Week 6

Qualitative Theory of ODEs

6.1 More Cauchy-Lipschitz and Intro to Continuous Dependence

10/31:

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

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

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

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

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

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

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

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

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

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

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

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

     - It follows inductively that

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

     - 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.

$$y_1(t) = 1 + \int_0^t (y_0(\tau))^2 \mathrm{d}\tau = 1 + t$$

$$y_2(t) = 1 + \int_0^t (1+\tau)^2 \mathrm{d}\tau = 1 + t + t^2 + \frac{t^3}{3}$$

$$y_3(t) = 1 + \int_0^t \left(1 + \tau + \tau^2 + \frac{\tau^3}{3}\right)^2 \mathrm{d}\tau = 1 + t + t^2 + t^3 + \frac{2t^4}{3} + \frac{t^5}{3} + \frac{t^6}{9} + \frac{t^7}{63}$$

     - 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 Shao is confusing the proof of this theorem with the proof of the Inverse Function Theorem).
  - The motivation for the last equality (the line above) is that if $F(x, f(x)) = 0$, then by the chain rule for partial derivatives,

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

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

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

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

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

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

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

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

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

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

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

circle terms cancel. □

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

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

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

  and the second order derivative must satisfy the second variational equation

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

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

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

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

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

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

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

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

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

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

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

  - It follows by substituting into the above differential equation that

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

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

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

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

- Solving, we get

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

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

- Therefore, our approximate solution is

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

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

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

6.3 Variational Examples

11/4: - We begin today with a more direct and less involved proof of the variation of parameters theorem.

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

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

Consider the Picard iteration sequence defined by

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

Differentiating we get

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

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

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

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

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

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

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

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

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

and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

    from which we can determine that

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

– Thus,

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

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

6.4 Chapter 2: Initial Value Problems

From Teschl (2012).

Section 2.4: Dependence on the Initial Condition

11/15:

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

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

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

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

for all $t \in [0, T]$.

If, in addition, $\alpha(s) \leq \alpha(t)$ for $s \leq t$, then

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

for all $t \in [0, T]$.

Proof. Let

$$\phi(t) := \exp\left(-\int_0^t \beta(s)\,\mathrm{d}s\right)$$

Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\phi(t)\int_0^t \beta(s)\psi(s)\,\mathrm{d}s\right) &= -\beta(t)\phi(t) \cdot \int_0^t \beta(s)\psi(s)\,\mathrm{d}s + \phi(t) \cdot \beta(t)\psi(t) \\ &= \beta(t)\phi(t)\left(\phi(t) - \int_0^t \beta(s)\phi(s)\,\mathrm{d}s\right) \\ &\leq \alpha(t)\beta(t)\phi(t)\end{aligned}$$

where the first equality holds by the product rule and the FTC, and the last inequality above holds by the first assumption in the statement of the lemma. Integrating the above inequality with respect to t and dividing the result by $\phi(t)$ shows that

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

It follows that

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

as desired.

The proof of the second claim is covered in Problem 2.11 (and is not applicable to course content). ☐

- A simple consequence of the generalized Grönwall's inequality.
  - If
    $$\psi(t) \leq \alpha + \int_0^t (\beta\psi(s) + \gamma)\,\mathrm{d}s$$
    for all $t \in [0,T]$, where $\alpha,\gamma \in \mathbb{R}$ and $\beta \geq 0$, then
    $$\psi(t) \leq \alpha\mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)$$
    for all $t \in [0,T]$.
  - See Problem 2.12 for the proof.
- We can now show that the IVP is well-posed.
- Theorem 2.8: Suppose $f,g \in C(U,\mathbb{R}^n)$ and let f be locally Lipschitz continuous in the second argument, uniformly with respect to the first. If $x(t),y(t)$ are respective solutions of the IVPs

$$\begin{aligned} \dot{x} &= f(t,x) \\ x(t_0) &= x_0 \end{aligned} \qquad\qquad \begin{aligned} \dot{y} &= g(t,y) \\ y(t_0) &= y_0 \end{aligned}$$

then

$$|x(t) - y(t)| \leq |x_0 - y_0|\mathrm{e}^{L|t-t_0|} + \frac{M}{L}(\mathrm{e}^{L|t-t_0|} - 1)$$

where L is the Lipschitz constant of $f : V \to \mathbb{R}^n$, $M = \|f - g\|$ for $f,g : V \to \mathbb{R}^n$, and $V \subset U$ contains $G(x),G(y)$.

Proof. WLOG let $t_0 = 0$. Then

$$\begin{aligned} |x(t) - y(t)| &\leq |x_0 - y_0| + \int_0^t |f(s,x(s)) - g(s,y(s))|\,\mathrm{d}s \\ &\leq |x_0 - y_0| + \int_0^t (L|x(s) - y(s)| + M)\,\mathrm{d}s \end{aligned}$$

Thus, taking

$$\underbrace{|x(t) - y(t)|}_{\psi(t)} \leq \underbrace{|x_0 - y_0|}_{\alpha} + \int_0^t (\underbrace{L}_{\beta}\underbrace{|x(s) - y(s)|}_{\phi(s)} + \underbrace{M}_{\gamma})\,\mathrm{d}s$$

we have by the above consequence of the generalized Grönwall's inequality that

$$|x(t) - y(t)| \leq |x_0 - y_0|\mathrm{e}^{Lt} + \frac{M}{L}(\mathrm{e}^{Lt} - 1)$$

as desired. ☐

- Establishing continuous dependence on the initial condition.
  - Denote the solution of the IVP by $\phi(t,t_0,x_0)$ to emphasize the dependence on the initial condition.
  - Then in the special case $f=g$ (i.e., where $M=0$), Theorem 2.8 implies that

$$|\phi(t,t_0,x_0)-\phi(t,t_0,y_0)|\leq|x_0-y_0|\mathrm{e}^{L|t-t_0|}$$

  - In other words, ϕ depends continuously on the initial value.
  - Of course, this bound blows up exponentially as t increases, but the linear equation $\dot{x}=x$ shows that we cannot define a better bound in general.
- We now formalize the above notion.
- Theorem 2.9: Suppose $f\in C(U,\mathbb{R}^n)$ is locally Lipschitz continuous in the second argument, uniformly with respect to the first. Around each point $(t_0,x_0)\in U$, we can find a compact set $I\times B\subset U$ such that $\phi(t,s,x)\in C(I\times I\times B,\mathbb{R}^n)$. Moreover, $\phi(t,t_0,x_0)$ is Lipschitz continuous with

$$|\phi(t,t_0,x_0)-\phi(s,s_0,y_0)|\leq|x_0-y_0|\mathrm{e}^{L|t-t_0|}+(|t-s|+|t_0-s_0|\mathrm{e}^{L|t-s_0|})M$$

where L is the Lipschitz constant of $f:V\to\mathbb{R}^n$, $M=\|f\|$ for $f:V\to\mathbb{R}^n$, and $V\subset U$ compact contains $I\times\phi(I\times I\times B)$.

Proof. By the Picard-Lindelöf theorem, there exists $V=[t_0-\varepsilon,t_0+\varepsilon]\times\overline{B_\delta(x_0)}$ such that $\phi(t,t_0,x_0)$ exists and is continuous for $|t-t_0|<\varepsilon$. It can be shown that $\phi(t,t_1,x_1)$ exists for $|t-t_1|\leq\varepsilon/2$, provided that $|t_1-t_0|\leq\varepsilon/2$ and $|x_1-x_0|\leq\delta/2$. Thus, choose $I=[t_0-\varepsilon/2,t_0+\varepsilon/2]$ and $B=\overline{B_{\delta/2}(x_0)}$.

Moreover, we have that

$$\begin{aligned}|\phi(t,t_0,x_0)-\phi(s,s_0,y_0)|&\leq|\phi(t,t_0,x_0)-\phi(t,t_0,y_0)|\\&\quad+|\phi(t,t_0,y_0)-\phi(t,s_0,y_0)|\\&\quad+|\phi(t,s_0,y_0)-\phi(s,s_0,y_0)|\\&\leq|x_0-y_0|\mathrm{e}^{L|t-t_0|}\\&\quad+\left|\int_{t_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\&\quad+\left|\int_s^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\end{aligned}$$

We estimated the first term using the note on continuous dependence directly preceding this theorem and proof. We can estimate the third term to be $M|t-s|$. We can estimate the second term as follows: Abbreviating $\Delta(t):=|\phi(t,t_0,y_0)-\phi(t,s_0,y_0)|$ and assuming WLOG that $t_0\leq s_0\leq t$, we have that

$$\begin{aligned}\Delta(t)&=\left|\int_{t_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\&=\left|\int_{t_0}^{s_0} f(r,\phi(r,t_0,y_0))\,\mathrm{d}r+\int_{s_0}^t f(r,\phi(r,t_0,y_0))\,\mathrm{d}r-\int_{s_0}^t f(r,\phi(r,s_0,y_0))\,\mathrm{d}r\right|\\&\leq\int_{t_0}^{s_0}|f(r,\phi(r,t_0,y_0))|\,\mathrm{d}r+\int_{s_0}^t|f(r,\phi(r,t_0,y_0))-f(r,\phi(r,s_0,y_0))|\,\mathrm{d}r\\&\leq|t_0-s_0|M+L\int_{s_0}^t\Delta(r)\,\mathrm{d}r\end{aligned}$$

Therefore, by Grönwall's inequality (as in Problem 2.12; note that $\gamma=0$ here),

$$\Delta(t)\leq|t_0-s_0|M\mathrm{e}^{L|t-s_0|}$$

as desired. ∎

- By Problem 1.8, we have $\phi(t,t_0,x_0)=\phi(t-t_0,0,x_0)$ for an autonomous system, so we may consider $\phi(t,x_0)=\phi(t,0,x_0)$.
- The previous result of *continuous* dependence on initial conditions is not good enough in every situation; sometimes, we need to be able to *differentiate* with respect to the initial condition.

11/23:

- Before we prove that a certain set of conditions will imply the existence of the derivative of a solution ϕ with respect to x, we will assume such a derivative exists and investigate some of its properties (so as to motivate the following theorem and proof).
  - Suppose $\phi(t,t_0,x)$ is differentiable with respect to x.
  - If we assume $f\in C^k(U,\mathbb{R}^n)$ for some $k\geq 1$ and ϕ is sufficiently regular that its partial derivatives commute as well, then we may write[2]

$$\begin{aligned}\dot{\phi}(t,t_0,x)&=f(t,\phi(t,t_0,x))\\ \frac{\partial\phi}{\partial t}&=f(t,\phi)\\ \frac{\partial^2\phi}{\partial x\partial t}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\\ \frac{\partial^2\phi}{\partial t\partial x}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\\ \frac{\partial}{\partial t}\frac{\partial\phi}{\partial x}&=\frac{\partial f}{\partial\phi}\frac{\partial\phi}{\partial x}\end{aligned}$$

  - Thus, if $\partial\phi/\partial x$ exists, then it satisfies the **first variational equation** with $A(t,x)=\partial f/\partial\phi$.
  - It also necessarily satisfies the corresponding integral equation

$$\underbrace{\frac{\partial\phi}{\partial x}(t,t_0,x)}_{y(t)}=\underbrace{\frac{\partial\phi}{\partial x}(t_0,t_0,x)}_{\mathbb{I}}+\int_{t_0}^{t}\underbrace{\frac{\partial f}{\partial\phi}(s,\phi(s,t_0,x))}_{A(s,x)}\cdot\underbrace{\frac{\partial\phi}{\partial x}(s,t_0,x)}_{y(s)}\,\mathrm{d}s$$

- **First variational equation**: The following differential equation. *Given by*

$$\dot{y}=A(t,x)y$$

  - Which is the correct first variational equation?? This one, or the one with $F(t;\mu)$ term?
- Indeed, we have shown that $\partial\phi/\partial x$ necessarily satisfies the above integral equation if it exists. But using fixed point techniques, we can further show that said equation *has* a solution (which is $\partial\phi/\partial x$) under certain conditions. We will not go into this argument in depth, but it will be used in our next theorem.
- Theorem 2.10: Suppose $f\in C^k(U,\mathbb{R}^n)$, $k\geq 1$. Around each point $(t_0,x_0)\in U$, we can find an open set $I\times B\subset U$ such that $\phi(t,s,x)\in C^k(I\times I\times B,\mathbb{R}^n)$. Moreover, $\partial\phi/\partial t\,(t,s,x)\in C^k(I\times I\times B,\mathbb{R}^n)$ and if D_k is a partial derivative of order k, then $D_k\phi$ satisfies the higher order variational equation obtained from

$$\frac{\partial}{\partial t}D_k\phi(t,s,x)=D_k\frac{\partial}{\partial t}\phi(t,s,x)=D_kf(t,\phi(t,s,x))$$

  by applying the chain rule repeatedly. In particular, this equation is linear in $D_k\phi$ and it also follows that the corresponding higher-order derivatives commute.

  Proof. To prove the claim, we induct on k. For the base case $k=1$, we need only prove that $\phi(t,x)$ is *differentiable* at every $x_1\in B$. We now define some terms that will be useful in our argument.

[2]Note that dot-denoted derivatives refer to derivatives with respect to t.

Using the trick on Teschl (2012, p. 7), add t to the dependent variables to make the ODE autonomous. This allows us to consider $\phi(t,x)=\phi(t,0,x)$ WLOG. Thus, we have by Theorem 2.9 that there exists a set $I\times B\subset U$ such that $\phi(t,x_0)\in C(I\times B,\mathbb{R}^n)$. We may take $I=(-T,T)$ and $B=B_\delta(x_0)$ for some $T,\delta>0$ such that $\overline{I\times B}\subset U$. Additionally, let $x_1=0$ WLOG. Furthermore, let $\phi(t):=\phi(t,x_1)$, $A(t):=A(t,x_1)$, and $\psi(t)$ denote the solution to the first variational equation $\dot{\psi}(t)=A(t)\psi(t)$ corresponding to the initial condition $\psi(t_0)=\mathbb{I}$ ($\psi(t)$ is guaranteed to exist by the aforementioned fixed point argument). Lastly, let

$$\theta(t,x):=\frac{\phi(t,x)-\phi(t)-\psi(t)x}{|x|}$$

The proof strategy is thus: If we can show that $\lim_{x\to x_1=0}\theta(t,x)=0$, then we will have proven that $\partial\phi/\partial x$ exists and is equal to ψ. To do so, we will derive a bound on $|\theta(t,x)|$ that we can more easily control. Let's begin.

Since $f\in C^1$, we have that

$$\begin{aligned}
f(y)-f(x)&=\int_x^y\frac{\partial f}{\partial x}(t)\,\mathrm{d}t\\
\frac{f(y)-f(x)}{y-x}&=\frac{1}{y-x}\int_x^y\frac{\partial f}{\partial x}(t)\,\mathrm{d}t\\
&=\int_0^1\frac{\partial f}{\partial x}(x+t(y-x))\,\mathrm{d}t\\
&=\frac{\partial f}{\partial x}(x)+\int_0^1\left(\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right)\mathrm{d}t\\
f(y)-f(x)&=\frac{\partial f}{\partial x}(x)(y-x)+|y-x|R(y,x)
\end{aligned}$$

where

$$|R(y-x)|\leq\max_{t\in[0,1]}\left\|\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right\|$$

Let's take another minute to justify the above algebraic manipulations from a geometric perspective. The first line should be a fairly straightforward application of the FTC. The second line rephrases the equality as two different ways of calculating the average slope of f on $[x,y]$; specifically, we may either do rise over run (LHS) or find the area under the derivative and divide by the "length" to get the average "height." The third line shrinks the domain of integration to an interval of unit length, compressing all of the information in $\partial f/\partial x$ along with it, and making it so that we no longer have to divide through by the "length." The fourth line adds and subtracts the starting point so that geometrically, we take the area in two parts: All of the area beneath the first point, and then all of the area that "changes." The last line sees us multiply both sides by $y-x$ and then take the integral to be a sort of "remainder." Alternatively, we may justify

$$f(y)=f(x)+\frac{\partial f}{\partial x}(x)(y-x)+\left(\int_0^1\left(\frac{\partial f}{\partial x}(x+t(y-x))-\frac{\partial f}{\partial x}(x)\right)\mathrm{d}t\right)(y-x)$$

by analogy: If $f(x)=x^2$ and thus $\mathrm{d}f/\mathrm{d}x=2x$, then 4^2 equals 2^2, plus the area under the curve $2x$ from $x=2$ to $x=4$ partitioned into the rectangle with base length $y-x=2$ and height $\mathrm{d}f/\mathrm{d}x\,(2)=4$ and the triangle sitting on top of the aforementioned rectangle and below the graph of $\mathrm{d}f/\mathrm{d}x$. Note that all of these justifications are taken from the perspective of f being a single-variable function; to justify the transformations in the multivariable case would likely be much more complicated and is beyond my grasp at the moment (recall that $\partial f/\partial x$ is actually a matrix!). Also note that as such, the norm in the bound on $|R(y,x)|$ given above is the matrix norm.

Since $f\in C^1$, its partial derivatives are uniformly continuous in a neighborhood of x_1. It follows that $\lim_{y\to x}|R(y,x)|=0$ where the argument converges uniformly in x in some neighborhood of $x_1=0$.

Using the above expression for $f(y)-f(x)$, we have that

$$
\begin{aligned}
\dot{\theta}(t,x) &= \frac{1}{|x|}[\dot{\phi}(t,x)-\dot{\phi}(t)-\dot{\psi}(t)x]\\
&= \frac{1}{|x|}[f(\phi(t,x))-f(\phi(t))-A(t)\psi(t)x]\\
&= \frac{1}{|x|}\left[\frac{\partial f}{\partial x}(\phi(t))(\phi(t,x)-\phi(t))+|\phi(t,x)-\phi(t)|R(\phi(t,x),\phi(t))-A(t)\psi(t)x\right]\\
&= A(t)\cdot\frac{\phi(t,x)-\phi(t)-\psi(t)x}{|x|}+\frac{|\phi(t,x)-\phi(t)|}{|x|}R(\phi(t,x),\phi(t))\\
&= A(t)\theta(t,x)+\frac{|\phi(t,x)-\phi(t)|}{|x|}R(\phi(t,x),\phi(t))
\end{aligned}
$$

Integrating and taking absolute values yields

$$
\begin{aligned}
|\theta(t,x)| &= |\theta(t,x)-\theta(0,x)|\\
&= \left|\int_0^t\left(A(s)\theta(s,x)+\frac{|\phi(s,x)-\phi(s)|}{|x|}R(\phi(s,x),\phi(s))\right)\mathrm{d}s\right|\\
&\leq \int_0^t\frac{1}{|x|}|\phi(s,x)-\phi(s,0)|\cdot|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s\\
&\leq \int_0^t\frac{1}{|x|}|x-0|\mathrm{e}^{L|s-0|}\cdot|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s\\
&\leq \mathrm{e}^{LT}\int_0^T|R(\phi(s,x),\phi(s))|\,\mathrm{d}s+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s
\end{aligned}
$$

Note that $\theta(0,x)=0$ since we have taken $t_0=x_1=0$ WLOG. Also note that we use the continuous dependence on initial conditions equation to get from line 2 to line 3. Lastly, note that from the next to the last to the last line, we use the inequality $\mathrm{e}^{Ls}\leq\mathrm{e}^{LT}$ to transform the exponential function into a constant that bounds it (recall that $s\leq T$ by definition).

Define

$$\tilde{R}(x)=\mathrm{e}^{LT}\int_0^T|R(\phi(s,x),\phi(s))|\,\mathrm{d}s$$

so that

$$|\theta(t,x)|\leq\tilde{R}(x)+\int_0^t\|A(s)\||\theta(s,x)|\,\mathrm{d}s$$

It follows by the Generalized Grönwall's inequality that

$$|\theta(t,x)|\leq\tilde{R}(x)\exp\left(\int_0^T\|A(s)\|\,\mathrm{d}s\right)$$

This combined with the fact from earlier that $\lim_{y\to x}|R(y,x)|=0$ and hence $\lim_{x\to 0}\tilde{R}(x)=0$ implies that $\lim_{x\to 0}\theta(t,x)=0$, as desired.

Additionally, $\partial\phi/\partial x$ is continuous as the solution to the first variational equation. This completes the base case.

Now suppose via strong induction that the claim holds for $1,\dots,k$, and let $f\in C^{k+1}$. Then $\phi(t,x)\in C^1$ and $\partial\phi/\partial x$ solves the first variational equation, as per the base case. But since $A(t,x)\in C^k$ and hence $\partial\phi/\partial x\in C^k$, we have by Lemma 2.3 that $\phi(t,x)\in C^{k+1}$. $\square$

- This theorem also allows us to handle dependence on parameters.

- In particular, if f depends on some paramenters $\lambda \in \Lambda \subset \mathbb{R}^p$ such that we are now solving the IVP

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

for a solution $\phi(t, t_0, x_0, \lambda)$, then we have the following result.

- Theorem 2.11: Suppose $f \in C^k(U \times \Lambda, \mathbb{R}^n)$, $k \geq 1$. Around each point $(t_0, x_0, \lambda_0) \in U \times \Lambda$ we can find an open set $I \times B \times \Lambda_0 \subset U \times \Lambda$ such that $\phi(t, s, x, \lambda) \in C^k(I \times I \times B \times \Lambda_0, \mathbb{R}^n)$.

Proof. Largely follows from Theorem 2.10. Noteworthy modificatinos: We add the parameters λ to the dependent variables and require $\dot{\lambda} = 0$ (i.e., that the parameters do not change with time and therefore uniquely determine a solution over time). □

Problems

11/15: **2.12.** Show that if

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

for all $t \in [0, T]$, where $\alpha, \gamma \in \mathbb{R}$ and $\beta \geq 0$, then

$$\psi(t) \leq \alpha \mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)$$

for all $t \in [0, T]$. *Hint*: Introduce $\tilde{\psi}(t) = \psi(t) + \gamma/\beta$.

Proof. Taking the hint and substituting $\psi(t) = \tilde{\psi}(t) - \gamma/\beta$, we get

$$\begin{aligned}
\tilde{\psi}(t) - \frac{\gamma}{\beta} &\leq \alpha + \int_0^t \left(\beta\left(\tilde{\psi}(t) - \frac{\gamma}{\beta}\right) + \gamma\right) \mathrm{d}s \\
\tilde{\psi}(t) &\leq \alpha + \frac{\gamma}{\beta} + \int_0^t \beta\tilde{\psi}(t) \, \mathrm{d}s
\end{aligned}$$

It follows by the generalized Grönwall's inequality that

$$\begin{aligned}
\tilde{\psi}(t) &\leq \alpha + \frac{\gamma}{\beta} + \int_0^t \left(\alpha + \frac{\gamma}{\beta}\right) \beta \exp\left(\int_s^t \beta \, \mathrm{d}r\right) \mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + \int_0^t (\alpha\beta + \gamma)\mathrm{e}^{\beta(t-s)} \, \mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + (\alpha\beta + \gamma)\mathrm{e}^{\beta t} \int_0^t \mathrm{e}^{-\beta s} \, \mathrm{d}s \\
&= \alpha + \frac{\gamma}{\beta} + (\alpha\beta + \gamma)\mathrm{e}^{\beta t} \left(\frac{1}{-\beta}\mathrm{e}^{-\beta t} - \frac{1}{-\beta}\mathrm{e}^0\right) \\
&= \alpha + \frac{\gamma}{\beta} + \left(\alpha + \frac{\gamma}{\beta}\right) \mathrm{e}^{\beta t}(1 - \mathrm{e}^{-\beta t}) \\
&= \alpha + \frac{\gamma}{\beta} + \left(\alpha + \frac{\gamma}{\beta}\right) (\mathrm{e}^{\beta t} - 1) \\
&= \alpha + \frac{\gamma}{\beta} + \alpha\mathrm{e}^{\beta t} - \alpha + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1) \\
&= \frac{\gamma}{\beta} + \alpha\mathrm{e}^{\beta t} + \frac{\gamma}{\beta}(\mathrm{e}^{\beta t} - 1)
\end{aligned}$$

Subtracting γ/β from both sides and returning the substitution yields the desired result. □