon/3L$ for all α. This guarantees that the oscillation of each f_k on any I_α is $\leq \varepsilon/3$ since if $x, y \in I_\alpha$ for some α, then

$$|f_\ell(x) - f_\ell(y)| \leq L|x-y| \leq L \cdot \frac{\varepsilon}{3L} = \frac{\varepsilon}{3}$$

Using the fact that $\{f_\ell\}$ is convergent and hence Cauchy on the rationals, pick N large enough so that $r_\alpha \in I_\alpha$ implies $|f_k(r_\alpha) - f_\ell(r_\alpha)| < \varepsilon/3$ for $k, \ell > N$. We will choose this N to be our N. Now let $t \in [a,b]$ be arbitrary. By their definition, we know $t \in I_\alpha$ for some α. Therefore,

$$\begin{aligned}
|f_k(t) - f_\ell(t)| &\leq |f_k(t) - f_k(r_\alpha)| + |f_k(r_\alpha) - f_\ell(r_\alpha)| + |f_\ell(r_\alpha) - f_\ell(t)| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon
\end{aligned}$$

as desired.

Lastly, we can prove that the limit function f of $\{f_\ell\}$ is L-Lipschitz as follows. Let $t, t' \in [a,b]$ be arbitrary. Then

$$\left|\frac{f(t) - f(t')}{t - t'}\right| = \lim_{k \to \infty} \left|\frac{f_k(t) - f_k(t')}{t - t'}\right| \leq \lim_{k \to \infty} \left|\frac{L|t - t'|}{t - t'}\right| = \lim_{k \to \infty} |L| = L$$

as desired. □

- Now we come to the proof of the Peano Existence Theorem.
- Theorem (Peano Existence Theorem): Let $f : [t_0, t_0 + a] \times \bar{B}(y_0, b) \to \mathbb{R}^n$ be bounded ($|f(t,z)| \leq M$) and continuous. Then the IVP
$$y'(t) = f(t, y(t)), \quad y(t_0) = b$$
has at least one solution for $t \in [t_0, t_0 + T]$ where $T = \min(a, b/M)$.

Proof. Since there is no Lipschitz condition, we use another strategy to find approximate solutions.

picture Fix $T = \min(a, b/M)$. We divide $[t_0, t_0 + T]$ into m congruent closed subintervals I_α ($\alpha = 0, \dots, m-1$), each of length $h_m = T/m$. Define a continuous function $y_m(t)$ as follows: The values at the nodes t_α (the intersection points of adjacent congruent subintervals) are defined inductively via

$$y_m(t_{\alpha+1}) = y_m(t_\alpha) + f(t_\alpha, y_m(t_\alpha))h_m$$

for $\alpha = 0, \dots, m-1$, and y_m is taken to be linear between the nodes[3]. The idea is that we replace the derivative $y'(t)$ by the difference quotient $[y(t+h) - y(t)]/h$. It follows by the construction that every function in the set $\{y_k(t) : [t_0, t_0 + T] \to \bar{B}(y_0, b)\}$ is piecewise linear (hence continuous), uniformly bounded, and uniformly M-Lipschitz continuous. Therefore, by the Arzelá-Ascoli theorem, $\{y_k\}$ contains a uniformly convergent subsequence $y_{m_k} \to y$.

It remains to verify that y is a solution to the integral equation

$$y(t) = y_0 + \int_{t_0}^{t} f(\tau, y(\tau))\mathrm{d}\tau$$

Observe that the domain of f is a closed and bounded subset of the real numbers. Thus, it is compact by the Heine-Borel theorem[4]. Moreover, since f is a continuous function on a compact domain, we

[3]Note that this construction is quite similar to that employed in Euler's method.
[4]Theorem 10.16 of Honors Calculus IBL.

have by the Heine-Cantor theorem[5] that f is uniformly continuous. Thus, for any $\varepsilon > 0$, there exists N such that if $m > N$, then

$$|f(t, y_m(t)) - f(t_\alpha, y_m(t_\alpha))| < \frac{\varepsilon}{T}$$

for all $\alpha = 0, \ldots, m-1$ and $t \in I_\alpha$. Additionally, observe that

$$y_{m_k}(t) = y_0 + \sum_{\alpha=0}^{m-1} \int_{t_\alpha}^{t_{\alpha+1}} \chi_t(\tau) f(t_\alpha, y_{m_k}(t_\alpha)) \mathrm{d}\tau$$

where $\chi_t(\tau)$ denotes the **characteristic function** of $[t_0, t]$. To see this, compare with the original inductive definition of $y_m(t_{\alpha+1})$. *picture* We thus see that y_0 in the above equation corresponds to $y_m(t_0) = y(t_0)$, as we would expect. We see that we are summing a series of side-by-side integrals so that in the end, we integrate over all of $[t_0, t_0 + T]$. We see that the characteristic function restricts us to integrating over the ODE only up until t, as we would want for an approximation $y_{m_k}(t)$ at t using Euler's method. And we see that since $f(t_\alpha, y_{m_k}(t_\alpha))$ is constant and $h_m = t_{\alpha+1} - t_\alpha$, the integral does take on the expected value $f(t_\alpha, y_m(t_\alpha))h_m$. Moving right along, we see that

$$\begin{aligned}
\left| y_{m_k}(t) - y_0 - \int_{t_0}^{t} f(\tau, y_{m_k}(\tau)) \mathrm{d}\tau \right| &\leq \sum_{\alpha=0}^{m-1} \int_{t_\alpha}^{t_{\alpha+1}} \chi_t(\tau) |f(t_\alpha, y_{m_k}(t_\alpha)) - f(\tau, y_{m_k}(\tau))| \mathrm{d}\tau \\
&< \int_{t_0}^{t_0+T} \chi_t(\tau) \cdot \frac{\varepsilon}{T} \mathrm{d}\tau \\
&= \int_{t_0}^{t} \frac{\varepsilon}{T} \mathrm{d}\tau \\
&= \varepsilon \cdot \frac{t - t_0}{T} \\
&\leq \varepsilon
\end{aligned}$$

Thus, by uniform convergence, $\int_{t_0}^{t} f(\tau, y_{m_k}(\tau)) \mathrm{d}\tau \to \int_{t_0}^{t} f(\tau, y(\tau)) \mathrm{d}\tau$ uniformly, so y does satisfy the integral equation, as desired. ☐

- **Characteristic function** (of $[a, b]$): The function defined as follows. *Denoted by* $\boldsymbol{\chi_{[a,b]}}$. *Given by*

$$\chi_{[a,b]}(t) = \begin{cases} 1 & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

- Utility of the Peano Existence Theorem: Proves the *existence* of a solution, but the proof is not constructive; it does not give an algorithm for finding the desired sequence. Nor does the PET make any statement on uniqueness.
- We now look to use a related method to define a sequence of functions that will converge to the desired solution of the ODE.
  - While the PET does not require it, in practice, most f we would be interested in will satisfy an additional Lipschitz condition.
  - Define the integral operator

$$\Phi[u] = y_0 + \int_{t_0}^{t} f(\tau, u(\tau)) \mathrm{d}\tau$$

    We will prove that Φ is a contraction on the function space. This will imply that $\Phi^N[u]$ converges across the entire interval $[t_0, t_0 + T]$ to the solution y for any $u : [t_0, t_0 + T] \to \bar{B}(y_0, b)$, giving us our desired computational strategy. Let's begin.

[5] Theorem 13.6 of Honors Calculus IBL.

- To prove that Φ is a contraction, it will suffice to show that $\left\|\Phi^j[u_1] - \Phi^j[u_2]\right\| \to 0$ as $j \to \infty$. Thus, we wish to put a bound on $\left\|\Phi^j[u_1] - \Phi^j[u_2]\right\|$ that decreases as j increases. To that end, we will prove that

  $$\left\|\Phi^j[u_1] - \Phi^j[u_2]\right\| \leq \frac{(LT)^j}{j!} \cdot \|u_1 - u_2\|$$

  for all j.

  - We induct on j. For the base case $j = 1$, we have that

    $$\begin{aligned}|\Phi[u_1](t) - \Phi[u_2](t)| &\leq \int_{t_0}^{t} L|u_1(\tau) - u_2(\tau)|\mathrm{d}\tau \\ &\leq L(t - t_0)\|u_1 - u_2\| \\ &\leq LT\|u_1 - u_2\| \\ &= \frac{(LT)^1}{1!} \cdot \|u_1 - u_2\|\end{aligned}$$

    for all t.
  - Now suppose inductively that $\left\|\Phi^j[u_1] - \Phi^j[u_2]\right\| \leq (LT)^j/j! \cdot \|u_1 - u_2\|$. Then we have that

    $$\begin{aligned}|\Phi^{j+1}[u_1](t) - \Phi^{j+1}[u_2](t)| &\leq \int_{t_0}^{t} L|\Phi^j[u_1](\tau) - \Phi^j[u_2](\tau)|\mathrm{d}\tau \\ &\leq \int_{t_0}^{t} L \cdot \frac{(LT)^j}{j!} \cdot \|u_1 - u_2\|\mathrm{d}\tau \\ &= \cdots \\ &\leq \frac{(LT)^{j+1}}{j!} \cdot \|u_1 - u_2\|\end{aligned}$$

    for all t, implying the desired result.
- We now estimate the error between y_m and y in terms of y_m, alone. Indeed, we have from the above that

  $$\begin{aligned}\left\|y_m - \Phi^N[y_m]\right\| &\leq \sum_{j=0}^{N-1} \left\|\Phi^j[y_m] - \Phi^{j+1}[y_m]\right\| \\ &\leq \|y_m - \Phi[y_m]\| \sum_{j=0}^{N-1} \frac{(TL)^j}{j!} \\ \|y_m - y\| &\leq \|y_m - \Phi[y_m]\|\mathrm{e}^{TL}\end{aligned}$$

  where we get from the second to the third line by letting $N \to \infty$.

  - The proof of the PET guarantees that $\|y_m - \Phi[y_m]\|$ is small when m is large, no matter whether y_m itself converges or not.
  - In fact, when $f \in C^1$, the error is estimated as

    $$\|y_m - y\| \leq \frac{LT\mathrm{e}^{TL}}{m}$$

    for $L = \|f\|_{C^1}$.

- Takeaway: This polygon method gives rise to an algorithm to solve ODEs. Theoretically, it converges much slower than the Picard iteration, but in practice, it has the advantage that we do not need to do any numerical integration. Indeed, to obtain the desired precision using the Picard iteration, the numerical integration will need more and more steps and the total accumulated error will not be less than this polygon method.

- Better difference methods include Runge-Kutta or Heun, but please refer to monographs on numerical ODEs for these.

7.2 Asymptotic Stability

11/9:
- Going forward, we restrict ourselves to autonomous ODEs $y' = f(y)$, where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field.
- For every $x \in \mathbb{R}^n$, the IVP
$$y' = f(y), \quad y(0) = x$$
has a unique maximal solution $\phi_t(x)$ for $t \in I_x$.
- **Orbit**: The following set. *Given by*
$$\{\phi_t(x) : t \in I_x\}$$
- Let $K \subset \mathbb{R}^n$ be compact.
  - Then there exists $T_K \in \mathbb{R}$ such that $\phi_t(x)$ is defined for all $x \in K$ and $|t| \leq T_K$.
  - Moreover, the map from $K \to \mathbb{R}^n$ defined by $x \mapsto \phi_t(x)$ is injective due to uniqueness (and therefore a **homeomorphism**). We get one such map for each t.
  - Similar to the diffeomorphism idea from Guillemin and Haine (2018).
- **Invariant set**: A subset of $\mathbb{R}^n$ such that any orbit starting within it never leaves it.
- Compact invariant sets are quite interesting.
- Proposition: Let $\Omega \subset \mathbb{R}^n$ be a domain with a piecewise smooth boundary $\partial\Omega$. Suppose $f(x)$ is transversal to $\partial\Omega$ and inward pointing: That is, if ν is the inward pointing unit normal, then $f(x) \cdot \nu(x) \geq 0$ for all $x \in \partial\Omega$. Then $\bar{\Omega}$ is an invariant set: That is, any orbit starting from a point $\bar{\Omega}$ exists throughout the time and never leaves $\bar{\Omega}$.

  Proof idea. $x \in \partial\Omega$ ensures that $\phi_t(x)$ must be in Ω for small t. Hence, it suffices to consider $x \in \Omega$. In that case, pick the smallest $T > 0$ such that $\phi_T(x) \in \partial\Omega$. Then by transversality it must turn back into Ω. $\square$

- This simple proposition is especially useful when establishing global attraction of the orbits.
- **Fixed point**: A point in $\mathbb{R}^n$ at which f evaluates to zero. *Denoted by* $\boldsymbol{x_0}$.
  - This means that the vector at x_0 is zero.
- **Lyapunov stable** (fixed point): A fixed point x_0 such that for any neighborhood $B(x_0, \varepsilon)$, there exists a neighborhood $B(x_0, \delta)$ such that $\phi_t(x) \in B(x_0, \varepsilon)$ for any $t \geq 0$ and $x \in B(x_0, \delta)$.
- **Asymptotically stable** (fixed point): A Lyapunov stable fixed point x_0 such that $\phi_t(x) \to x_0$ as $t \to +\infty$ for $x \in B(x_0, \delta)$.
- Example of a system that is Lyapunov stable but not asymptotically stable: The system
$$y' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} y$$
where A denotes a rotation.
  - The orbits are concentric circles and never converge to 0.
- Investigation: The local behavior near a fixed point.
  - Consider $y' = f(y)$ as a perturbation of the linearized system $y' = f'(x_0)y$. In this case,
$$f(x) = f'(x_0)(x - x_0) + O(|x - x_0|^2)$$
as $x \to x_0$.

- Theorem: Let $f(x_0) = 0$. If the eigenvalues of the linearization $A = f'(x_0)$ all have negative real parts, then the fixed point $x = x_0$ is asymptotically stable.

  Proof. WLOG let $x_0 = 0$. Write $f(x) = Ax + g(x)$, where $g(x) = O(|x|^2)$. Since every $\lambda \in \sigma(A)$ has negative real part, there exist $a, C > 0$ (let $C > 1$ WLOG) such that
  $$|\mathrm{e}^{tA}x| \leq C\mathrm{e}^{-at}|x|$$
  The C arises because the matrix norm of e^{tA} is bounded as $t \to +\infty$ if all eigenvalues are negative. The e^{-at} arises similarly, and reflects the exponential decrease in magnitude happening along all subspaces on which e^{tA} acts.

  Let δ be such that $|g(x)| \leq a|x|/2C$ when $|x| \leq \delta$. Now consider the IVP
  $$y' = Ay + g(y), \quad y(0) \in \bar{B}\left(0, \frac{\delta}{2C}\right)$$
  Then at least for small t (i.e., t such that $|y(t)| \leq \delta$),
  $$|y(t)| \leq C\mathrm{e}^{-at}|y(0)| + \frac{a}{2C}\int_0^t \mathrm{e}^{-a(t-\tau)}|y(\tau)|\mathrm{d}\tau$$
  It follows from Grönwall's inequality that
  $$\mathrm{e}^{at}|y(t)| \leq C|y(0)|\mathrm{e}^{at/2}$$
  hence
  $$|y(t)| \leq \frac{\delta}{2}\mathrm{e}^{-at/2} < \delta$$
  Hence, any orbit of the system starting from $\bar{B}(0, \delta/2C)$ stays in $\bar{B}(0, \delta)$. So the maximal time of existence T is $+\infty$. This is because if not then, then the IVP starting from $y(T)$ is still solvable, contradicting the definition of T. Thus, we have proven that
  $$|y(t)| \leq \frac{\delta}{2}\mathrm{e}^{-at/2}$$
  for all $t \geq 0$ as long as $|y(0)| \leq \delta/2C$. □

- This is the last rigorous proof given in this course.
- A similar theorem:
- Theorem: Let $f(0) = 0$. If one of the eigenvalues of $A = f'(0)$ has positive real part, then the fixed point $x = 0$ is not Lyapunov stable.
- Initial application: Nonlinear mechanical system with frictions, e.g., ideal pendulum with friction.
  $$ml\theta'' + b\theta' = -mg\sin\theta$$
  - Substitute $\eta = b/ml$ and $\omega = \theta'$ to get a nonlinear system
    $$\begin{pmatrix}\theta\\ \omega\end{pmatrix}' = \begin{pmatrix}\omega\\ -\eta\omega - g/l\sin\theta\end{pmatrix}$$
  - At the equilibrium position $(\theta, \omega) = (0, 0)$, we have
    $$A = \begin{pmatrix}\frac{\partial}{\partial\theta}(\omega) & \frac{\partial}{\partial\omega}(\omega)\\ \frac{\partial}{\partial\theta}(-\eta\omega - g/l\sin\theta) & \frac{\partial}{\partial\omega}(-\eta\omega - g/l\sin\theta)\end{pmatrix} \approx \begin{pmatrix}0 & 1\\ -g/l & -\eta\end{pmatrix}$$
    i.e.,
    $$\begin{pmatrix}\theta\\ \omega\end{pmatrix}' = \begin{pmatrix}0 & 1\\ -g/l & -\eta\end{pmatrix}\begin{pmatrix}\theta\\ \omega\end{pmatrix} + O(|\theta|^2 + |\omega|^2)$$

    - Since $\eta > 0$, the eigenvalues have a common negative real part, so the equilibrium is asymptotically stable.
  - At the equilibrium $(\pi, 0)$, we have

$$\begin{pmatrix} \theta \\ \omega \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ g/l & -\eta \end{pmatrix} \begin{pmatrix} \theta - \pi \\ \omega \end{pmatrix} + O(|\theta - \pi|^2 + |\omega|^2)$$

    - For $\eta \geq 0$, there is one positive and one negative eigenvalue, so this equilibrium is unstable.
  - These results should make intuitive sense: If a pendulum is resting at the bottom, that is a stable equilibrium. If a pendulum is resting at the top, that is not a stable equilibrium.

7.3 Applications of the Lyapunov Method

11/11:

- Purely imaginary eigenvalues can still lead to Lyapunov stability.
- **Lyapunov function** (of a system $y' = f(y)$ with fixed point x_0 near x_0): A continuous real function on $\mathbb{R}^n$ such that the following two axioms hold. *Denoted by* $\boldsymbol{L}$.
  1. $L(x_0) = 0$ and $L(x) > 0$ for all $x \in \mathring{B}(x_0, \delta) = B(x_0, \delta) \setminus \{x_0\}$.
  2. $\dot{L}(x) = \nabla L(x) \cdot f(x) \leq 0$ for all $x \in \mathring{B}(x_0, \delta) = B(x_0, \delta) \setminus \{x_0\}$.
- Since

$$\frac{\mathrm{d}}{\mathrm{d}t} L(\phi_t(x)) = \nabla L(\phi_t(x)) \cdot f(\phi_t(x))$$

  the second condition is equivalent to saying that the function L is decreasing along the orbits starting near x_0.
- **Strict** (Lyapunov function): A Lyapunov function for which the decreasing is strict.
- Theorem: For the autonomous system $y' = f(y)$, a fixed point x_0 is
  1. Stable if there is a Lyapunov function near it;

     Proof. Pick a small number $\delta > 0$. Let[6]

$$m := \min\{L(x) : |x - x_0| = \delta\}$$

     Since x_0 does not satisfy $|x - x_0| = \delta > 0$, we know from the first constraint on Lyapunov functions that $L(x) > 0$ for all x satisfying said relation. Thus, $m > 0$. Consequently, any orbit starting from $\{x \mid L(x) < m\} \cap B(x_0, \delta)$ can never meet $\partial B(x_0, \delta)$ since $L(x)$ is decreasing along any orbit (and we would have to go up to get to the boundary). So $L(\phi_t(x)) < m$ for all $x \in \{x \mid L(x) < m\} \cap B(x_0, \delta)$. But this means that $\{x \mid L(x) < m\} \cap B(x_0, \delta)$ is in fact an invariant set. Therefore, x_0 is Lyapunov stable. $\square$
  2. Asymptotically stable if there is a strict Lyapunov function near it.

     Proof. If $x \in \{x \mid L(x) < m\} \cap B(x_0, \delta)$, then $L(\phi_t(x))$ is strictly decreasing. As $t \to +\infty$, $\phi_t(x)$ has a partial limit z_0, say $\phi_{t_k}(x) \to z_0$ (Lemma 6.6 of Teschl (2012)). If $z_0 \neq x_0$, then the orbit $\{\phi_t(z_0) \mid t \in I_{z_0}\}$ is not a single point: Since L is a strict Lyapunov function, we have $L(\phi_t(z_0)) < L(z_0)$ for all $t > 0$. When k is large, $\phi_{t_k}(x)$ is close to z_0, so by continuity,

$$L(\phi_{t+t_k}(x)) = L(\phi_t(\phi_{t_k}(x))) < L(z_0)$$

     But this contradicts $L(\phi_t(x)) > L(z_0)$ (which we must have if there are arbitrarily large t such that $\phi_t(x)$ is close to z_0). Therefore, $x_0 = z_0$. $\square$

[6]Intuitively (in 2D), we take a ring around x_0, find the nonzero value of $L(x)$ at each point on the ring, and take the minimum among them. Imagine a circular valley with hills rising all around the bottommost point; we are essentially looking for the hill that rises the least.

- If all eigenvalues of A have negative real parts, then the perturbed system

$$y' = Ay + g(y)$$

  has a strict Lyapunov function around the fixed point $x = 0$.

  - This observation yields another proof of the stability theorem.

- Advantage of the Lyapunov function: Can be constructed globally and thus gives us global information on the system.

- Examples in studying the global behavior of a phase portrait:

  - Consider a mass point moving along the real axis in a potential field $U(x)$. Then

$$mx'' = -U'(x)$$

    - The total energy
$$E = \frac{m}{2}|x'|^2 + U(x)$$
      is always a constant along any solution.
    - Introducing the velocity allows us to obtain a planar system
$$\begin{pmatrix} x \\ v \end{pmatrix}' = \begin{pmatrix} v \\ -U'(x)/m \end{pmatrix}$$
    - Thus, $E(x, v)$ is a global Lyapunov function.
    - Any fixed point of the system must be of the form $(x_0, 0)$, where $U'(x_0) = 0$.
      - Intuitively, this means that the velocity must be zero (that makes sense) and the position must be such that we are at a critical point of the potential.
    - Because of this, the linearization at a fixed point must be of the following form.
$$\begin{pmatrix} v \\ -U'(x)/m \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -U''(x_0)/m & 0 \end{pmatrix} \begin{pmatrix} x - x_0 \\ v \end{pmatrix} + O(|x - x_0|^2 + |v|^2)$$
    - Thus, $(x_0, 0)$ is Lyapunov stable if U has a nondegenerate local minimum at x_0 and unstable if U has a nondegenerate local maximum at x_0.
      - In the former case, the orbits near $(x_0, 0)$ are closed curves, corresponding to periodic oscillations near x_0 (e.g., harmonic oscillator and ideal pendulum again).

  - Prey-predator model with capacity:
$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} (1 - y - \lambda x)x \\ \alpha(x - 1 - \mu y)y \end{pmatrix}$$
    $\alpha, \lambda, \mu > 0$.
    - x is the number of rabbits and y is the number of wolves.
    - Different ranges of λ induce different global behavior (thus, this is an example of **bifurcation**).
    - General observation: $(x, y) = (0, 0)$ is a saddle point since the linearization there is $\text{diag}(1, -\alpha)$.
    - For $x = 0$ or $y = 0$, the equation is of separable form; the positive x, y-axes are invariant sets.
      - Implication: No orbit in the first quadrant can escape it (compatible with meaning as population).
    - Jacobian:
$$\begin{pmatrix} 1 - y - 2\lambda x & -x \\ \alpha y & \alpha(x - 1) - 2\alpha\mu y \end{pmatrix}$$

- When $\lambda, \mu = 0$, we're back to the Lotka-Volterra system, where there is a single fixed point $(1,1)$.
  - In that case,
    $$(y - \log y - 1) + \alpha(x - \log x - 1)$$
    is a Lyapunov function.
  - However, it is not a strict Lyapunov function since it is constant along any orbit.
  - Moreover, the function is convex, so all level sets are closed curves around the fixed point.
  - This is, indeed, the behavior we observe in Figure 2.1.
- Other cases: $\lambda \geq 1$.
  - There is only one additional fixed point of interest: $(1/\lambda, 0)$. Note that there are other fixed points, but these do not lie in the first quadrant and thus we are not interested.
  - For $\lambda > 1$, the fixed point is stable (a sink) and when $\lambda = 1$, one eigenvalue is 0 since the linearization at that point is $\text{diag}(-1, \alpha(1/\lambda - 1))$.
- $0 < \lambda < 1$.
  - $(1/\lambda, 0)$ becomes a saddle point, and there is a third fixed point
    $$(x_0, y_0) = \left(\frac{1+\mu}{1+\mu\lambda}, \frac{1-\lambda}{1+\mu\lambda}\right)$$
- More on this case in Chapter 7 of Teschl (2012). This is relevant here!

7.4 Chapter 2: Initial Value Problems

From Teschl (2012).

Section 2.6: Extensibility of Solutions

12/6:
- Investigating the maximal interval on which a solution to an IVP can be defined.
- Not really something we covered in class (certainly not from a theoretical point of view).

Section 2.7: Euler's Method and the Peano Theorem

- Mostly review from class; a few interesting points noted below.
- We can derive the Peano proof technique from Taylor's theorem by approximating
  $$\phi(t_0 + h) = x_0 + \dot{\phi}(t_0)h + o(h) = x_0 + f(t_0, x_0)h + o(h)$$
  eliminating the error term, and rearranging.
- **Euler's method**: A method for approximating the solution to an ODE via
  $$x_h(t_{m+1}) = x_h(t_m) + f(t_m, x_h(t_m))h, \quad t_m = t_0 + mh$$
  using linear interpolation in between.
- **Equicontinuous** (family of functions): A family of functions $\{x_m\}$ such that for every $\varepsilon > 0$, there is a $\delta > 0$ such that if $|t - s| < \delta$ and $m \in \mathbb{N}$, then
  $$|x_m(t) - x_m(s)| \leq \varepsilon$$
  - Note that each function in an equicontinuous family of functions is uniformly continuous.

- Theorem 2.18 (Arzelà-Ascoli): Suppose the sequence of functions $x_m(t) \in C(I, \mathbb{R}^n)$, $m \in \mathbb{N}$, on a compact inverval I is equicontinuous. If the sequence x_m is bounded, then there is a uniformly convergent subsequence.
- Theorem 2.19 (Peano): Suppose f is continuous on $V = [t_0, t_0 + T] \times \overline{B_\delta(x_0)} \subset U$ and denote the maximum of $|f|$ by M. Then there exists at least one solution of the initial value problem

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

for $t \in [t_0, t_0 + T_0]$ which remains in $\overline{B_\delta(x_0)}$, where $T_0 = \min(T, \delta/M)$. The analogous result holds for the interval $[t_0 - T_0, t_0]$.
- The Euler algorithm is not the most effective one available today.
  - Variations on it usually take more terms in the Taylor expansion, resulting in an algorithm that converges faster but requires more calculations at each step.
  - A good compromise between more terms (but not too many more terms) is the **Runge-Kutta algorithm**.
  - Even better ones appear in the literature on numerical methods for ODEs.
- **Runge-Kutta algorithm**: An algorithm which approximates $\phi(t_0 + h)$ up to the fourth order in h, setting $t_m = t_0 + hm$ and $x_m = x_n(t_m)$ to yield

$$x_{m+1} = x_m + \frac{h}{6}(k_{1,m} + 2k_{2,m} + 2k_{3,m} + k_{4,m})$$

where

$$k_{1,m} = f(t_m, x_m) \qquad k_{2,m} = f(t_m + \tfrac{h}{2}, x_m + \tfrac{h}{2} \cdot k_{1,m})$$
$$k_{3,m} = f(t_m + \tfrac{h}{2}, x_m + \tfrac{h}{2} \cdot k_{2,m}) \qquad k_{4,m} = f(t_{m+1}, x_m + hk_{3,m})$$

Problems

2.23. Heun's method (or **improved Euler**) is given by

$$x_{m+1} = x_m + \frac{h}{2}(f(t_m, x_m) + f(t_{m+1}, y_m)), \quad y_m = x_m + f(t_m, x_m)h$$

Show that using this method, the error during one step is of $O(h^3)$, provided $f \in C^2$:

$$\phi(t_0 + h) = x_0 + \frac{h}{2}(f(t_0, x_0) + f(t_1, y_0)) + O(h^3)$$

Note that this is not the only possible scheme with this error order since

$$\phi(t_0 + h) = x_0 + \frac{h}{2}(f(t_1, x_0) + f(t_0, y_0)) + O(h^3)$$

as well.

7.5 Chapter 6: Dynamical Systems

From Teschl (2012).

Section 6.1: Dynamical Systems

- Good intuition for what a dynamical system is.
- **Semigroup**: An algebraic structure consisting of a set together with an associative internal binary operation on it.
  - Thus, like a **group**, a semigroup's operation is associative. However, we do not postulate the existence of an identity element or inverses in this case.
- **Dynamical system**: The action of a semigroup G with identity element e[7] on a set M.
  - In particular, a dynamical system is a map $T : G \times M \to M$ which sends $(g, x) \mapsto T_g(x)$ such that
$$T_g \circ T_h = T_{g \circ h} \qquad\qquad T_e = \mathbb{I}$$
  - Intuition: We often think of a dynamical system very similar to a diffeomorphism, in that as we slide t up and down, a set of points gets distorted according to some field. Here, we're taking the formalization of time *acting on* the points to move them around.
  - This is an incredibly minimal/broad/general definition; the dynamical systems we're interested in usually have far more structure.
- **Invertible dynamical system**: A dynamical system for which G is a group.
- **Discrete dynamical system**: A dynamical system for which $G \in \{\mathbb{N}_0, \mathbb{Z}\}$.
- **Continuous dynamical system**: A dynamical system for which $G \in \{\mathbb{R}^+, \mathbb{R}\}$.
- Example: Iterated map, i.e., f^n.
- Example: The flow of an autonomous differential equation, where $T_t = \Phi_t$ and $G = \mathbb{R}$; we consider this example in the next section.

Section 6.2: The Flow of an Autonomous Equation

- Herein, we consider the system
$$\dot{x} = f(x), \quad x(0) = x_0$$
- For the remainder of Teschl (2012), we assume $f \in C^k(M, \mathbb{R}^n)$ $(k \geq 1)$.
  - We also assume M is an open subset of $\mathbb{R}^n$.
- Such a system can be regarded as a **vector field** on $\mathbb{R}^n$.
  - Solutions are curves in M which are tangent to the vector field at each point.
- **Integral curve**: A solution to an autonomous IVP. *Also known as* **trajectory**.
  - We say "ϕ is an integral curve at x_0" if $\phi(0) = x_0$.
- By Theorem 2.13: Every point $x \in M$ has an associated (unique) **maximal integral curve**.
- **Maximal integral curve** (at x): The unique integral curve at x, the domain of which is a **maximal interval**. *Denoted by* $\boldsymbol{\phi_x}$.
- **Maximal interval**: The interval for an integral curve at x containing all other possible intervals on which the integral curve can be defined. *Denoted by* $\boldsymbol{I_x}$, $\boldsymbol{(T_-(x), T_+(x))}$.

[7] So a **monoid**?? A monoid is, by definition, an algebraic structure consisting of a set together with an associative internal binary operation and an identity element.

- We define a set which contains information about the maximal interval of the integral curve at x for all x:

$$W = \bigcup_{x \in M} I_x \times \{x\} \subset \mathbb{R} \times M$$

- **Flow** (of a differential equation): The map from W to M which pairs every starting point x and time t to the point to which the differential equation will have moved x after time t has elapsed. *Denoted by* $\mathbf{\Phi}$. *Given by*

$$(t,x) \mapsto \phi(t,x)$$

where $\phi(t,x)$ is the maximal integral curve at x.

- Notation: We sometimes write

$$\Phi(t,x) = \Phi_x(t) = \Phi_t(x)$$

- If $\phi(\cdot)$ is the maximal integral curve at x, then $\phi(\cdot + s)$ is the maximal integral curve at $y = \phi(x)$ and $I_x = s + I_y$. It follows that for all $x \in M$ and $s \in I_x$, we have

$$\Phi(s+t,x) = \Phi(t,\Phi(s,x))$$

for all $t \in I_{\Phi(s,x)} = I_x - s$.

- We now state formally the ideas we've just developed informally.

- Theorem 6.1: Suppose $f \in C^k(M,\mathbb{R}^n)$. For all $x \in M$, there exists an interval $I_x \subset \mathbb{R}$ containing 0 and a corresponding unique maximal integral curve $\Phi(\cdot,x) \in C^k(I_x,M)$ at x. Moreover, the set W defined as above is open and $\Phi \in C^k(W,M)$ is a (local) flow on M, that is,

$$\begin{gathered}\Phi(0,x) = x\\ \Phi(t+s,x) = \Phi(t,\Phi(s,x)), \quad x \in M, \quad x, t+s \in I_x\end{gathered}$$

Proof. Given. $\square$

- Example: Let $M = \mathbb{R}$ and $f(x) = x^3$. Then $W = \{(t,x) \mid 2tx^2 < 1\}$[8] and

$$\Phi(t,x) = \frac{x}{\sqrt{1-2x^2t}}$$

We have $T_-(x) = -\infty$ and $T_+(x) = 1/(2x^2)$.

- **Fixed point**: A point at which f evaluates to 0. *Denoted by* $\boldsymbol{x_0}$.

- Lemma 6.2 (Straightening out vector fields): Suppose $f(x_0) \neq 0$. Then there is a local coordinate transform $y = \varphi(x)$ such that $\dot{x} = f(x)$ is transformed to

$$\dot{y} = (1,0,\dots,0)$$

Section 6.3: Orbits and Invariant Sets

- Some of the definitions herein come up in class, some do not, but many are interesting and IMO grant a deeper understanding of dynamical systems.

- **Orbit** (of x): The image under the flow of the maximal interval of the maximal integral curve at x. *Denoted by* $\boldsymbol{\gamma(x)}$. *Given by*

$$\gamma(x) = \Phi(I_x \times \{x\})$$

- $y \in \gamma(x)$ implies $y = \Phi(t,x)$ for some t, and hence (by Theorem 6.1) $\gamma(x) = \gamma(y)$.

[8] This condition is equivalent to all (t,x) such that $1 - 2x^2t > 0$, i.e., that the denominator of the flow is positive.

- Implication: Distinct orbits are disjoint.
  - Formalism: The orbits partition M, i.e., we have an equivalence relation on M defined by $x \sim y$ iff $\gamma(x) = \gamma(y)$.
- **Fixed point** (of Φ): A point $x \in M$ for which $\gamma(x) = \{x\}$. *Also known as* **singular point**, **stationary point**, **equilibrium point**.
- **Regular point** (of Φ): A point $x \in M$ that is not a fixed point of Φ.
- If x is a regular point, then $\Phi(\cdot, x) : I_x \hookrightarrow M$[9].
- **Forward** (orbit of x): The image under the flow of the *positive portion* of the maximal interval of the maximal integral curve at x. *Denoted by* $\boldsymbol{\gamma_+(x)}$. *Given by*

$$\gamma_+(x) = \Phi((0, T_+(x)) \times \{x\})$$

- **Backward** (orbit of x): The image under the flow of the *negative portion* of the maximal interval of the maximal integral curve at x. *Denoted by* $\boldsymbol{\gamma_-(x)}$. *Given by*

$$\gamma_-(x) = \Phi((T_-(x), 0) \times \{x\})$$

- Relating the orbit, forward orbit, and backward orbit:

$$\gamma(x) = \gamma_-(x) \cup \{x\} \cup \gamma_+(x)$$

- **Periodic point** (of Φ): A point $x \in M$ for which there exists $T > 0$ such that $\Phi(T, x) = x$.
- **Period** (of a periodic point x): The lower bound on the set of T for which $\Phi(T, x) = x$. *Denoted by* $\boldsymbol{T(x)}$. *Given by*

$$T(x) = \inf\{T > 0 \mid \Phi(T, x) = x\}$$

- The continuity of Φ guarantees that

$$\Phi(T(x), x) = x$$

  for $T(x)$ as defined.
- By the flow property (Theorem 6.1), we have

$$\Phi(t, T(x), x) = \Phi(t, x)$$

- **Periodic orbit**: An orbit for which one point (hence all points) of the orbit is/are periodic. *Also known as* **closed orbit**.
  - Reason for the moniker "closed orbit:" x is periodic iff $\gamma_+(x) \cap \gamma_-(x) \neq \emptyset$, i.e., if the forward orbit joins the negative orbit and "closes" the loop.
- Classification of the orbits of f:
  1. Fixed orbits (corresponding to a periodic point with period zero).
  2. Regular periodic orbits (corresponding to a periodic point with positive period).
  3. Non-closed orbits (not corresponding to a periodic point).
- **Positive lifetime** (of x): The positive ending limit point of the maximal interval of x. *Denoted by* $\boldsymbol{T_+(x)}$. *Given by*

$$T_+(x) = \sup I_x$$

[9]Notation: $\hookrightarrow$ indicates an injective function.

- **Negative lifetime** (of x): The negative ending limit point of the maximal interval of x. *Denoted by* $\boldsymbol{T_-(x)}$. *Given by*
$$T_-(x) = \inf I_x$$
- $\boldsymbol{\sigma}$ **complete** (point): A point $x \in M$ for which $T_\sigma(x) = \sigma\infty$, where $\sigma \in \{\pm\}$.
- **Complete** (point): A point $x \in M$ that is both $+$ and $-$ complete.
- Lemma 6.3: Let $x \in M$ and suppose that the forward (resp. backward) orbit lies in a compact subset C of M. Then x is $+$ (resp. $-$) complete.
- Periodic points are complete.
- **Complete** (vector field): A vector field in which all points are complete.
- f complete implies Φ is globally defined, that is, $W = \mathbb{R} \times M$.
- $\boldsymbol{\sigma}$ **invariant**: A set $U \subset M$ such that $\gamma_\sigma(x) \subset U$ for all $x \in U$, where $\sigma \in \{\pm\}$.
- $C \subset M$ a compact σ invariant set implies (by Lemma 6.3) that all points in C are σ complete.
- Lemma 6.4:
  1. Arbitrary intersections and unions of σ invariant sets are σ invariant. Moreover, the closure of a σ invariant set is again σ invariant.
  2. If U, V are invariant, so is the complement $U \setminus V$.

  Proof. Given. □

- Goal: Describe the long-term asymptotics of solutions.
  - Tool: We introduce the set where an orbit eventually accumulates.
- $\boldsymbol{\omega_\pm}$**-limit set** (of x): The set of all points $y \in M$ for which there exists a sequence $\{t_n\}$ that converges to $\pm\infty$ and satisfies $\Phi(t_n, x) \to y$. *Denoted by* $\boldsymbol{\omega_\pm(x)}$.
- By definition, $\omega_\pm(x)$ is empty unless x is $\pm$ complete.
- $y \in \gamma(x)$ implies $\omega_\pm(x) = \omega_\pm(y)$.
  - This is because the hypothesis shows that $y = \Phi(t, x)$ for some t, so
$$\Phi(t_n, x) = \Phi(t_n - t, \Phi(t, x)) = \Phi(t_n - t, y)$$
  - Hence, $\omega_\pm(x)$ depends only on the orbit $\gamma(x)$.
- Lemma 6.5: The set $\omega_\pm(x)$ is a closed invariant set.

  Proof. Given. □

- Example: For $\dot{x} = -x$, $\omega_+(x) = \{0\}$ for all $x \in \mathbb{R}$ since every solution converges to 0 as $t \to +\infty$. Moreover, $\omega_-(x) = \emptyset$ for $x \neq 0$ and $\omega_-(0) = \{0\}$.
- Conclusion: Even for x complete, the set $\omega_\pm(x)$ can be empty.
- Lemma 6.6: If $\gamma_\sigma(x)$ is contained in a compact set C, then $\omega_\sigma(x)$ is nonempty, compact, and connected.

  Proof. Given. □

- Lemma 6.7: Suppose $\gamma_\sigma(x)$ is contained in a compact set. Then we have

$$\lim_{t\to\sigma\infty} d(\phi(t,x),\omega_\sigma(x)) = 0$$

Proof. Given. □

- Teschl (2012) works through an example that proves that the compactness requirement is necessary.
- **Minimal** (set): A nonempty, compact, σ invariant set that contains no proper σ invariant subset possessing these three properties.
- Examples:
  - The $\omega_\pm$-limit sets are minimal for all $x \in \omega_\pm(x)$.
  - A periodic orbit.
    - In 2D, this is the only example by Corollary 7.12.
    - In three or more dimensions, orbits can be dense on a compact hypersurface, meaning that the hypersurface cnnot be split into smaller *closed* invariant sets.
- Lemma 6.8: Every nonempty, compact σ invariant set $C \subset M$ contains a minimal σ invariant set.

  If in addition C is homeomorphic to a closed m-dimensional disk (where m is not necessarily the dimension of M), it contains a fixed point.

Proof. Given. □

Section 6.4: The Poincaré Map

- Never covered in class.

Section 6.5: Stability of Fixed Points

- Herein, we continue investigating the long-term behavior of the dynamical system

$$\dot{x} = f(x), \quad x(0) = x_0$$

  - In particular, we investigate whether or not a solution is **stable**.
- **Stable** (fixed point): A fixed point x_0 of $f(x)$ such that for any given neighborhood $U(x_0)$, there exists another neighborhood $V(x_0) \subset U(x_0)$ such that any solution starting in $V(x_0)$ remains in $U(x_0)$ for all $t \geq 0$. *Also known as* **Lyapunov stable**[10].
  - If a solution remains in $U(x_0)$ for all $t \geq 0$, it remains in the compact set $\overline{U(x_0)}$ for all $t \geq 0$.
  - Thus, by Lemma 6.3, said solution exists for all positive times.
- **Unstable** (fixed point): A fixed point which is not stable.
- **Asymptotically stable** (fixed point): A fixed point x_0 of $f(x)$ that is stable and for which there exists a neighborhood $U(x_0)$ such that

$$\lim_{t\to\infty} |\phi(t,x) - x_0| = 0$$

  for all $x \in U(x_0)$.

[10]Teschl (2012) uses the alternate spelling "Liapunov;" I will continue using "Lyapunov" without further comment.

- **Exponentially stable** (fixed point): A fixed point x_0 of $f(x)$ for which there exist constants $\alpha, \delta, C > 0$ such that
$$|\phi(t,x) - x_0| \leq C\mathrm{e}^{-\alpha t}|x - x_0|$$
when $|x - x_0| \leq \delta$.
- Exponential stability implies both stability and asymptotic stability.
- Example: Consider $\dot{x} = ax$ in $\mathbb{R}^1$. Then $x_0 = 0$ is stable iff $a \leq 0$ and exponentially stable iff $a < 0$.
- These definitions of stability agree with those we introduced for linear autonomous systems in Section 3.2.
- Teschl (2012) goes over an alternate stability criterion adapted from Section 1.5.
- If $f'(x_0) \neq 0$, the stability of x_0 can be read off from the derivative of f at x_0 alone.
  - More generally, a fixed point is exponentially stable if this is true for the corresponding linearized system (the proof is not directly presented in Teschl (2012) but is rather spread out, making it not of much use to me rn).
- Theorem 6.10 (Exponential stability via linearization): Suppose $f \in C^1$ has a fixed point x_0 and suppose that all eigenvalues of the Jacobian matrix at x_0 have negative real part. Then x_0 is exponentially stable.
- **Bifurcation theory**: The systematic study of small changes in an ODEs parameters that induce large changes in qualitative behavior.
  - Theorem 2.11 asserts that provided f depends smoothly on μ, so does the flow. Nevertheless, very small changes in parameters can induce large changes in the qualitative behavior.
  - A few examples follow.
- **Pitchfork bifurcation**: A stable fixed point for $\mu \leq 0$ which becomes unstable and splits off two stable fixed points at $\mu = 0$. *picture*
  - Example: $\dot{x} = \mu x - x^3$.
- **Transcritical bifurcation**: Two stable fixed points for $\mu \neq 0$ which collide and exchange stability at $\mu = 0$. *picture*
  - Example: $\dot{x} = \mu x - x^2$.
- **Saddle-node bifurcation**: One stable and one unstable fixed point for $\mu < 0$ which collide at $\mu = 0$ and vanish. *picture*
  - Example: $\dot{x} = \mu + x^2$.
- Rest of the chapter: Good criteria for the stability of $\dot{x} = f(x)$ (since it cannot be solved explicitly in general).

Section 6.6: Stability via Lyapunov's Method

- For a fixed point x_0 of f and an open neighborhood $U(x_0)$ of x_0, we may define the following.
- **Lyapunov function**: A continuous function $L : U(x_0) \to \mathbb{R}$ which is zero at x_0, positive for $x \neq x_0$, and satisfies
$$L(\phi(t_0)) \geq L(\phi(t_1))$$
where $t_0 < t_1$ and $\phi(t_j) \in U(x_0) \setminus \{x_0\}$ $(j = 0, 1)$ for any solution $\phi(t)$.
- **Strict Lyapunov function**: A Lyapunov function for which the central inequality in the above definition is strict.

- Claim: If L is strict, $U(x_0) \setminus \{x_0\}$ cannot contain any periodic orbits.

  Proof. Suppose for the sake of contradiction that $\gamma(x) \subset U(x_0) \setminus \{x_0\}$. Since $\gamma(x)$ is a periodic orbit, $\phi(0,x) = \phi(T(x),x)$ where $T(x) > 0$ by definition. Letting $t_0 = 0$ and $t_1 = T(x)$, we have by the definition of a strict Lyapunov function that

  $$L(\phi(0,x)) > L(\phi(T(x),x)) = L(\phi(0,x))$$

  contradicting the fact that L is well-defined. □

- ***S_δ***: The following set. *Given by*

  $$S_\delta = \{x \in U(x_0) \mid L(x) \leq \delta\}$$

  - S_δ contains x_0.
  - In general, S_δ need not be closed since it can share boundary with $U(x_0)$. In such a case, orbits can escape through this part of the boundary.
  - Restricting S_δ to closed versions, though, we get the following lemma.

- Lemma 6.11: If S_δ is closed, then it is positively invariant.

- Lemma 6.12: For every $\delta > 0$, there is an $\varepsilon > 0$ such that $S_\varepsilon \subset B_\delta(x_0)$ and $B_\varepsilon(x_0) \subset S_\delta$.

- Implication: Given any neighborhood $V(x_0)$, we can find an ε such that $S_\varepsilon \subset V(x_0)$ is positively invariant. But this just means that x_0 is stable, and we have proven the following[11].

- Theorem 6.13 (Lyapunov): Suppose x_0 is a fixed point of f. If there is a Lyapunov function L, then x_0 is stable.

- Theorem 6.14 (Krasovskii-LaSalle principle): Suppose x_0 is a fixed point of f. If there is a Lyapunov function L which is not constant on any orbit lying entirely in $U(x_0) \setminus \{x_0\}$, then x_0 is asymptotically stable. This is for example the case if L is a strict Lyapunov function. Moreover, every orbit lying entirely in $U(x_0)$ converges to x_0.

- Theorem 6.15: Let $L : U \to \mathbb{R}$ be continuous and bounded from below. If for some x we have $\gamma_+(x) \subset U$ and

  $$L(\phi(t_0,x)) \geq L(\phi(t_1,x))$$

  for $t_0 < t_1$, then L is constant on $\omega_+(x) \cap U$.

- Most Lypunov functions are differentiable.

- If L is differentiable, then $L(\phi(t_0)) \geq L(\phi(t_1))$ for all $t_0 < t_1$ iff

  $$\frac{\mathrm{d}}{\mathrm{d}t} L(\phi(t,x)) = \nabla(L)(\phi(t,x))\dot{\phi}(t,x) = \nabla(L)(\phi(t,x))f(\phi(t,x)) \leq 0$$

- **Lie derivative** (of L along f): The following expression. *Given by*

  $$\nabla(L)(x) \cdot f(x)$$

- **Constant of motion**: A function for which the Lie derivative vanishes and, hence, is constant on every orbit.

- Theorem 6.15 implies that all $\omega_\pm$-limit sets are contained in the set where the Lie derivative of L vanishes.

[11] Clever pedagogical tool: Teschl (2012) weaves any really important proofs into the flow of the text so that you can't gloss over it.

- Example: Consider the system
$$\dot{x} = y \qquad\qquad \dot{y} = -x$$
with function
$$L(x,y) = x^2 + y^2$$
  - For $x \in \mathbb{R}^2$ arbitrary, the Lie derivative is
$$\nabla(L)(x) \cdot f(x) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \cdot \begin{pmatrix} y \\ -x \end{pmatrix} = 2xy - 2xy = 0$$
  - Thus, L is a Lyapunov function.
  - In particular, L is a constant of motion.
  - Thus, by Theorem 6.13, the origin is stable.
  - Every level set $L(x,y) = \delta$ corresponds to an orbit, so the system is not asymptotically stable.
- Takeaway:
  - Extract properties of Lyapunov functions from the fact that they are monotonically decreasing on all orbits.
  - Prove that a function is a Lyapunov function using the Lie derivative.

Section 6.7: Newton's Equation in One Dimension

- Goal: Illustrate the results of Chapter 6 with a specific example.
- Recall that the motion of a particle moving in one dimension under the external force field $f(x)$ is described by Newton's equation
$$\ddot{x} = f(x)$$
- **Phase space**: The set $\mathbb{R}^2$, as referred to by physicists.
- **Phase point**: A point of the form $(x, \dot{x})$ in the phase space.
- **Phase curve**: A solution to the ODE.
- By the Picard-Lindelöf theorem (Theorem 2.2), precisely one phase curve passes through each phase point.
- **Kinetic energy**: The following quadratic form. *Denoted by* $\boldsymbol{T(\dot{x})}$. *Given by*
$$T(\dot{x}) = \frac{\dot{x}^2}{2}$$
- **Potential energy**: The following function. *Denoted by* $\boldsymbol{U(x)}$. *Given by*
$$U(x) = -\int_{x_0}^{x} f(\xi)\mathrm{d}\xi$$
  - Only determined up to an arbitrary constant.
- **Energy** (of a Newtonian system): The sum of the kinetic and potential energies. *Denoted by* $\boldsymbol{E}$. *Given by*
$$E = T(\dot{x}) + U(x)$$

- E is constant along a phase curve. Indeed, if $x(t)$ satisfies $\ddot{x} = f(x) = U'(x)$, i.e., $\ddot{x} - f(x) = 0$, then
$$\frac{\mathrm{d}E}{\mathrm{d}t} = \dot{x}\ddot{x} + U'(x)\dot{x} = \dot{x}(\ddot{x} - f(x)) = 0$$
as desired.

- The solution corresponding to the initial conditions $x(0) = x_0$, $\dot{x}(0) = x_1$ can be given implicitly.
  - First off, we have that
$$\begin{aligned}
E &= T(\dot{x}) + U(x)\\
&= \frac{\dot{x}^2}{2} + U(x)\\
\sqrt{2(E - U(x))} &= \frac{\mathrm{d}x}{\mathrm{d}t}\\
\int_0^t \mathrm{d}\tau &= \int_{x_0}^x \frac{\mathrm{d}\xi}{\sqrt{2(E - U(\xi))}}\\
t &= \operatorname{sign}(x_1)\int_{x_0}^x \frac{\mathrm{d}\xi}{\sqrt{2(E - U(\xi))}}
\end{aligned}$$
    - Why do we input $\operatorname{sign}(x_1)$ between the next-to-last and last lines??
  - Second, since E is constant along the solution, its value at the start will be the same as its value at any other point. Thus, we can use the starting initial conditions to calculate it, as follows.
$$E = \frac{{x_1}^2}{2} + U(x_0)$$
  - If $x_1 = 0$, then $\operatorname{sign}(x_1)$ must be replaced with $-\operatorname{sign}(U'(x_0))$.
- Theorem 6.16: Newton's equation has a fixed point if and only if $\dot{x} = 0$ and $U'(x) = 0$ at this point. Moreover, a fixed point is stable if $U(x)$ has a local minimum there.
  - If $U(x)$ has a local minimum at x_0, the energy $E - U(x_0)$ can be used as a Lyapunov function; we subtract $U(x_0)$ so that $E - U(x_0)$ evaluates to zero at x_0.
- A fixed point cannot be asymptotically stable (due to conservation of energy).
- Example: Mathematical pendulum.
$$\ddot{x} = -\sin(x)$$
  - x describes the displacement angle from the position at rest ($x = 0$).
  - x should be understood modulo 2π.
  - We have that
$$U(x) = -\int_{x_0}^x (-\sin(\xi))\mathrm{d}\xi = \cos(x_0) - \cos(x)$$
    - Since the constant is arbitrary, we may take $U(x) = 1 - \cos(x)$ for ease. This has the additional advantage that the energy is never negative.
  - We now begin the rigorous investigation.
  - We restrict our attention to the interval $x \in (-\pi, \pi]$. Thus, the fixed points are $x = 0, \pi$.
  - By Theorem 6.16 and the fact that $U(0)$ is a minimum, 0 is a stable fixed point.
  - As before in the general case, $E(x, \dot{x}) =$ constant gives invariant level sets.
    - $E = 0$: The corresponding level set is the equilibrium position $(x, \dot{x}) = (0, 0)$.
    - $0 < E < 2$: The level sets are homeomorphic to circles. Since these circles contains no fixed points, they are regular periodic orbits.

- $E = 2$: The level set consists of the fixed point π and two non-closed orbits connecting $-\pi$ and π. This is a **separatrix**.
- $E > 2$: The level sets are again closed orbits (due to our modulo 2π perspective).

– In a neighborhood of the equilibrium position $x = 0$, the system is approximated by the linearization $\sin(x) = x + O(x^2)$ given by

$$\ddot{x} = -x$$

and referred to as the **harmonic oscillator**.

- Here, we have

$$E = \frac{\dot{x}^2}{2} + \frac{x^2}{2}$$

so the phase portrait consists of circle centered at 0.

– More generally, if $U'(x_0) = 0$ and $U''(x_0) = \omega^2/2 > 0$, then we should approximate our system with

$$\ddot{y} = -\omega^2 y, \quad y(t) = x(t) - x_0$$

– Lastly, if we use the momentum $p = \dot{x}$ (units chosen such that $m = 1$) and the location $q = x$ as coordinates, then the energy

$$H(p, q) = \frac{p^2}{2} + U(q)$$

is called the **Hamiltonian**.

- In this case, the equations of motion are

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \dot{p} = -\frac{\partial H}{\partial q}$$

- This formalism is called **Hamiltonian mechanics**.
- It is useful for systems with more than one degree of freedom.
- See Section 8.3 for more.

7.6 Chapter 7: Planar Dynamical Systems

From Teschl (2012).

Section 7.1: Examples from Ecology

- Teschl (2012) derives via ecological reasoning the **Lotka-Volterra** predator-prey equations.
- **Lotka-Volterra predator-prey equations**: The following system of differential equations. *Given by*

$$\dot{x} = (1 - y)x \qquad \dot{y} = \alpha(x - 1)y$$

for $\alpha > 0$.

- Two fixed points.
  - $(0, 0)$ gives rise to invariant subspaces along the x- and y-axes. Indeed,

$$\Phi(t, (0, y)) = (0, y\mathrm{e}^{-\alpha t}) \qquad \Phi(t, (x, 0)) = (x\mathrm{e}^{t}, 0)$$

    - Since no other solution can cross these lines, the first quadrant $Q = \{(x, y) \mid x, y > 0\}$ is invariant. This is the region we are interested in.
  - $(1, 1)$ is the other fixed point.

  - Let's eliminate t from the ODEs to get a single first-order equation for the orbits.
  - Writing $y = y(x)$, we infer from the chain rule that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^{-1} = \alpha\frac{(x-1)y}{(1-y)x}$$

  - This equation is separable. Solving it yields

$$L(x,y) = f(y) + \alpha f(x) = \text{constant}$$

    where

$$f(a) = a - 1 - \log(a)$$

    - ➢ Note that log denotes the natural logarithm.
  - f cannot be inverted in terms of elementary functions. However, f is convex with global minimum at $x = 1$, and $f \to \infty$ as $a \to 0, +\infty$. It follows that the level sets are portions of this curve near the bottom of the well in both dimensions, and thus they are compact.
  - The exchange of energy from one to the other and back again also indicates that each orbit is periodic surrounding the fixed point $(1, 1)$.

- Theorem 7.1: All orbits of the Lotka-Volterra equations in Q are closed and encircle the only fixed point $(1, 1)$.
- Modification: Let's assume each species' population can only grow so fast. Then we get

$$\dot{x} = (1 - y - \lambda x)x \qquad \dot{y} = \alpha(x - 1 - \mu y)y$$

  for $\alpha, \lambda, \mu > 0$.
- We now have four fixed points:

$$(0,0) \qquad (\lambda^{-1}, 0) \qquad (0, -\mu^{-1}) \qquad \left(\frac{1+\mu}{1+\mu\lambda}, \frac{1-\lambda}{1+\mu\lambda}\right)$$

  - The third lies outside of $\bar{Q}$, so we disregard it.
  - The fourth lies outside of $\bar{Q}$ if $\lambda > 1$. Thus, let's start with the case $\lambda \geq 1$ so that we only have to deal with one new fixed point.
  - $\lambda \geq 1$. *picture*
    - Our new fixed point is $(\lambda^{-1}, 0)$.
    - It is a hyperbolic sink if $\lambda > 1$.
    - If $\lambda = 1$, one eigenvalue is 0 and we need a more thorough investigation.
    - Idea: Split Q into regions where $\dot{x}, \dot{y}$ have definite signs and then use the elementary observation in Lemma 7.2.
    - The regions where $\dot{x}, \dot{y}$ have definite signs are separated by the two lines

$$L_1 = \{(x,y) \mid y = 1 - \lambda x\} \qquad L_2 = \{(x,y) \mid \mu y = x - 1\}$$

      - ➢ We derive these by setting $1 - y - \lambda x = 0$ and $x - 1 - \mu y = 0$.
    - Label the regions in Q enclosed by these lines from left to right by Q_1, Q_2, Q_3.
    - Observe that the lines are transversal, i.e., can only be crossed in the direction from $Q_3 \to Q_2$ and $Q_2 \to Q_1$. This can be seen from the solution curves in the picture.
    - Suppose we start at $(x_0, y_0) \in Q_3$.
      - ➢ Additional constraint: $x \leq x_0$ (the flow is to the left??).
      - ➢ By Lemma 7.2: Either the trajectory enters L_2 or it converges to a fixed point in $\bar{Q}_3$. The latter case can only happen if $(\lambda^{-1}, 0) \in \bar{Q}_3$, i.e., if $\lambda = 1$.

- Similarly, starting in Q_2 either gets you across L_1 or to $(\lambda^{-1}, 0)$.
- Starting in Q_1 must take you to the fixed point.
- Thus, every trajectory converges to the fixed point.

– Let $0 < \lambda < 1$.
  - We apply the same strategy as before.
  - We have four regions this time. Let Q_4 be the new (bottom) one. We can only pass through these in the order $Q_4 \to Q_3 \to Q_2 \to Q_1 \to Q_4$.
  - Thus, we have to rule out the periodic case this time.
  - For simplicity's sake, let
$$(x_0, y_0) = \left(\frac{1+\mu}{1+\mu\lambda}, \frac{1-\lambda}{1+\mu\lambda}\right)$$
  - To do so, introduce (inspired by the original) the Lyapunov function
$$L(x, y) = \gamma_1 f(\tfrac{y}{y_0}) + \alpha\gamma_2 f(\tfrac{x}{x_0})$$
where, as before, $f(a) = a - 1 - \log(a)$.
  - We seek constraints on γ_1, γ_2 that will make L strict.
  - Calculate
$$\dot{L} = \frac{\partial L}{\partial x}\dot{x} + \frac{\partial L}{\partial y}\dot{y} = -\alpha\left(\frac{\lambda\gamma_2}{x_0}\bar{x}^2 + \frac{\mu\gamma_1}{y_0}\bar{y}^2 + \left(\frac{\gamma_2}{x_0} - \frac{\gamma_1}{y_0}\right)\bar{x}\bar{y}\right)$$
where
$$\dot{x} = (-\bar{y} - \lambda\bar{x})x \qquad \dot{y} = \alpha(\bar{x} - \mu\bar{y})y \qquad \bar{x} = x - x_0 \qquad \bar{y} = y - y_0$$
  - The the RHS will be negative if we choose $\gamma_1 = y_0$ and $\gamma_2 = x_0$, so choose this, and then L is strictly decreasing, so all orbits starting in Q converge to the fixed point (x_0, y_0).

- Lemma 7.2: Let $\phi(t) = (x(t), y(t))$ be the solution of a planar system. Suppose U is open and $\bar{U}$ is compact. If $x(t), y(t)$ are strictly monotone in U, then either ϕ hits the boundary at some finite time $t = t_0$ or $\phi(t)$ converges to a fixed point $(x_0, y_0) \in \bar{U}$.

- Therefore, after all of that, we have proven the following.

- Theorem 7.3: Suppose $\gamma \geq 1$. Then there is no fixed point of
$$\dot{x} = (1 - y - \lambda x)x \qquad \dot{y} = \alpha(x - 1 - \mu y)y$$
in Q and all trajectories in Q converge to the point $(\lambda^{-1}, 0)$.

  If $0 < \lambda < 1$, then there is only one fixed point $(\frac{1+\mu}{1+\mu\lambda}, \frac{1-\lambda}{1+\mu\lambda})$ in Q. It is asymptotically stable and all trajectories in Q converge to this point.

- Ecological interpretation: Predators can only survive if their growth rate is positive at the limiting population λ^{-1} of the prey species.

- Teschl (2012) discusses cooperative and competing species.

Week 8

Stability Grab Bag

8.1 Midterm 2 Review

11/14:

- Still 3 problems total and 5 points each.
  - The problems will be calculations based on the basic concepts.
  - Figure out the stable and unstable subspaces of some finite systems.
  - Figure out whether or not a system is stable.
  - Prove whether or not a function is planar linear
- Starting with the classification of planar linear autonomous systems.
  - We have $y' = Ay$ where A is a 2×2 real matrix.
  - As a result of the realness, the eigenvalues behave regularly, i.e., there are only finitely many types of eigenvalues. These are...
    1. Real, nonzero, same sign. Depending on the sign, we'll either have a source or a sink. The orbits will be a distorted graph of a power function. If asked to investigate the phase portrait, then we need to figure out the stable and unstable subspaces and clearly indicate a basis. If asked to draw, we need to clearly indicate which subspaces are stable and unstable. We also need to clearly indicate the direction of the phase lines. First case: Everything is stable; second case: Everything is unstable. We draw the eigenspaces as well with arrows on the "axes." Figure 5.3a-5.3b.
    2. Real, different sign. One stable and one unstable subspace. We need to clearly indicate how the axes are tilted. Figure 5.3c.
    3. A is similar to the Jordan block with zero eigenvalues and 1 in the upper right hand corner. Then
    $$A \sim \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathrm{e}^{tA} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$
    4. Purely imaginary eigenvalues. These must appear in a conjugate pair. The phase diagram will be concentric ellipses, and we essentially have the harmonic oscillator equation. If we have to sketch, we must show how the ellipses are tilted.
    5. Complex eigenvalues $\sigma \pm i\beta$. Either we have a spiral source or a spiral sink. It's meaningless to indicate how the spiral tilts here, so don't bother trying. Determining whether they spin clockwise or counterclockwise. If
    $$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

then our fundamental solution is

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and we rotate counterclockwise. Since $A^2 = -\mu^2 I_2$, $\mathrm{e}^{tA} = I_2 \cos \mu t + \mu^2 I_2 \sin t$. Negative reverses everything. Harmonic oscillator goes counterclockwise.

- There is an online website that gives us phase portraits for an equation. We can use this to help develop intuition.
- If you have a set of eigenvectors, how do you know how to tilt it?
  - Shao goes over examples of eigenvalues and eigenvectors.
- This is not something you need to memorize, but something you need to be able to recover.

- This is not a course for math majors; thus, there will not be proofs concerning the contraction mapping principle. We will not be asked to show existence, uniqueness, continuous difference, or differentiability with respect to parameters.
- We do need to know Grönwall's inequality, however.
- Grönwall's inequality: If $\phi : [p, T] \to \mathbb{R}$ and

$$\phi(t) \leq b + a \int_0^t \phi(\tau) \mathrm{d}\tau$$

then

$$\phi(t) \leq b\mathrm{e}^{at}$$

  - Usually stated in the integral form, and we usually only need a special case.
  - We may need to prove this; the proof mimics the derivation of the Duhamel formula.
  - $a, b \in \mathbb{R}$.
  - We need to memorize the proof.
  - We also need to be able to recognize when we can and should use it. Let $\phi(t) = \Phi'(t) \leq b + a\Phi(t)$, $\Phi(0) = 0$. Then $\phi(t) \leq b + a \int_0^t \varphi(\tau) \mathrm{d}\tau$.
  - Use it when we want to bound a function that satisfies either an integral or a differential quantity.
  - This is the only proof in the theory of ODE systems we need to memorize.
- We need to master the methods to compute perturbation series.
  - Suppose our IVP depends on a parameter μ differentiably.

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t; y(t; \mu); \mu), y(t_0) = x(\mu), \mu \approx 0$$

  - If the parameter is close to zer, then you should be able to compute the μ-derivative with respect to the parameter.
  - By Taylor expanding with respect to the parameter, you should be able to recover solutions that are close to the actual.

$$y(t; \mu) = y_0(t) + y_1(t)\mu + y_2(t)\mu^2 + O(\mu^3)$$

  - We are typically satisfied with approximations to the second order.
  - We expand our ODE into a Taylor series of μ. The differentiability with respect to parameters theorem (see Lecture 6.2 or Theorem 2.11 in Teschl (2012)) tells us that this is legitimate.

$$\frac{\mathrm{d}}{\mathrm{d}t}(y_0(t)) = f(t; y_0(t); 0), y_0(t_0) = x(0)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(y_1(t)) = \frac{\partial f}{\partial z} y_1(t) + \frac{\partial f}{\partial \mu}, y_1(t) = \frac{\partial x}{\partial \mu}$$

- Just know the basic Taylor expansions (trig ones and exponential functions; usually we'll stick to polynomials, though).
- Use the ansatz $y(t;\mu) = y_0(t) + y_1(t)\mu + y_2(t)\mu^2 + O(\mu^3)$.
- Substitute $y(t;\mu)$ into $f(t, y(t;\mu);\mu)$. Expand $f(t;y(t;\mu);\mu)$ into a Taylor series of μ. Balance the coefficients of $\mu^0, \mu^1, \mu^2, \dots$.
- Then you will get a series of equations that is theoretically solvable. Then a sequence of ODEs for $y_0(t), y_1(t), y_2(t), \dots$.
- Your ODEs for $y_1, y_2, \dots$ should not involve μ (because they are coefficients in the Taylor expansion with respect to μ. Coefficients of a Taylor series shouldn't involve the argument); if it does, there is something going wrong.
- As for the initial value, $y_0(t_0) + y_1(t_0)\mu + y_2(t_0)\mu^2 + \cdots$. This implies that something equals $x(\mu)$. The Taylor coefficients of $x(\mu)$ at $\mu = 0$.
- These are the general steps you use to find the perturbative series expansion.
- The computations on the exam will not be too heavy.
- If you're still unclear on the calculation, look through the HW answer keys.

- Conclusion: The Grönwall's inequality is something we need to remember from the theory; the perturbative procedure is something we need to be able to do.
- Why do we expand with respect to μ?
  - We do it with respect to μ because our function is a function of μ. Differentiability and smallness imply we can use the Taylor series.
- Shao reiterates: Definitely read through the key to HW5!!! All the steps you will need to do are done completely and in detail.
- There will be things that are in HW6 (the one due Friday) that will appear on the exam because we have discussed these things in lecture.
- The definitions of Lyapunov stability and asymptotic stability. These will appear in the exam. We need to *clearly* remember the definitions.
- Consider $y' = f(y)$, $f(x_0) = 0$ (an autonomous system with a fixed point; we can transform our system via $(y - x_0)' = f(x_0 + (y - x_0))$ to translate our fixed point to zero; implies $y' = f(x_0 + y)$, $y = 0$ is a fixed point). We should be able to determine the asymptotic stability near x_0 by computing the linearization (i.e., the Jacobian $f'(x_0)$) at the fixed point.
  - Regarding determining stability near x_0, remember the following theorem.
  - Theorem: If all eigenvalues of $f'(x_0)$ have negative real parts, then x_0 is asymptotically stable. If at least one eigenvalue has real part greater than zero, then x_0 is not Lyapunov stable.
  - We should be able to apply the above criterion in practice.
  - We should also be able to reproduce the proof of the first part of Lyapunov's theorem (related to a question in HW6).
  - Lyapunov functions: $f(x_0) = 0$. Definition:
    1. $L(x)$ is C^1 near x_0, $L(x_0) = 0$. $L(x) > 0$ for x near x_0.
    2. $\nabla L(x) \cdot f(x) \leq 0$ for x near x_0 iff $L(\phi_t(x)) \leq L(x)$, $t \geq 0$. If $L(\phi_t(x))$ is always strictly decreasing, then it is a strict Lyapunov function.
  - Theorem (Lyapunov's theorem): Usually, we can explicitly determine a Lyapunov function:
    1. If there is a Lyapunov function near the fixed point, then it is Lyapunov stable. For trajectories starting at nearby points, the trajectory can never excape nearby points.
    2. If there is a strict Lyapunov function, then it is asymptotically stable.

    - We need to be able to apply this theorem in practice; we don't need to know the proof.

- Examples of Lyapunov functions: Newton's second law.
  - Suppose you have a particle moving within a potential field with potential function U, i.e.,

$$mx'' = -U'(x)$$

  - Then by a standard process, you can convert it to a planar linear system by introducing the variable v (the velocity), i.e.,

$$\begin{pmatrix} x \\ v \end{pmatrix}' = \begin{pmatrix} v \\ -U(x)'/m \end{pmatrix}$$

  - Then $E(x,v) = \frac{m}{2}v^2 + U(x)$ is constant along the orbits, that is,

$$\nabla E(x,v) \cdot \begin{pmatrix} v \\ -U'(x)/m \end{pmatrix} = 0$$

    - The gradient of the energy function is orthogonal to the vector field.
  - $E(x,v)$ is a Lyapunov function (global). This happens and induces a fixed point exactly where the velocity is zero and the function takes on a critical value.
  - Linearization at the fixed point $(x_0, 0)$ is

$$\begin{pmatrix} 0 & 1 \\ -\frac{U''(x_0)}{m} & 0 \end{pmatrix}$$

    So $E(x_0, v) > E(x_0, 0)$ for $x \sim x_0$, $v \sim 0$ iff U takes a minimum at x_0. The energy function cannot always stay larger than the energy at the fixed point. Satisfies second Lyapunov condition, but not the first.
  - One question: Classification of planar linear autonomous systems, one on Grönwall, one on qualitative asymptotic analysis using Lyapunov. Three questions total. There will also be some questions (parts of questions, I guess) on perturbative series.

8.2 Misc. Stability Tools

11/18:

- Let $y' = f(y)$, where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth function and x_0 is a fixed point (i.e., $f(x_0) = 0$).
- **Hyperbolic** (fixed point of f): A fixed point $x_0 \in \mathbb{R}^n$ for which $f'(x_0)$ has neither purely imaginary nor zero eigenvalues.
- If x_0 is a hyperbolic fixed point of A, then we know that for the linear system $y' = Ay$, the eigenvalues of A are never purely imaginary by definition.
  - This allows us to decompose $\mathbb{R}^n$ into the direct sum of the **stable subspace** and the **unstable subspace** of the system.
- **Stable subspace** (of x_0 under A): The space of all generalized eigenvectors of A corresponding to eigenvalues λ with $\operatorname{Re}\lambda < 0$. *Also known as* **attracting subspace**. *Denoted by* $\mathbf{E}_{\boldsymbol{s}}$.
- **Unstable subspace** (of x_0 under A): The space of all generalized eigenvectors of A corresponding to eigenvalues λ with $\operatorname{Re}\lambda > 0$. *Also known as* **repelling subspace**. *Denoted by* $\mathbf{E}_{\boldsymbol{u}}$.
- But what if f is not a linear transformation? Then we cannot guarantee the subspace structure, so we need to generalize.
- **Stable subset** (of x_0 under f): The set of all vectors attracted to x_0. *Also known as* **attracting subset**. *Denoted by* $\boldsymbol{W_s(x_0)}$. *Given by*

$$W_s(x_0) = \{x \in \mathbb{R}^n \mid \phi_t(x) \to x_0 \text{ as } t \to +\infty\}$$

- **Unstable subset** (of x_0 under f): The set of all vectors repelled from x_0. *Also known as* **repelling subset**. *Denoted by* $\boldsymbol{W_u(x_0)}$. *Given by*

$$W_u(x_0) = \{x \in \mathbb{R}^n \mid \phi_t(x) \to x_0 \text{ as } t \to -\infty\}$$

- Notice that if $f = A$, then the stable (resp. unstable) subset equals the stable (resp. unstable) subspace.
- Theorem (stable manifold theorem): Let $y' = f(y)$ and let x_0 be a hyperbolic fixed point of f. Then there exists a neighborhood $U(x_0)$ of x_0 such that $U(x_0) \cap W_s(x_0)$ is a smooth submanifold of dimension $\dim \mathbb{E}_s[f'(x_0)]$ that is tangent to $\mathbb{E}_s[f'(x_0)]$ at x_0. An analogous statement holds for $U(x_0) \cap W_u(x_0)$.
- ***k*-dimensional smooth submanifold** (of $\mathbb{R}^n$): A subset congruent to the graph

$$G = (w_1, \dots, w_k; h_1(w), \dots, h_{n-k}(w))$$

  of some smooth function $h : \mathbb{R}^k \to \mathbb{R}^{n-k}$.
- Example: (w, w^2) is a 1-dimensional submanifold of $\mathbb{R}^2$.
  - We know it as the graph of the unit parabola.
- Example: $\left(w_1, w_2, \sqrt{1 - (w_1^2 + w_2^2)}\right)$ is a 2-dimensional submanifold of $\mathbb{R}^3$.
  - In particular, it is the positive hemisphere of the unit two-sphere.
- **Homeomorphism**: A continuous, invertible function with continuous inverse.
  - Essentially, it's a coordinate change function.
- Theorem (Hartman-Grobman Theorem): Let $y' = f(y)$, x_0 a hyperbolic fixed point, and $A = f'(x_0)$. Then there exists a neighborhood $U(x_0)$ and a homeomorphism $h : U(x_0) \to B(x_0, d)$ such that

$$h \circ \phi_t = \mathrm{e}^{tA} \circ h$$

  for $|t|$ small.

Figure 8.1: Hartman-Grobman Theorem visualization.

  - In laymen's terms: Near the hyperbolic fixed point, the orbits are just slight distortions of the linearized system.
- Corollary: Suppose A, B are matrices with no purely imaginary eigenvalues. Then the flows of A, B are topologically conjugate iff $\dim \mathbb{E}_s(A) = \dim \mathbb{E}_s(B)$ (equivalently, iff $\dim \mathbb{E}_u(A) = \dim \mathbb{E}_u(B)$).
- Example: Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Figure 8.2: Topologically conjugate flows.

- Consider the linear autonomous systems $y' = Ay$ and $x' = Bx$.
- Then since

$$e^{tA} = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} \qquad e^{tB} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}$$

we know that the flows are

$$y(t) = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} t \\ 1 \end{pmatrix}$$
$$x(t) = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- Since both A, B have no purely imaginary eigenvalues, these flows will be topologically conjugate by the Corollary.
  - Indeed, we can kind of see that one is a distortion of the other in Figure 8.2.
- We can't expect the coordinate change from Hartman-Grobman to be smooth, but it will exist.

- Example: Let

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -x + y + 3y^2 \\ y \end{pmatrix}$$

  - We can solve this to get

$$\phi_t \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} ze^{-t} + w\sinh(t) + w^2(e^{2t} - e^{-t}) \\ we^t \end{pmatrix}$$

  - Notice that the origin 0 is a fixed point.
  - From this, we can determine that (how??)

$$W_s(0) = x\text{-axis} \qquad W_u(0) = \left\{ \left(\frac{y}{2} + y^2, y \right) \,\middle|\, y \in \mathbb{R} \right\}$$

- What if we can't solve the system in the above example explicitly?
  - Take the Jacobian at 0:

$$A = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

  - Find its stable and unstable subspaces. Calculate eigenvalues and eigenvectors to be

$$\lambda_1 = -1 \qquad \lambda_2 = 1$$
$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

  - Thus, we get v_1 (the x-axis) as the stable subspace, and v_2 as the unstable subspace.

- General procedure for planar systems:
  1. Find all fixed points.
  2. Determine the stability of the fixed points. If hyperbolic, then apply the stable manifold and Hartman theorems. If the eigenvalues are purely imaginary, try to find a Lyapunov function.
  3. Decompose the plane into regions in which the monotonicity of
  $$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$
  is determined, i.e., the signs of the two components of the vector field are determined. This step requires more improvisation.

- Example:
$$\begin{pmatrix} \theta \\ \omega \end{pmatrix}' = \begin{pmatrix} \omega \\ -\sin\theta \end{pmatrix}$$

  picture

  - We only care where $-\pi < \theta < \pi$. The fixed points are
  $$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad\qquad \pm\begin{pmatrix} \pi \\ 0 \end{pmatrix}$$
  - At 0, the linearization has purely imaginary eigenvalues. We have Lyapunov function
  $$E(\theta,\omega) = \frac{1}{2}\omega^2 + (1-\cos\theta)$$
  - At $(\pi, 0)$, the linearization is
  $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
  which has eigenvalues and eigenvectors
  $$\lambda_1 = 1 \qquad\qquad \lambda_2 = -1$$
  $$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad\qquad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
  - Thus, we get orbits around 0 in the θ, ω plane and two subspaces that converge/diverge to $(\pi, 0)$. All of these lines are compatible tangentially.

8.3 Chapter 9: Local Behavior Near Fixed Points

From Teschl (2012)

Section 9.1: Stability of Linear Systems

12/6:
- Goal for the chapter: "Show that a lot of information on the stability of a flow near a fixed point can be read off by linearizing the system around the fixed point" (Teschl, 2012, p. 253).
- Recall the stability discussion for linear systems
$$\dot{x} = Ax$$
from Section 3.2.

  - Additionally, our definition from Section 6.5 is invariant under a linear change of coordinates, so we may work in JCF.
  - Recall that the long-term behavior is determined by the real part of the eigenvalues.
  - "In general, it depends on the initial condition, and there are two linear manifolds $E^+(e^A)$ and $E^-(e^A)$ such that if we start in $E^+(e^A)$ (resp. $E^-(e^A)$), then $x(t) \to 0$ as $t \to +\infty$ (resp. $t \to -\infty$)" (Teschl, 2012, p. 253).

Section 9.2: Stable and Unstable Manifolds

- Goal: Transfer results from the previous section to nonlinear equations.
- **Stable set** (of a fixed point): The set of all points converging to the fixed point x_0 for $t \to +\infty$. *Denoted by* $\boldsymbol{W^+(x_0)}$. *Given by*

$$W^+(x_0) = \{x \in M \mid \lim_{t \to +\infty} |\Phi(t,x) - x_0| = 0\}$$

- **Unstable set** (of a fixed point): The set of all points converging to the fixed point x_0 for $t \to -\infty$. *Denoted by* $\boldsymbol{W^-(x_0)}$. *Given by*

$$W^-(x_0) = \{x \in M \mid \lim_{t \to -\infty} |\Phi(t,x) - x_0| = 0\}$$

- Both the stable and unstable sets are invariant under the flow.
- We know that for small t, the solutions are adequately described by the linearization, but what about for large t?
  - In this section, we generalize the Section 6.5 result for $n = 1$ stability and $A = f'(x_0)$ to higher dimensions.
- **Hyperbolic** (fixed point of f): A fixed point x_0 for which the linearization $f'(x_0)$ has no eigenvalues with zero real part.
  - Note that this is equivalent to the definition from class: The "zero real part" condition can be divided into two cases (equal to zero and nonzero but purely imaginary). This definition says no "zero real part;" that definition says not either of the latter two cases.

Week 9

Periodicity

9.1 Periodic Solutions of Planar Systems

11/28:

- Last lectures: Special solutions to planar systems. Usually encountered in applications of ODEs (e.g., the homework). If we encounter ODEs in our physical sciences lives, we may need to remember this.
- There will be a Monday office hours that coincides with the regular lecture time.
- Special (periodic) solutions of planar systems.
- For a planar autonomous system
$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix}$$
a periodic solution is equivalent to a closed orbit.
- So geometrically, we'll be studying closed orbits. Analytically, we'll be studying periodic systems.
- Simple examples: Harmonic oscillator
$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} y \\ -x \end{pmatrix}$$
- Less trivial example: The pendulum
$$\begin{pmatrix} \theta \\ \omega \end{pmatrix}' = \begin{pmatrix} \omega \\ -\sin\theta \end{pmatrix}$$
- We can easily find Lyapunov functions for these two systems; both functions are the energy function.
- Different orbital graphs: The first one is concentric circles; the second one is only sometimes periodic. *picture*
- In general, these periodic cycles are dense in the plane.
- However, we can also, at the other extreme, have isolated periodic solutions, referred to as **limit cycles**.
- Simplest example:
$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -y - [1-(x^2+y^2)]x \\ x - [1-(x^2+y^2)]y \end{pmatrix}$$
  - Hard to see the behavior in Cartesian coordinates; much easier in polar. If we use $r = \sqrt{x^2+y^2}$ and $\theta = \arctan(y/x)$. We will see equations of this type in our homework.

– Differentiating $r = \sqrt{x^2 + y^2}$ implicitly with respect to time, we get

$$\begin{aligned}
\frac{\mathrm{d}r}{\mathrm{d}t} &= \frac{xx'}{\sqrt{x^2+y^2}} + \frac{yy'}{\sqrt{x^2+y^2}} \\
&= -\frac{xy}{\sqrt{x^2+y^2}} + \frac{1-r^2}{r}x^2 + \frac{xy}{\sqrt{x^2+y^2}} + \frac{1-r^2}{r}y^2 \\
&= r(1-r^2)\cos^2\theta + r(1-r^2)\sin^2\theta \\
&= r(1-r^2)
\end{aligned}$$

– Differentiating $\theta = \arctan(y/x)$ implicitly with respect to time, we get

$$\begin{aligned}
\frac{\mathrm{d}\theta}{\mathrm{d}t} &= \frac{-\frac{y}{x^2}x'}{1+\left(\frac{y}{x}\right)^2} + \frac{\frac{1}{x}y'}{1+\left(\frac{y}{x}\right)^2} \\
&= \frac{-yx'}{x^2+y^2} + \frac{xy'}{x^2+y^2} \\
&= 1
\end{aligned}$$

– Thus, we can transform the original equation to

$$\begin{pmatrix} r \\ \theta \end{pmatrix}' = \begin{pmatrix} r(1-r^2) \\ 1 \end{pmatrix}$$

for $r > 0$ and $\theta \in \mathbb{R}$.

– This ODE can be explicitly solved, but since we are interested in qualitative behavior, we will not do that.

– The unit circle partitions the xy-plane into two parts (inside and outside). *picture*

– Let's start on the unit circle. Then we just spiral around on it with constant velocity ($r' = 0$ and $\theta' = 1$).

– Let's now start inside. Since $\theta(t) = t + \theta(0)$, and r' is positive, we get a spiral that approaches the unit circle.

– If we start outside, we get a spiral that starts outside and spirals toward the unit circle.

– Thus, in this case, the unit circle is the unique limit cycle of the system.

- Shao suggests we search for iodine clock videos on YouTube to help with the homework. The iodine clock is described by the limit cycles.

- Historical remark: David Hilbert posed 23 questions at the beginning of the 20th century. The 16th one asked about planar polynomial systems. For these, is it possible to estimate the number of limit cycles. Even for the case of quadratic polynomials, the question is still open! For quadratics, we know that there can be 1, 2, 3, or 4 cycles, but we have no idea whether or not there is an upper bound. This is a central open problem in the study of ODEs.

- Basic theorem in this area is as follows.

- Theorem (Poincaré-Bendixson Theorem): Let $\Omega \in \mathbb{R}^2$ be open, $f(x)$ a vector field on Ω. Fix $x \in \Omega$. Define

$$\omega(x) := \{z \in \Omega \mid \text{there is a sequence } t_n \to +\infty \text{ such that } \phi_{t_n}(x) \to z\}$$

and

$$\alpha(x) := \{z \in \Omega \mid \text{there is a sequence } t_n \to -\infty \text{ such that } \phi_{t_n}(x) \to z\}$$

That is, if you reverse the direction of time, $\alpha(x)$ collects all of the points. Also, let $\omega(x) \subset \Omega$ be compact and nonempty. In particular, there are three mutually exclusive cases for these limit sets.

1. $\omega(x)$ (or $\alpha(x)$[1]) is a fixed point.
2. $\omega(x)$ is a limit cycle.
3. $\omega(x)$ consists of finitely many fixed points, together with curves joining these fixed points.

Proof. The proof relies largely on algebraic topology, so we will not go into it in detail or even sketch it. The statement will be sufficient for our purposes. □

- Theorem (Annulus theorem of Bendixson): Suppose C_1, C_2 are closed simple planar curves such that geometrically, one contains the other. We call the annular region (between the two curves) A. Suppose $f(x)$ is a planar vector field which points inward at every point of ∂A (the boundary of A). Then the annular region A is an invariant region of the plane. In particular, if A does not contain any fixed points, then it must contain a limit cycle. As before, curves within and without spiral towards it. *picture*
- Can produce beautiful diagrams: Zhifen Zhang's example — Xiao and Zhang (2008) — is the 4 limit cycle one. After her research, mathematicians found a family with four limit cycles.
- System from the homework: Consider the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} a - x - \frac{4xy}{1+x^2} \\ b(x - \frac{xy}{1+x^2}) \end{pmatrix}$$

  picture

  - We take $a, b > 0$.
  - Every vector points in toward a, which points straight upward.
  - On the boundary,

$$\begin{pmatrix} a - x - \frac{4xy}{1+x^2} \\ b(x - \frac{xy}{1+x^2}) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} > 0$$

    so we have a strict Lyapunov function.
  - Thus, any orbit in the first quadrant can never escape, reflecting our expectation that the concentration can never be negative.
  - Spoiler: Looks like a rectangular region.
  - We are also asked to find the fixed point and investigate its stability.
- First step: Find fixed points.
- Second step: Lyapunov functions.
- Third step: Divide the vector field into positive and negative regions. Requires some improvization.
- Fourth step: Check for an annular region.
- Note: Note of these arguments can be generalized to higher dimensional systems. The regularity of the Poincaré-Bendixson system disappears as we go to higher dimensions.
- Lorenz system: Oversimplified 3D quadratic system that serves as a simplification model for weather systems.
  - Lorenz was an astronomer.
  - He discovered that even if the systems are very close together, the orbits will be separated indefinitely as time evolves on. It's not just about being stable or unstable, but overall global chaotic behavior.

[1] We just have to reverse the time.

  - Mathematicians quickly discovered that the Lorenz system has a **strange attractor**. An attractor for a planar system can only be closed orbits. In the Lorenz system, we have a strange type of butterfly. The dimension of the butterfly is not even an integer. Have to introduce Hausdorff measure to understand length and area in the more general framework of curves or surfaces.
  - No more detail, but a classical example of a chaotic ODE system worth mentioning. This phenomenon cannot appear for planar systems; even increasing the dimension by 1 can lead to chaos.
  - No widely accepted definition of mathematical chaos, but generally accepted ones are very irregular global attractors and points that are arbitrarily close and arbitrarily far from each other.
  - Fractorial sets.
- Last lecture (Wednesday): Another example that has a limit cycle.
- Friday will be a review.

9.2 Nonlinear Oscillation

11/30:

- Friday: Review and questions session.
- Monday: OH during class time in the class room and normal OH.
- Exam time: Wednesday, December 7 from 7:30-9:30 AM. Will happen in this classroom; Shao hopes to finish grading the same day.
- Last HW and review outline is due 12/2. 130 total points for the grade. 90/130 gives an A. That's 69% and above.
  - Minus 10, minus 10, for the grading below A range.
- Last example that arises naturally.
- RLC circuit: Nonlinear oscillation. Occurred in resonance. Nonlinear oscillation in K.
- The circuit that we're interested in still is an RLC circuit (see Figure 5.5b).
- We assume that the resistor satisfies Ohm's law, i.e., $I_R = V_R/R$. This is the linear case. Causes the voltage to decay exponentially fast, where the resistor acts as a kind of damping factor.
- However, for some more delicate cases, Ohm's law might fail and we might get a nonlinear replacement $V_R = R(I_R)$. We let R be any suitable function. We still have Kirchoff's law, i.e., that I_R, I_C, and I_L are all the same and we can refer to them as I.
- The system we're interested in is
$$\begin{cases} LI' &= -V_C - R(I)\\ CV' &= I \end{cases}$$
- Comparing to the linear case, we just want to replace V_R/R in the same system with $R(I)$. After some suitable scaling, we can get the system
$$\begin{cases} x' &= y - f(x)\\ y' &= -x \end{cases}$$
- This is a classical nonlinear oscillation model, typically referred to as the **Liénard equation**.
- The first case of the Liénard equation studied in detail was when
$$f(x) = \mu\left(\frac{x^3}{3} - x\right)$$

  - Discovered by the engineer van der Pol while he was investigating RLC circuits, and hence referred to as the **van der Pol oscillator**.
- Theorem: Suppose $f(x)$ is odd, i.e., $f(x) = -f(-x)$. Also suppose that $f(x) < 0$ for $x \in (0, \alpha)$ for some $\alpha > 0$ (i.e., $f(x)$ is less than zero for sufficiently small positive values of x). Moreover, suppose that $\liminf_{x\to\infty} f(x) > 0$ (i.e., f does not diverge). Lastly, suppose that $f(x)$ is increasing for $x > \alpha$. Then the conclusion is that the Liénard system has a unique closed orbit. Every other orbit except for the trivial orbit $(0,0)$ will be attracted to this periodic solution as the system evolves ($t \to +\infty$).
  - This explains the term "nonlinear oscillation." As long as we are away from the equilibrium, we converge to it??
- Wrt the van der Pol oscillator, we sketch the boundary first and then put in the limit cycle. The origin is an unstable fixed point, and any origin starting from the origin will spiral and approximate the limit cycle.
  - ?? is guaranteed by the Poincaré-Bendixson theorem.
- The proof does not use any advanced math, but it is complex, so we'll sketch it. The proof is in Teschl (2012) if we're interested. It only involves calculus.
- End of lecture.
- First problem, second section of HW: First integral: Conservative law.
- What is the correct first variational equation?
- Difference between not Lyapunov stable and completely unstable?
  - Something is repelled away vs. everything is repelled away.
- The iodine clock is not super hard, but it needs some improvisation. That's what we talked about at the end of last lecture.
- Friday will be a review of everything since the second midterm.

9.3 Question Session

12/2:

- The actual Lecture 9.1 system is

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -y - [1 - (x^2 + y^2)]x \\ x - [1 - (x^2 + y^2)]y \end{pmatrix}$$

- Cannot derive a phase portrait spiral from a potential energy function.
  - Relevant to 3c.
  - Idea is to show that the potential energy well is decreasing along any trajectory.
- $r = \sqrt{c_0}$.
  - Limit cycle is omega and alpha limit cycle.
- Solving for the stable and unstable manifolds.
  - Stable set consists of points which are attracted to the equilibrium. Curves are not attracted or repelled.
  - Stable subset: Points (z, w) such that $(x, y) \to (0, 0)$ as $t \to +\infty$. Stable subset: necessitates taking $w = 0$ and then z can be anything, so x axis.

- Unstable: As $t \to -\infty$, w can be anything and $y(t) \to 0$. The $+t$ terms will go to zero as $t \to -\infty$, and then we must have $z - w/2 = 0$. Put it in the form

$$\begin{aligned} x(t) &= (w^2(e^{2t} - e^t) + \frac{we^t}{2}) + (z - \frac{w}{2})e^{-t} \\ &= ((we^t)^2 + \frac{we^t}{2}) - w^2e^t + (z - \frac{w}{2})e^{-t} \\ &= (y^2 + \frac{y}{2}) - w^2e^t + (z - \frac{w}{2})e^{-t} \end{aligned}$$

- There is a typo in the original form. There should be e^{-t} for the last rightmost term above. We will converge and diverge along the manifolds.
- I have a confusion in the stable/unstable subset definition? Unstable subset isn't the set of all points with orbits that diverge as $t \to +\infty$; it's the set of all points that diverge away from x_0.

• Solving for the stable and unstable manifolds of a planar ODE given the explicit solution.

- We will treat

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} -x + y + 3y^2 \\ y \end{pmatrix}$$

from Lecture 8.2.
- The correct flow is as follows (there was a typo in class).

$$\phi_t \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} ze^{-t} + w\sinh(t) + w^2(e^{2t} - e^{-t}) \\ we^t \end{pmatrix}$$

- The stable manifold is going to be the set of all $x \in \mathbb{R}^2$ such that $\phi_t(x) \to 0$ as $t \to +\infty$. Approach: Find values of z, w such that the solution converges to zero componentwise.
  - Since $y(t) = we^t$, we must have $w = 0$; otherwise, we will get exponential divergence as $t \to +\infty$.
  - Thus, $x(t) = ze^{-t}$. This function converges to zero for any value of z, so we may let z be arbitrary.
  - But the set of all points

$$\begin{pmatrix} z \\ 0 \end{pmatrix}$$

is the x-axis!
- Unstable manifold: We need to find the set of all points $x \in \mathbb{R}^2$ such that $\phi_t(x) \to 0$ as $t \to -\infty$. Approach: Again, go by components.
  - $y(t) = we^t$ will converge to 0 as $t \to -\infty$ for all w, so this component does not put any restrictions on w. Note that it also does not put any restrictions on z since it does not even contain z.
  - Working with the other one, we expand and combine all e^{at} terms for $a > 0$ and all e^{bt} terms for all $b < 0$.

$$\begin{aligned} x(t) &= ze^{-t} + w\sinh(t) + w^2(e^{2t} - e^{-t}) \\ &= ze^{-t} + w \cdot \frac{e^t - e^{-t}}{2} + w^2e^{2t} - w^2e^{-t} \\ &= ze^{-t} + \frac{w}{2}e^t - \frac{w}{2}e^{-t} + w^2e^{2t} - w^2e^{-t} \\ &= \left[w^2e^{2t} + \frac{w}{2}e^t\right] + \left[z - \frac{w}{2} - w^2\right]e^{-t} \end{aligned}$$

  - The left term above will clearly converge to 0 as $t \to -\infty$.

- However, the right term will diverge to ∞ as $t \to -\infty$ unless $z - w/2 - w^2 = 0$, so we take this to be our condition.
- Indeed, this implies that $z = w/2 + w^2$ is a constraint on z, but w can still take on any value, so our solution is
$$W_u(0) = \left\{ \left(\frac{y}{2} + y^2, y \right) \,\middle|\, y \in \mathbb{R} \right\}$$
as desired.

- Number 5:
  - What is a vector field in 1d? Vectors pointing in the positive or negative x direction (just a function).
  - Set of points should be a subset of the real line (an interval).
  - You can only approach the zero.
- Number 3 energy term.
  - Multiply both sides by x' to get
$$x'x'' = \frac{1}{2}(x')^2$$
  - We have
$$\begin{aligned} 0 &= x'' + bx' + U'(x) \\ &= x'x'' + b|x'|^2 + x'U'(x) \\ &= \left(\frac{1}{2}(x')^2 \right)' + b|x'|^2 + (U(x))' \\ -b|x'|^2 &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2}(x')^2 + U(x) \right) \end{aligned}$$
- Lyapunov stuff. Was a question in HW6.
  - ...
  - Intuitive justification for this Lyapunov function?
  - Most natural way is to look at when A is diagonalizable.
  - Get expressions with negative eigenvalues.
  - We have that the sum is equal to $\frac{\mathrm{d}}{\mathrm{d}t}\langle y, Dy \rangle$. The INP is equal to $2\langle Dy, Dy \rangle$.
  - So it's a weighted norm.
- We'll be allowed to bring the JNF notes to the exam!

9.4 Office Hours (Shao)

- Question 3(1): How do we derive the energy function?
  - We don't actually need to give a final expression for the energy function; just show that it's always decreasing.
- Question 3(2): How should we apply the stable manifold theorem and Hartman linearization theorem?
  - By the stable manifold theorem, we can determine source, sink, saddle.
  - By the Hartman linearization theorem, we can further characterize the saddle point by saying that all orbits not on the stable/unstable manifolds (which we don't have to find) approach the fixed point and then diverge away.

- Submit a sketch of the local behavior of each one with my answer!!

- Question 4(3): What does this even mean?
  - Solution is $1/3(x^3 + y^3) + xy = a$.
  - As $a \to 1/3$, the orbits have shape very similar to an ellipse.
  - We need $a = 0$ because we must have the system pass through zero.
  - This is related to the Lotka-Voterra model from Lecture 2.1.
  - F is called a first integral.
    - To be more general, in an undamped Newtonian system, the energy function is a first integral of the system. Think like quantum mechanics and vibrational energy levels being equivalent to segments of the energy parabola. More generally, though, this perspective applies to all mechanical systems via Hamiltonian mechanics. We don't need to understand Hamiltonian mechanics for this course, though, because we've defined phase spaces independently (a phase space is like the (θ, ω) plane for the harmonic oscillator).
    - We talked about first integrals when we discussed the Kepler problem. The Kepler problem is covered in Section 8.5 of Teschl (2012).
  - We take

$$\begin{aligned}
\frac{x'}{y'} &= \frac{-x-y^2}{x^2+y} \\
(x^2+y)\frac{\mathrm{d}x}{\mathrm{d}t} &= (-x-y^2)\frac{\mathrm{d}y}{\mathrm{d}t} \\
(x^2+y)\frac{\mathrm{d}x}{\mathrm{d}t} + (x+y^2)\frac{\mathrm{d}y}{\mathrm{d}t} &= 0
\end{aligned}$$

    Solve this for F, which will be a polynomial.

- Question 4(4): What does this even mean?
  - For the $c = 0$ case, the stable set is the part of the curve tangent to the x-axis and the unstable set is the part of the curve tangent to the y-axis. We get a change from stable to unstable at $(0, 0)$ and $(-1.5, -1.5)$.

- Enzyme kinetics 1(1): What does "two first integrals" mean?
  - First conservation law: Conservation of the substrate in all its forms. Second: Conservation of the enzyme in all its forms.
  - Sum of the second and third equations equals zero.
  - Sum of first, third, and fourth is zero.
  - We have

$$0 = \frac{\mathrm{d}[\mathrm{E}]}{\mathrm{d}t} + \frac{\mathrm{d}[\mathrm{ES}]}{\mathrm{d}t}$$

    This is also obvious from their definitions. Thus, $[\mathrm{E}] + [\mathrm{ES}] = [\mathrm{E}]_0$ is a first integral.
  - This is not an equation of exact form because of 3 derivatives, but it still implies a conservation law/has a first integral:

$$0 = \frac{\mathrm{d}[\mathrm{S}]}{\mathrm{d}t} + \frac{\mathrm{d}[\mathrm{ES}]}{\mathrm{d}t} + \frac{\mathrm{d}[\mathrm{P}]}{\mathrm{d}t}$$

    That first integral is $[\mathrm{S}] + [\mathrm{ES}] + [\mathrm{P}] = [\mathrm{S}]_0$

- Iodine Clock: 2(1)?
  - Ben knows what's going on here.
  - There are curves that divide the first quadrant into regions of definite sign.
  - Definite sign: Neither component changes sign in a certain region

9.5 Chapter 7: Planar Dynamical Systems

From Teschl (2012).

Section 7.2: Examples from Electrical Engineering

12/6:

- Consider an RLC circuit again, but this time with a resistor of arbitrary characteristic

$$V_R = R(I_R)$$

  - Ohm's law asserts that $R(I_R) = RI_R$, where the resistance R is a constant.
  - But what about a stranger, more sophisticated element? We must have $R(0) = 0$ (since there's no potential difference if there's no current) for any characteristic, but other than that, we're pretty free.

- **Diode**: A circuit element that lets the current pass in only one direction.
  - For example, the characteristic of a diode is given by

$$V = \frac{kT}{q}\log\left(1 + \frac{I}{I_L}\right)$$

    where I_L is the leakage current, q is the charge of an electron, k is the Boltzmann constant, and T is the absolute temperature.
  - Implications: In the positive direction, very little voltage gives a large current; in the negative direction, you will get almost no current even for fairly large voltages.

- As in class, we obtain the system

$$\begin{aligned}
L\dot{I}_L &= -V_C - R(I_L)\\
C\dot{V}_C &= I_L
\end{aligned}$$

  where $R(0) = 0$ and $L, C > 0$.

- Additional note.
  - Kirchoff's laws and the substitutions from Section 3.3 imply

$$\begin{aligned}
I_LV_L + I_CV_C + I_RV_R &= 0\\
LI_L\dot{I}_L + CV_C\dot{V}_C + I_RR(I_R) &= 0\\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{L}{2}I_L^2 + \frac{C}{2}V_C^2\right) &= -I_RR(I_R)
\end{aligned}$$

  - Conclusion: The energy dissipated in the resistor has to come from the inductor and capacitor.

- **Liénard's equation**: The result of scaling the above system. *Given by*

$$\begin{aligned}
\dot{x} &= y - f(x)\\
\dot{y} &= -x
\end{aligned}$$

- The additional note now reads

$$\frac{\mathrm{d}}{\mathrm{d}t}W(x, y) = -xf(x)$$

  where

$$W(x, y) = \frac{x^2 + y^2}{2}$$

  - If $xf(x) > 0$ in a neighborhood of $x = 0$, then W is a Lyapunov function and hence $(0,0)$ is stable.

- Theorem 7.5: Suppose $xf(x) \geq 0$ for all $x \in \mathbb{R}$ and $xf(x) > 0$ for $0 < |x| < \varepsilon$. Then every trajectory of Liénard's equation converges to $(0,0)$.

  Proof. Given. □

- Conversely, if $xf(x) < 0$ for $0 < |x| < \varepsilon$, then $(0,0)$ is unstable (and the distance to the fixed point will actually grow).
- Teschl (2012) works through proving the main theorem from lecture (Theorem 7.8 below).
- Theorem 7.8: Suppose f satisfies requirements (i)-(iii) below.
  (i) f is odd, that is, $f(-x) = -f(x)$.
  (ii) $f(x) < 0$ for $0 < x < \alpha$ ($f(\alpha) = 0$ without restriction).
  (iii) $\liminf_{x\to\infty} f(x) > 0$ and, in particular, $f(x) > 0$ for $x > \beta$ ($f(\beta) = 0$ without restriction).
  (iv) $f(x)$ is monotone increasing for $x > \alpha$ (i.e., $\alpha = \beta$).

  Then Liénard's equation has at least one periodic orbit encircling $(0,0)$.

  If in addition (iv) holds, this periodic orbit is unique and every trajectory (except $(0,0)$) converges to this orbit as $t \to \infty$.

- The classical application of this theory is to **van der Pol's equation**.
- **Van der Pol's equation**: The following ODE, which models a triode circuit. *Given by*

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = 0, \quad \mu > 0$$

- We can show that van der Pol's equation is equivalent to Liénard's equation with

$$f(x) = \mu\left(\frac{x^3}{3} - x\right)$$

- Therefore, by Theorem 7.8, van der Pol's equation has a unique periodic orbit and all trajectories converge to this orbit as $t \to \infty$.

Section 7.3: The Poincaré-Bendixson Theorem

- In all previous examples, the solutions of ODEs have either converged to a fixed point or a periodic orbit.
  - This is normal behavior, and in this section we will classify all possible $\omega_\pm$-limit sets (for planar systems).
  - Note that the difference between $\mathbb{R}^2$ and $\mathbb{R}^n$ ($n \geq 3$) arises from the validity of the **Jordan Curve Theorem** in $\mathbb{R}^2$ and its being false in higher dimensions.
- **Jordan curve**: A homeomorphic image of the circle S^1. *Denoted by* $\boldsymbol{J}$.
- **Jordan Curve Theorem**: Every Jordan curve dissects $\mathbb{R}^2$ into two connected regions. In particular, $\mathbb{R}^2 \setminus J$ ahs two components.
- Teschl (2012) builds up to proving the Poincaré-Bendixson theorem.
- Theorem 7.16 (generalized Poincaré-Bendixson): Let M be an open subset of $\mathbb{R}^2$ and $f \in C^1(M, \mathbb{R}^2)$. Fix $x \in M$, $\sigma \in \{\pm\}$, and suppose $\omega_\sigma(x) \neq \emptyset$ is compact, connected, and contains only finitely many points. Then one of the following cases holds.

(i) $\omega_\sigma(x)$ is a fixed orbit.
(ii) $\omega_\sigma(x)$ is a regular periodic orbit.
(iii) $\omega_\sigma(x)$ consists of (finitely many) fixed points $\{x_j\}$ and non-closed orbits $\gamma(y)$ such that $\omega_\pm(y) \in \{x_j\}$.

- Teschl (2012) gives an example of the third case.
- Lemma 7.17: The interior of every periodic orbit must contain a fixed point.
- **Limit cycle**: A periodic orbit attracting other orbits.
- Lemma 7.18: Let $\gamma(y)$ be an isolated regular periodic orbit (such that there are no other periodic orbits within a neighborhood). Then every orbit $\gamma(x)$ starting sufficiently close to $\gamma(y)$ will have either $\omega_-(x) = \gamma(y)$ or $\omega_+(x) = \gamma(y)$.
- Example: In general, the system

$$\dot{x} = -y + f(r)x \qquad\qquad \dot{y} = x + f(r)y$$

becomes

$$\dot{r} = rf(r) \qquad\qquad \dot{\theta} = 1$$

for any function f.

9.6 Chapter 8: Higher Dimensional Dynamical Systems

From Teschl (2012).

Section 8.1: Attracting Sets

- Not much of relevance here.
- A bit on the Duffing equation from HW7.
- **Topologically transitive** (set): A closed invariant set Λ such that for any two open sets $U, V \subset \Lambda$, there is some $t \in \mathbb{R}$ such that $\Phi(t, U) \cap V \neq \emptyset$.
- **Attractor**: An attracting set which is topologically transitive. *Denoted by* $\boldsymbol{\Lambda}$.

Section 8.2: The Lorenz Equation

- **Lorenz equation**: One of the most famous dynamical systems which exhibits chaotic behavior. *Given by*

$$\begin{aligned}\dot{x} &= -\sigma(x - y)\\ \dot{y} &= rx - y - xz\\ \dot{z} &= xy - bz\end{aligned}$$

where $\sigma, r, b > 0$.

  - "Lorenz arrived at these equations when modelling a two-dimensional fluid cell between two parallel plates which are at different temperatures. The corresponding situation is described by a complicated system of nonlinear partial differential equations. To simplify the problem, he expanded the unknown functions into Fourier series with respect to the spacial coordinates and set all coefficients except for three equal to zero. The resulting equation for the three time dependent coefficients is [the above]. The variable x is proportional to the intensity of convective motion, y is proportional to the temperature difference between ascending and descending currents, and z is proportional to the distortion from linearity of the vertical temperature profile" (Teschl, 2012, p. 234).

- **Strange attractor**: An attractor that has a complicated set structure.

References

Guillemin, V., & Haine, P. J. (2018). *Differential forms* [https://math.mit.edu/classes/18.952/2018SP/files/18.952_book.pdf].

Teschl, G. (2012). *Ordinary differential equations and dynamical systems* [https://www.mat.univie.ac.at/~gerald/ftp/book-ode/ode.pdf].

Xiao, D., & Zhang, Z. (2008). On the existence and uniqueness of limit cycles for generalized liénard systems. *J. Math. Anal. Appl.*, *343*(1).