Differential Equations 2020/21
MA 209

Extra Notes 2

Continuity of Functions
Existence and Uniqueness of Solutions of ODEs
The Lipschitz Conditions

These extra notes provide some extra information related to Chapter 4 of the Lecture Notes.

2.1 Functions and their limits

Some important general definitions for functions are:

* If $f : S \longrightarrow T$, where S and T are non-empty sets, then S is called the *domain of* f and T is the *range of* f.
* For $R \subseteq S$, the *set of attainable values of* f *on* R or the *image of* R *under* f is

$$f(R) \,=\, \{\, y \in T \mid f(x) = y \text{ for some } x \in R \,\} \,=\, \{\, f(x) \mid x \in R \,\}.$$

- Most of you will have a notion of what it means for a function from $\mathbb{R}$ to $\mathbb{R}$ to have a limit. Definitions for these concepts usually include notions such as "x approaches a" or "x approaches a from above". But if we are dealing with functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we must be careful what we mean if we say "x approaches a" since x and a are points in some higher dimensional space. We will use the following definition:

  Definition
  Given a function $f : S \longrightarrow \mathbb{R}^m$ where $S \subseteq \mathbb{R}^n$. Then we say that $f(x) \longrightarrow y$ as $x \longrightarrow a$, where $y \in \mathbb{R}^m$ and $a \in \mathbb{R}^n$, if for every $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x \in S$ with $0 < \|x - a\| < \delta$ we have $\|f(x) - y\| < \varepsilon$.

  Here $\|\,.\,\|$ is the normal *norm* in $\mathbb{R}^n$: if $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, then $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

- For one-dimensional functions we have an additional definition:

  Definition
  Given a function $f : \mathbb{R} \longrightarrow \mathbb{R}^m$. Then we say that $f(t) \longrightarrow y$ as $t \longrightarrow \infty$, where $y \in \mathbb{R}^m$, if for all $\varepsilon > 0$ there exists an $M \in \mathbb{R}$ such $\|f(t) - y\| < \varepsilon$ for all $t > M$.

- For any $x \in \mathbb{R}^n$ and real number $r > 0$, the *open ball $B(x,r)$ with centre x and radius r* is the set $B(x,r) = \{\, y \in \mathbb{R}^n \mid \|x - y\| < r \,\}$.

  If $D \subseteq \mathbb{R}^n$, then a point $x \in D$ is an *interior point of* D if there is an $r > 0$ so that $B(x,r) \subseteq D$.

2.2 Continuous functions

We use the following definitions for a function to be continuous:

* A function $f : S \longrightarrow \mathbb{R}^m$, where $S \subseteq \mathbb{R}^n$, is said to be *continuous at $x_0 \in S$* if $f(x) \longrightarrow f(x_0)$ as $x \longrightarrow x_0$.
* And f is *continuous on S* if f is continuous at every point in S.

Using the definition from the previous subsection, a more extended definition would be:

* A function $f : S \longrightarrow \mathbb{R}^m$, where $S \subseteq \mathbb{R}^n$, is said to be *continuous at $x_0 \in S$* if for all $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x \in S$ with $0 < \|x - x_0\| < \delta$ we have $\|f(x) - f(x_0)\| < \varepsilon$.

- If $f : S \longrightarrow \mathbb{R}^m$, then we can think of f as being defined by m functions $f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$,

  one for each coordinate. It can be shown that this means that:

  * *f is continuous at $x_0 \in S$ if and only if each f_i is continuous at $x_0 \in S$.*

2.3 Existence of Solutions

For the remainder of these notes we suppose that we are given a function $f : D \longrightarrow \mathbb{R}^n$ for some $D \subseteq \mathbb{R}^{n+1}$ and a point $(t_0, x_0) \in D$. And we are looking for solutions $x(t)$ to the following initial-value problem

$$(1) \qquad x' = f(t,x) \qquad \text{with} \qquad x(t_0) = x_0.$$

(The reason to allow f to be defined on a subset D of $\mathbb{R}^{n+1}$, is so that we also can consider equations like $x' = x/t$ for $t > 0$.)

- Following the definitions and observations from the previous subsection, we can write

$$f(t,x) = \begin{bmatrix} f_1(t, x_1, x_2, \ldots, x_n) \\ f_2(t, x_1, x_2, \ldots, x_n) \\ \vdots \\ f_n(t, x_1, x_2, \ldots, x_n) \end{bmatrix},$$

and f is continuous on D if and only if each $f_i(t, x_1, \ldots, x_n)$ is continuous on D.

- The following theorem guarantees that most initial-value problems have a solution.

  Theorem 2.1
  If $f(t,x)$ is continuous on D and (t_0,x_0) is an interior point of D, then there exist t_s,t_e with $t_s < t_0 < t_e$ and a function $x : (t_s,t_e) \longrightarrow \mathbb{R}^n$ that solves the initial-value problem (1) for all $t \in (t_s,t_e)$.

  The above is equivalent to Theorem 4.5.1 in the Lecture Notes (although formulated a little different). It is a lot more general than Theorem 4.1.1 in the Lecture Notes.

  The proof of Theorem 2.1 is fairly tricky. So we only give a sketch of a possible way to prove the theorem. This in fact describes a way to find an approximate solution. The construction is known as the *Cauchy-Euler construction*.

  Sketch of proof We will only show that there is a solution on the interval $[t_0,t_e)$ for some $t_e > t_0$. The same ideas can be used to show that there is a solution for t below t_0 as well.

  We will assume that all points considered below are in D. This can be achieved by choosing the value of α below appropriately.

  Fix a positive number α and set $t_\alpha = t_0 + \alpha$. We will construct an approximate solution on $[t_0,t_\alpha]$. Next, for a positive integer N, divide the interval between t_0 and t_α into N equal parts. So write $\Delta_t = \alpha/N$ and define $N+1$ time points $t_r = t_0 + r\Delta_t$ for $r = 0,1,\ldots,N$. We now form the corresponding sequence of points x_r, $r = 0,1,\ldots,N$, defined by

  $$x_r \;=\; x_{r-1} + (t_r - t_{r-1})\, f(t_{r-1},x_{r-1}) \qquad \text{for } r = 1,\ldots,N.$$

  Finally, for a time $t \in [t_0,t_\alpha)$ we know that $t \in [t_{r-1},t_r)$, for exactly one $r \in \{1,2,\ldots,N\}$. So for all $t \in [t_0,t_\alpha)$ we can define the approximate solution $x^N(t)$ as follows:

  $$x^N(t) \;=\; x_{r-1} + (t - t_{r-1})\, f(t_{r-1},x_{r-1}) \qquad \text{where } r \text{ is the integer so that } t \in [t_{r-1},t_r).$$

  In order to understand what is happening, make the following observations about x^N:

  - For $t \in [t_0,t_1)$, $x^N(t)$ is nothing more than the line starting in x_0 going in the direction $f(t_0,x_0)$. Note that the solution of the ODE is a function $x(t)$ with $x(t_0) = x_0$ and $x'(t_0) = f(t_0,x_0)$. So for $t \in [t_0,t_1)$, $x^N(t)$ is the linear approximation of the solution x we are looking for.
  - The linear approximation from the first step goes from $t = t_0$ until $t = t_1$. Then we are in the point x_1 and we start using a new direction $f(t_1,x_1)$. So for $t \in [t_1,t_2)$, $x^N(t)$ is the linear approximation of a solution x that would start with $x(t_1) = x_1$.
  - The process continues; between times t_{r-1} and t_r we follow a straight line starting at x_{r-1} and with direction $f(t_{r-1},x_{r-1})$.

  So the function $x^N(t)$ consists of a sequence of linear pieces, each piece chosen to give a reasonable approximation of a possible solution of the ODE at the starting point of the piece. This approach is often used in computer software to find approximations of solutions, or for instance to draw the graphs of solutions when the solution itself is not explicitly known.

  From this point, we should continue the proof by showing that for α small enough, if we let $N \longrightarrow \infty$ (which is the same as $\Delta_t \downarrow 0$), then x^N converges to some differentiable function x

which is the solution of (1). This part of the proof involves analysis of uniform convergence of $x^N(t)$, etc. We will skip that and just believe that by doing the construction above using smaller and smaller steps Δ_t we eventually get a solution. ■

- Note that Theorem 2.1 only guarantees a solution over a certain interval $I = (t_s, t_e)$ with $t_0 \in I$. This interval very much depends on the exact form of the ODE. For instance the one-dimensional ODE

$$x' = x^2 - 1, \qquad x(0) = 0,$$

has the solution $x(t) = \dfrac{1 - e^{2t}}{1 + e^{2t}}$, for all $t \in \mathbb{R}$ (so we can take $I = \mathbb{R}$). But the very similar looking ODE

$$y' = y^2 + 1, \qquad y(0) = 0,$$

has the solution $y(t) = \tan(t)$, for $-\frac{1}{2}\pi < t < \frac{1}{2}\pi$. So here the solution is only valid for the interval $I = (-\frac{1}{2}\pi, \frac{1}{2}\pi)$.

- Theorem 2.1 guarantees that most ODEs have a solution. But that doesn't mean the solution has to be unique. For instance, the ODE

$$x' = \tfrac{3}{2}\sqrt[3]{x}, \qquad x(0) = 0,$$

has solutions $x(t) = 0$ for all $t \in \mathbb{R}$, but also $x(t) = |t|^{3/2}$ for all $t \in \mathbb{R}$, and many others.

So in order to make sure that there is a unique solution, we must put some extra conditions on the ODE, in particular on the expression $f(t, x)$.

2.4 Uniqueness of Solutions - The Lipschitz Condition

The different forms of the Lipschitz Conditions (*locally Lipschitz*, *globally Lipschitz*, *locally Lipschitz in x uniform with respect to t*) are defined in Section 4.4 of the Lecture Notes. The most important one is the following.

* **Definition**
Let $D \subseteq \mathbb{R}^{n+1}$ be some domain in which (t_0, x_0) is an interior point. Then the function $f(t, x)$ defined on D *satisfies the Lipschitz condition on D* if there exists a constant L such that

$$\|f(t, x) - f(t, y)\| \leq L\|x - y\| \qquad \text{for all } (t, x), (t, y) \in D.$$

With this condition, we can formulate the most important general result on the uniqueness of ODEs:

* **Theorem 2.2**
Let $f(t, x)$ satisfy the Lipschitz Condition on some domain D in which (t_0, x_0) is an interior point. Then there exist t_s, t_e with $t_s < t_0 < t_e$ such that the differential equation $x' = f(t, x)$, with initial value $x(t_0) = x_0$, has a unique solution on (t_s, t_e).

To prove this theorem, we first would need to show that there is at least one solution. But that follows from the earlier results, since it can be shown that if a function satisfies a Lipschitz Condition, then it is continues. So we only need to prove that that solution is unique.

That second part of the proof can be found in Section 4.3.2 of the Lecture Notes, but also at the end of these extra notes. It is quite some work, although nothing incredibly complicated is happening. Nevertheless, we won't spend much time with the proof, and hence it is not considered examinable material.

- Although the Lipschitz Condition is not too complicated, in practice it is quite hard to find out if a function satisfies the Lipschitz Condition. The following condition is often useful.

  * **Theorem 2.3**
  Suppose $f(t,x)$ is continuous differentiable on some open convex domain $D \subseteq \mathbb{R}^{n+1}$ with $(t_0,x_0) \in D$ and that there exists some constant K such that the partial derivatives with respect to the x-coordinates satisfy:

  $$\left|\frac{\partial f_i(t,x)}{\partial x_j}\right| \le K \qquad \text{for all } i=1,\dots,n,\ j=1,\dots,n \text{ and } (t,x)\in D.$$

  Then f satisfies the Lipschitz Condition on D.

2.5 Proof of Uniqueness under the Lipschitz Condition

Before really starting with the proof, we first rewrite the standard ODE from (1) in a somewhat different form. To find this, suppose we have a solution $x(t)$ for (1). Integrating both sides of the differential equation from t_0 to t we get

$$(2) \qquad \int_{t_0}^{t} x'(\tau)\,\mathrm{d}\tau = \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau.$$

You must realise that the functions x and f are in fact multi-dimensional. So in reality we have $x(s) = \big[x_1(s),\dots,x_n(s)\big]$, and hence we should read

$$\int_{t_0}^{t} \dot{x}(\tau)\,\mathrm{d}\tau = \int_{t_0}^{t} \big[x_1(\tau),\dots,x_n(\tau)\big]\,\mathrm{d}\tau = \left[\int_{t_0}^{t} x_1(\tau)\,\mathrm{d}\tau,\dots,\int_{t_0}^{t} x_n(\tau)\,\mathrm{d}\tau\right],$$

where each of the integrals $\int_{t_0}^{t} x_i(\tau)\,\mathrm{d}\tau$ is a normal, one-dimensional integral.

Now recall that for a differentiable function $f:\mathbb{R}\longrightarrow\mathbb{R}$ we have $\int_{t_0}^{t} f'(\tau)\,\mathrm{d}\tau = f(t)-f(t_0)$.

Then we find that the equation in (2) is equivalent to $x(t)-x(t_0) = \int_{t_0}^{t} f(\tau,x(\tau))\,\mathrm{d}\tau$. Entering the initial value $x(t_0)=x_0$, we get the so-called *Volterra's Integral Equation*:

$$(3) \qquad x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau.$$

In fact, we have shown that solving the initial-value differential equation in (1) is equivalent to solving the integral equation in (3).

- We next need a preliminary lemma.

  Lemma 2.4
  Let $\varphi : [t_0, t_e) \longrightarrow \mathbb{R}$, where $t_e > t_0$, be continuous on $[t_0, t_e)$ and satisfy $\varphi(t) \geq 0$ for all $t \in [t_0, t_e)$. Suppose there is some constant $K \geq 0$ so that

  $$0 \leq \varphi(t) \leq K \int_{t_0}^{t} \varphi(\tau)\,\mathrm{d}\tau \qquad \text{for all } t \in [t_0, t_e).$$

  Then $\varphi(t) = 0$ for all $t \in [t_0, t_e)$.

  Proof If $K = 0$, then we immediately get $0 \leq \varphi(t) \leq 0$, hence $\varphi(t) = 0$ for all $t \in [t_0, t_e)$. So from now on we assume $K > 0$.

  For $t \in [t_0, t_e)$ write $\Phi(t) = \int_{t_0}^{t} \varphi(\tau)\,\mathrm{d}\tau$. Since $\varphi(\tau) \geq 0$ for al $\tau \in [t_0, t_e)$, we also have $\Phi(t) \geq 0$. Also, $\Phi(t_0) = 0$ and $\Phi(t)$ is continuous differentiable with $\Phi'(t) = \varphi(t)$. Hence the inequality in the lemma can be written as

  $$0 \leq \Phi'(t) \leq K \cdot \Phi(t) \qquad \text{for all } t \in [t_0, t_e).$$

  The second part is the same as $\Phi'(t) - K\Phi(t) \leq 0$. After multiplying with the positive value e^{-Kt} we get $e^{-Kt}\Phi'(t) - e^{-Kt}K\Phi(t) \leq 0$, which is the same as $\frac{\mathrm{d}}{\mathrm{d}t}\big[e^{-Kt}\Phi(t)\big] \leq 0$. Now take the integral from t_0 to t on both sides to get:

  $$e^{-Kt}\Phi(t) - e^{-Kt_0}\Phi(t_0) = \int_{t_0}^{t} \frac{\mathrm{d}}{\mathrm{d}t}\big[e^{-K\tau}\Phi(\tau)\big]\,\mathrm{d}\tau \leq \int_{t_0}^{t} 0\,\mathrm{d}\tau = 0.$$

  (Recall that for a differentiable function ψ we have $\int_{a}^{b} \psi'(\tau)\,\mathrm{d}\tau = \psi(b) - \psi(a)$.) But since $\Phi(t_0) = 0$ we must conclude $e^{-Kt}\Phi(t) \leq 0$. Since e^{-Kt} is positive, it must be the case that $\Phi(t) \leq 0$. Together with the inequality $0 \leq K\Phi(t)$, hence $0 \leq \Phi(t)$ (since $K > 0$), we must conclude $\Phi(t) = 0$ for all t. But then also $\varphi(t) = \Phi'(t) = 0$ for all $t \in [t_0, t_e)$. ∎

- **Proof of Theorem 2.3** We only consider the interval $[t_0, t_e)$. The interval $(t_s, t_0]$ can be done similarly, but we must take care of the signs of the integrals when $t < t_0$.

  Suppose there are two solutions $x(t)$ and $y(t)$ of (1) valid on $[t_0, t_e)$ for some $t_e > t_0$. We will show that if $f(t, x)$ satisfies the Lipschitz Condition, then we must have $x(t) = y(t)$ for all $t \in [t_0, t_e)$.

  Let L be the constant corresponding to the Lipschitz Condition of $f(t, x)$. From the integral equation formulation in (3) we find

  $$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau \qquad \text{and} \qquad y(t) = x_0 + \int_{t_0}^{t} f(\tau, y(\tau))\,\mathrm{d}\tau.$$

  Subtracting, we see that

  $$x(t) - y(t) = \int_{t_0}^{t} \big[f(\tau, x(\tau)) - f(\tau, y(\tau))\big]\,\mathrm{d}\tau.$$

Taking the norm of both sides we get

$$0 \le \|x(t) - y(t)\| = \left\| \int_{t_0}^{t} \big[f(\tau, x(\tau)) - f(\tau, y(\tau))\big] \,\mathrm{d}\tau \right\|.$$

Now we use that for integrable functions $a : \mathbb{R} \longrightarrow \mathbb{R}^n$ we have $\left\| \int\limits_{t_0}^{t} a(\tau)\,\mathrm{d}\tau \right\| \le \int\limits_{t_0}^{t} \|a(\tau)\|\,\mathrm{d}\tau$. This should require a proof, but if you recall that the integral is the limit of a large sum, and using the triangle inequality for the norms of sums, I hope you will believe this. Anyway, applying this inequality and the Lipschitz Condition we find

$$\begin{aligned} 0 \le \|x(t) - y(t)\| &= \left\| \int_{t_0}^{t} \big[f(\tau, x(\tau)) - f(\tau, y(\tau))\big] \,\mathrm{d}\tau \right\| \\ &\le \int_{t_0}^{t} \big\| f(\tau, x(\tau)) - f(s, y(\tau)) \big\| \,\mathrm{d}\tau \\ &\le \int_{t_0}^{t} L\|x(\tau) - y(\tau)\|\,\mathrm{d}\tau = L \int_{t_0}^{t} \|x(\tau) - y(\tau)\|\,\mathrm{d}\tau. \end{aligned}$$

Now use Lemma 2.4 with $K = L$ and $\varphi(t) = \|x(t) - y(t)\|$, and we find $\|x(t) - y(t)\| = 0$ for all $t \in [t_0, t_e)$, hence $x(t) = y(t)$, as required. ■

Exercises

1 Consider the function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ given by $f(t, x) = \begin{bmatrix} x_1 + 1 \\ x_2^2 + t^2 \end{bmatrix}$.

(a) Prove that f is locally Lipschitz.

(b) Show that f is not globally Lipschitz on $\mathbb{R}^3$.

2 Suppose the function $f(t, x)$, where $f : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$, satisfies the Lipschitz Condition on a certain domain $D \subseteq \mathbb{R}^{n+1}$, and let $g : \mathbb{R} \longrightarrow \mathbb{R}^n$ be any function.

Prove that h defined by $h(t, x) = f(t, x) + g(t)$ also satisfies the Lipschitz Condition on D.

3 (a) Suppose the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ satisfies the Lipschitz Condition on the whole space $\mathbb{R}^2$. Give an example that shows that this does not guarantee that f is continuous on $\mathbb{R}^2$.

(b) Suppose that the function $f(t, x)$ is globally Lipschitz on $\mathbb{R}^2$. I.e. there is a constant L such that

$$|f(t, x) - f(s, y)| \le L\|(t, x) - (s, y)\| \qquad \text{for all } (t, x), (s, y) \in \mathbb{R}^2.$$

Prove that this means that f is continuous on $\mathbb{R}^2$.