MATH 27300 (Basic Theory of Ordinary Differential Equations) Notes

Steven Labalme

October 20, 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**: Any 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 n** (ODE): An ODE for which the n^{th} derivative of y is the highest-order derivative involved (and is involved).
- $y'=f(t,y)$ or $Y^{(n)}=F(t,Y,Y',\dots,Y^{(n-1)})$.
  - We can transform this second form into the first form via

$$y=\begin{pmatrix} Y \\ Y' \\ \vdots \\ Y^{(n-1)} \end{pmatrix} \qquad f(t,y)=\begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ F(t,y_1,y_2,\dots,y_{(n-1)}) \end{pmatrix}$$

making $y' = f(t,y)$ equal to the system of equations

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_{n-1}' &= F(t, y_1, \dots, y_{n-1}) \end{aligned}$$

- Think about this conversion more.

- Thus, we mainly focus on equations of the form $y' = f(t,y)$, 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

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

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

$$y' = f(y)$$

  - More equivalence w/ vector-valued functions?
- **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

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

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-Lindelof 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**: ...
- **Dynamical system**: A law under which a particle evolves over time. $y' = f(t, y)$, IVP is LWP
- Consider $\Phi(t, x)$ such that

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

- **Steady flow**: A vector field on a manifold contained in $\mathbb{R}^2$ or $\mathbb{R}^3$ that does not vary with time.
  - A velocity 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).
  - 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

$$\begin{aligned} R' &= k_1R - aWR \\ W' &= -k_2W + bWR \end{aligned}$$

- 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.
- 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* brachisto "shortest" + chrone "time."
  - 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

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

  - The time it takes for the particle to traverse an infinitesimal section of track of arc length ℓ is ℓ/v.
  - The track should be given by $y = y(x)$.
  - Arc length

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

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

$$\int_0^a \frac{\sqrt{1 + (y'(x))^2}\,\mathrm{d}x}{\sqrt{2gy(x)}}$$

  - We also have $y(0) = 0$ and $y(a) = y_0$.
  - Functionals: Mapping from a function space to numbers; we want to find y such that the above integral is minimized.

- 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.
- Take a function h, vanishing at a, b.
- Let $f(t) = J[y + th]$. Then

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

  and hence

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

- 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.
- 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,

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

- This gives us

  $$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}[F(y, y')] &= \frac{\partial F}{\partial z}(y, y') \cdot y' + \frac{\partial F}{\partial w}(y, y') \cdot y'' \\ \frac{\partial F}{\partial z}(y, y')y &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}(y, y') \right] y' \\ \frac{\mathrm{d}}{\mathrm{d}x}[F(y, y')] &= \underbrace{\frac{\partial F}{\partial w}(y, y')}_{U} \cdot \underbrace{y''}_{V'} + \underbrace{\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}(y, y') \right]}_{U'} \cdot \underbrace{y'}_{V} \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial F}{\partial w}(y, y') \cdot y' \right] \end{aligned}$$

- This reduces to the first-order equation

$$F(y,y') - \frac{\partial F}{\partial w}(y,y') \cdot y' = A$$

- Since $F(x,z,w)$ is known, we have that

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

- Plugging into the E-L equation gives us

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

  where the second line above is a separable differential equation.

- The solution is the **cycloid**

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

  where the specific parameters come from the boundary values.

- **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.

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).

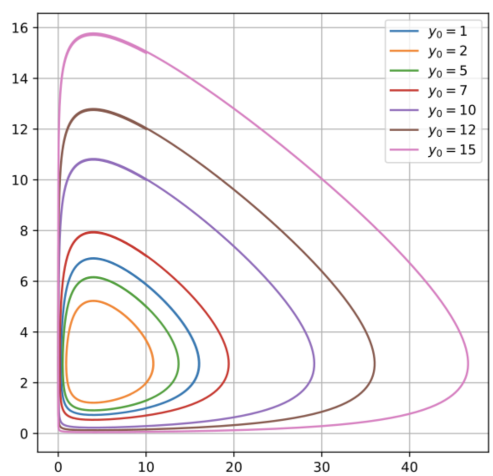

Figure 2.1: Lotka-Volterra solution curves.

- 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 following variation on the E-L equation 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

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

- For linear homogeneous equations $\mathrm{d}y/\mathrm{d}t = p(t)y$, we have

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

- 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
$$x'' = -\frac{k}{m} x$$
- 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.
- Example:
  - Consider
$$y' - \mu y = A\mathrm{e}^{\lambda t}$$
  - By the Duhamel equation, we have that a particular solution is of the form
$$A \int_0^t \mathrm{e}^{\mu(t-\tau)} \mathrm{e}^{\lambda\tau} \mathrm{d}\tau$$

- 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 $\mu = \lambda$.

- 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

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -k_1[\mathrm{E}][\mathrm{S}] + k_{-1}[\mathrm{ES}] \tag{1}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{E}] = -k_1[\mathrm{E}][\mathrm{S}] + (k_{-1} + k_2)[\mathrm{ES}] \tag{2}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ES}] = k_1[\mathrm{E}][\mathrm{S}] - (k_{-1} + k_2)[\mathrm{ES}] \tag{3}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{P}] = k_2[\mathrm{ES}] \tag{4}$$

- We can reduce these rate laws to the 2D system

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -k_1([\mathrm{E}]_0 - [\mathrm{ES}])[\mathrm{S}] + k_{-1}[\mathrm{ES}] \tag{5}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{ES}] = k_1([\mathrm{E}]_0 - [\mathrm{ES}])[\mathrm{S}] - (k_{-1} + k_2)[\mathrm{ES}] \tag{6}$$

- QSSA: Quasi steady-state assumption.
  - Assume that $[\mathrm{E}]_0/[\mathrm{S}]_0 \lll 1$.
  - Also assume that $\mathrm{d}[\mathrm{ES}]/\mathrm{d}t \approx 0$.
- Then

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

  where $k_M = (k_{-1} + k_2)/k_1$.
- Sub in the above to Equation 5:

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{S}] = -\frac{k_2[\mathrm{E}]_0[\mathrm{S}]}{k_M + [\mathrm{S}]}$$

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

$$\int_{[\mathrm{S}]_0}^{[\mathrm{S}]} -\frac{(k_M + z)\,\mathrm{d}z}{z v_{\max}} = \int_0^t \mathrm{d}t$$
$$-\frac{k_M}{v_{\max}} \log\frac{[\mathrm{S}]}{[\mathrm{S}]_0} - \frac{1}{v_{\max}}([\mathrm{S}] - [\mathrm{S}]_0) = t$$
$$-\frac{k_M}{v_{\max}} \frac{[\mathrm{S}]}{[\mathrm{S}]_0} \mathrm{e}^{-v_{\max}^{-1}([\mathrm{S}] - [\mathrm{S}]_0)} = \mathrm{e}^t$$

- The above equation is of the following form, for $x > 0$, $w(x) \sim s$, $x \sim 0$, and $w(x) \sim \log x$??

$$w(x)\mathrm{e}^{w(x)} = x$$

- 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:
$$E = \frac{1}{2}m|x'|^2 + \frac{1}{2}kx^2$$

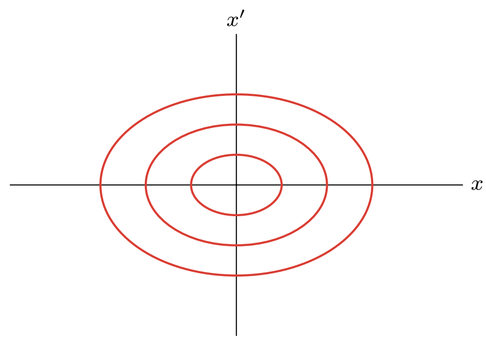

Figure 2.2: Conservation of mechanical energy in the harmonic oscillator.

  - 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:
$$\begin{aligned}0 &= l\theta' + g\sin\theta\\ &= \ell\theta''\theta' + g\sin\theta\cdot\theta'\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Bigg(\underbrace{\frac{\ell}{2}|\theta'|^2 - g\cos\theta}_{E}\Bigg)\end{aligned}$$
- 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.
- From Newton's second and third law, we get
$$m_1 x_1'' = \frac{U'(|x_1 - x_2|)^{x_1 - x_2}}{|x_1 - x_2|} \qquad m_2 x_2'' = \frac{U'(|x_1 - x_2|)^{x_1 - x_2}}{|x_1 - x_2|}$$
- 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
$$\frac{m_1}{M}x_1 + \frac{m_2}{M}x_2$$
    moves inertially.
    - Define the center of mass to be the origin.
- Conservation of angular momentum:
$$[\mu(x_1 - x_2)' \times (x_1 - x_2)]' = 0$$
    - $\mu = m_1m_2/(m_1 + m_2)$.
    - $\times$ indicates the cross product.
    - $L = \mu(x_1 - x_2) \times (x_1 - x_2)$.
- Conservation of mechanical energy:
$$\begin{aligned}\mu q'' + U'(|q|)\frac{q}{|q|} &= 0\\ \frac{\mu}{2}|q'|^2 + U(|q|) &= E\end{aligned}$$
    - $q = x_1 - x_2$.
- Introduce polar coordinates (r, ϕ).
    - Then $r^2\phi'/2 = \ell_0$, $r = r(\phi)$, and $\mathrm{d}r/\mathrm{d}\phi = r'(t)/\phi'(t)$.
    - It follows that
$$\frac{\mu}{2}(|r'|^2 + |\phi'|^2) + U(r) = E$$
    - Then
$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 + r^2 = \frac{2G\mu r^3}{\ell_0^2} + \frac{2Er^4}{\mu\ell_0^2}$$
    - The substitution $\mu = 1/r$ yields
$$\left(\frac{\mathrm{d}\mu}{\mathrm{d}\phi}\right)^2 + \mu^2 = \frac{2G\mu}{\ell_0^2}\mu + \frac{2E}{\mu\ell_0^2}$$

- Differentiating again gives

$$2\frac{\mathrm{d}\mu}{\mathrm{d}\phi}\frac{\mathrm{d}^2\mu}{\mathrm{d}\phi^2}+2\frac{\mathrm{d}\mu}{\mathrm{d}\phi}\mu=\frac{2G\mu}{\ell_0^2}\frac{\mathrm{d}\mu}{\mathrm{d}\phi}$$

- Substituting $\mu=\cos(t)$ gives

$$\frac{\mathrm{d}^2\mu}{\mathrm{d}\phi^2}+2\mu=\frac{2G\mu}{\ell_0^2}$$

  or

$$r=\frac{1}{G\mu/\ell_0^2+\varepsilon\cos(\phi-\phi_0)}$$

  - This is a conic section!

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 $x_1, \dots, x_m \in V$ such that 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* $\mathbf{dim}\, \boldsymbol{V}$.

- We usually denoted a basis as an ordered n-tuple since the order often matters (for orientation?).
- Notational onventions.
  - 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 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 three (??) 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:
  $$\det\left(Q^{-1}AQ\right) = \det\left(Q^{-1}\right)\det(A)\det(Q) = \det(A)\det\left(Q^{-1}\right)\det(Q) = \det(A)$$

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* $\boldsymbol{J_d(\lambda)}$.
  - 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.
- We have that

$$
\begin{gathered}
J_d(\lambda)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}\lambda\\0\\\vdots\\0\end{pmatrix} = \lambda\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \lambda e_1\\
J_d(\lambda)\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}1\\\lambda\\0\\\vdots\\0\end{pmatrix} = e_1 + \lambda e_2\\
\vdots\\
J_d(\lambda)e_{d-1} = e_{d-2} + \lambda e_{d-1}
\end{gathered}
$$

- 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

$$
Q^{-1}AQ = \begin{pmatrix} J_{h_1}(\lambda_1) & & \\ & J_{h_{d_1}}(\lambda_1) & \\ & & \ddots \end{pmatrix}
$$

  - We have that $h_1 + \cdots + h_{d_1} = \alpha_1$??
  - 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: $K_i \cap K_j = \{0\}$. Let $j_i = \dim K_i$. Take the direct sum of all K_i. Then $j_1 + \cdots + j_m = n$.

(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

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

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

(b) 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 constitutes the Jordan basis.

(c) 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 $k_{i,1} + \cdots + k_{i,\gamma} = d_i$.

- **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.
- $\boldsymbol{j_\lambda}$: The dimension of the generalized eigenspace of λ. *Given by*

$$j_\lambda = \dim K_\lambda$$

- The JNF computation can be really heavy; we'll only ever compute 2×2 or 3×3 versions.
- Example:
  - 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

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

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

– Thus,

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

– 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 $Q = (v|u)$[1] 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 $\gamma_\lambda = 1$.

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

- In the first case, we have two eigenvectors v_1, v_2. We can find the third (generalized) eigenvector by solving $(A-\lambda I)u = v_1$ and $(A-\lambda I)u = v_2$ (only one of these will have a solution). WLOG let the first equation have a solution. Then $Q = (v_1|u|v_2)$ and the JNF is

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

- 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 $A^n = Q^{-1}\Lambda^n Q$.
- What if A is not diagonalizable?
  - Then we convert to A to Jordan normal form $A = Q^{-1}BQ$. 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 $i \geq m$).

$$[J_d(\lambda)]^m = \binom{m}{0}\lambda^m I_d + \binom{m}{1}\lambda^{m-1}N_d + \cdots + \binom{m}{m}N_d^m$$

- Example: When $d = 3$, then

$$\begin{pmatrix} \lambda & 1 & 0 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}^3 = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & m(m-1)\lambda^{m-2} \\ & \lambda^m & m\lambda^{m-1} \\ & & \lambda^m \end{pmatrix}$$

- 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:
$$f(z) = c_0 I + c_1 A + c_2 A^2 + \cdots$$
- **Radius of convergence**: The number R such that the series converges for $\|A\| < R$.
- **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 $A = Q^{-1}BQ$. Then
$$\begin{aligned} f(A) &= f(Q^{-1}BQ) \\ &= c_0 I + c_1(Q^{-1}BQ) + c_2(Q^{-1}BQ)^2 + \cdots \\ &= Q^{-1}(c_0 I + c_1 B + c_2 B^2 + \cdots)Q \\ &= Q^{-1} f(B) Q \end{aligned}$$
  - Going even further,
$$B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \implies f(B) = \begin{pmatrix} f(B_1) & 0 \\ 0 & f(B_2) \end{pmatrix}$$
  - In particular, if A is diagonalizable, then
$$f(A) = Q^{-1} \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} Q$$
- 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

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

- 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*

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

  - 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.

- 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, $y(t) = \mathrm{e}^{tA} y_0$.

- **Simplex**: A higher-dimensional generalization of a triangle.

- We now consider the inhomogeneous equation. Before, we used an integrating factor. We will now do that again.

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

  - 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: ??
- Rewrite $A = QBQ^{-1}$, where B is in JNF.
  - Shao reviews some facts of JNF from previous lectures.
- 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
$$\begin{aligned}\mathrm{e}^{tA}y_0 &= Q\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{e}^{t} & t\mathrm{e}^{t} \\ 0 & 0 & \mathrm{e}^{t} \end{pmatrix}Q^{-1}y_0 \\ &= \begin{pmatrix} \mathrm{e}^t - 3t\mathrm{e}^t & 2t\mathrm{e}^t & t\mathrm{e}^t \\ & \vdots & \\ & \vdots & \end{pmatrix}\begin{pmatrix} y_0^1 \\ y_0^2 \\ y_0^3 \end{pmatrix}\end{aligned}$$
- **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): The space of all (generalized) eigenvectors corresponding to the stable eigenvalues.
- **Unstable** (subspace of the system): The space of all (generalized) eigenvectors corresponding to the unstable eigenvalues.
- 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??
  - If $\sigma_j < 0$, then $|\mathrm{e}^{tA}P_jy_0| \to 0$ at an exponential rate.
- Similarly, if you're working with an unstable eigenvalue, then $\sigma_j > 0$ implies $|\mathrm{e}^{tA}P_jy_0| \to +\infty$ at an exponential rate.

  - The rate of growth depends on σ_j.
- Along the stable subspaces, your points will be attracted to zero.
- Along the unstable subspaces, your points will be repelled from zero.
- The stable subspace of our example is

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

- 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??
- Let $x(t)$ be a higher order scalar ODE.
  - Then we can make a system out of it:

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

  - $F[p]$ is the **Frobenius** matrix.
  - The transpose of this matrix is a very special matrix called the **companion** matrix $C[p] = F[p]^T$.
  - Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$. Then $\chi_{C[p]} = p(z)$.

    Proof. We have that

$$\begin{aligned}\chi_{C[p]}(z) &= \det(zI - C[p])\\ &= z(z^{n-1} + a_{n-1}(z^{n-2} + a_{n-2}(z^{n-3} + \cdots)))\\ &= p(z)\end{aligned}$$

    as desired. ☐
  - Roots of $p(z)$ are the eigenvalues of $F[p]$ and $C[p]$.
  - We have that $C[p]e_i = e_{i+1}$ for $i = 1, \dots, n-1$ and

$$C[p]e_n = -a_0 e_1 - \cdots - a_{n-1}e_n$$

    which implies that if $r(z)/\deg r < n$ nullifies $C[p]$, then necessarily $r(z) = p(z)$ since $(z-\lambda_j)^{<\alpha_j}$??
- 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.

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

  - 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 Zhao 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-Lindelof 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-Lindelof 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:

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

  - Zhao 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 $t^k\mathrm{e}^{t\lambda} \to 0$).