Week 6

???

6.1 More Cauchy-Lipschitz and Intro to Continuous Dependence

10/31:

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

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

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

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

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

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

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

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

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

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

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

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

     - It follows inductively that

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

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

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

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

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

     - It follows by induction that

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

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

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

  Then

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

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

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

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

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

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

Proof. We have that

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

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

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

Combining what we've obtained, we have

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

as desired. ☐

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

6.2 Differentiability With Respect To Parameters

11/2:

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

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

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

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

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

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

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

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

- By Cauchy-Lipschitz, for small
$$T \leq \min\left(a, \frac{b}{M}, \frac{1}{2L}\right)$$
and $\mu \in B(0,r)$ (r small), the solution *exists* on $[t_0, t_0+T]$ and its value does not escape $\bar{B}(x(0), b+C)$.
  - We now aim to show that the solution is *differentiable* wrt. μ on this interval.
- If $y(t;\mu)$ satisfies $y'(t;\mu) = f(t, y(t;\mu);\mu)$ and if the Jacobian matrix $J = \partial y/\partial\mu$ exists, then J satisfies the **first variation equation**.
- **First variation equation**: The following linear differential equation. *Given by*
$$\frac{\mathrm{d}}{\mathrm{d}t} \underbrace{\frac{\partial y}{\partial \mu}(t;\mu)}_{J(t;\mu)} = \underbrace{\frac{\partial f}{\partial z}(t, y(t;\mu);\mu)}_{A(t;\mu)} \cdot \underbrace{\frac{\partial y}{\partial \mu}(t;\mu)}_{J(t;\mu)} + \frac{\partial f}{\partial \mu}(t, y(t;\mu);\mu), \quad \frac{\partial y}{\partial \mu}(t_0,\mu) = \frac{\partial x}{\partial \mu}(\mu)$$
- The first variation equation has a unique solution, but we do not yet know that $y(t;\mu)$ is even differentiable with respect to μ. We presently verify this claim.
- Theorem[1]: $y(t;\mu)$ is C^1 in μ and $\partial y/\partial\mu\,(t;\mu)$ satisfies the first variation equation.

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

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

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

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

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

  circle terms cancel. ☐

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

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

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

and the second order derivative must satisfy the second variational equation

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

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

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

of $y(t;\mu)$ about 0 gives an approximation of said function up to order k in μ.

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

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

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

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

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

- Simple example (more complex ones next lecture):

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

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

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

  - It follows by substituting into the above differential equation that

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

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

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

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

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

  - Where do the other initial conditions (all zero) come from??
- Therefore, our approximate solution is

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

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

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