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.
- 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.

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

  - 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*

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

- **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 $y_0^1, \ldots, y_0^n$ (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 $e^{tA} = Qe^{tB}Q^{-1}$.
- 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

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

  - 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 $e^{tA}v$. The entries consist of linear combinations of functions of the form $t^k e^{t\lambda}$. If the real part is greater than zero, the solution grows exponentially fast in the t direction (notice how $t \to \infty$ implies $t^k e^{t\lambda} \to \infty$). Otherwise, the solution decays exponentially fast (notice how $t \to \infty$ implies $t^k e^{t\lambda} \to 0$).

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.

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

  - 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. $\square$

- 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!