Existence and Uniqueness

The Picard-Lindelöf Theorem

Shamanth Sreekanth

I came across this particular theorem while exploring nonlinear dynamics and chaos. I'm sure most math majors are familiar with it, but we engineers don't usually come across proof-based math. It is however, an essential theorem, with a very clever solution that I think will be appreciated if explained right. The intent of this manuscript, therefore, is to break down the proof step-by-step to an audience that presumably has basic knowledge of first-order differential equations and calculus. By no means is this an original proof, only an original narration.

1 Introduction:

We address a fundamental question in ordinary differential equations (ODEs). Suppose we have a system that evolves in time, and we are given only the following conditions:

- a well-defined starting point: $y(x_0) = y_0$
- how the function changes near that point: $\frac{dy}{dx} = f(x, y)$

They form what is called an initial value problem (IVP). It defines how the system evolves and anchors the system's trajectory with a fixed point, which, as a result, serves as an important prerequisite for prediction and control. Now, how do we know for sure if

- a solution really exists for those conditions, and
- the solution is unique (i.e., there aren't multiple conflicting ways the system can evolve in)?

The Picard-Lindelöf theorem says, if $f(x, y)$ is

- continuous around the point (x_0, y_0), and
- Lipschitz continuous with respect to y,

then,

- there exists one solution $y(x)$ that solves the differential equation near x_0, and
- this solution is unique.

Definition 1 *Continuity*
Continuity ensures that the function behaves predictably and smoothly in a neighborhood around the point, without jumps or breaks. A function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ is said to be continuous at a point $(x_0, y_0) \in D$ if, as the point (x, y) approaches (x_0, y_0), the value of $f(x, y)$ approaches $f(x_0, y_0)$. Basically, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$$

then,

$$|f(x, y) - f(x_0, y_0)| < \varepsilon$$

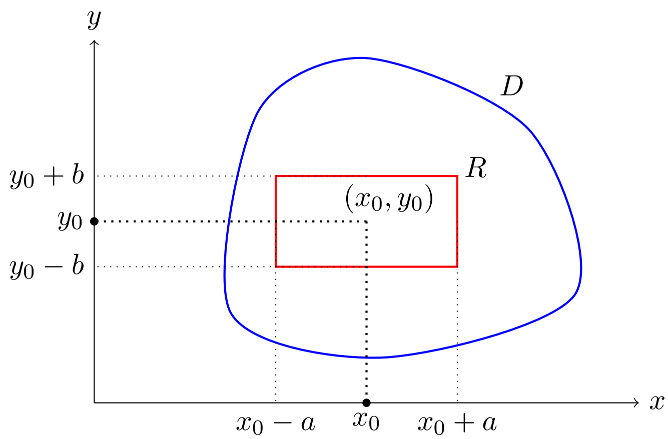
Definition 2 *Lipschitz Continuity*
This implies that the rate of change of f with respect to y is uniformly bounded across the domain. Lipschitz continuity is stronger than ordinary continuity and is essential for guaranteeing

uniqueness of solutions to differential equations. Basically, fixing a point on the x-axis, we vary y, to see how sharply the function changes. A function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ is said to be Lipschitz continuous in the variable y if there exists a constant $L > 0$ such that for all $(x, y_1), (x, y_2) \in D$,

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

With this in hand, we can now state the theorem in full mathematical rigor.

2 Theorem and proof:



Statement: Let D be a domain in $\mathbb{R}^2$, and let $f : D \to \mathbb{R}$ be a real-valued function that satisfies the following conditions:

(i) f is continuous on D,

(ii) $f(x, y)$ is Lipschitz continuous with respect to y on D, with Lipschitz constant $L \geq 0$.

Let (x_0, y_0) be an interior point of D. Let $a > 0$, $b > 0$ be some constants such that the rectangle

$$R = \{(x, y) \in \mathbb{R}^2 \mid |x - x_0| \leq a, \ |y - y_0| \leq b\} \subset D$$

Define

$$M = \max_{(x,y)\in R} f(x, y), \quad h = \min\left(a, \frac{b}{M}\right)$$

Then, the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0$$

has a unique solution $y(x)$ on the interval $|x - x_0| \leq h$.

Let's look at the statement a little carefully. $D \subseteq \mathbb{R}^2$ is an open set in a 2-dimensional plane where our function $f(x, y)$ exists. It's essentially our playground for now. The more important thing to focus on is R, which a rectangle such that its breadth is $2a$ and its height is $2b$, with the center of the rectangle being (x_0, y_0). We assume it houses the two prerequisite properties of the theorem: $f(x, y)$ being continuous and Lipschitz continuous, thereby bounding our function to it. This becomes our "workspace".

We define h similarly. It is the maximum interval $|x - x_0| \leq h$ inside which the solution is guaranteed to exist. Therefore it is essential for h to be defined in a way where the function stays within R and doesn't go out of bounds. We obviously know a is the most we can move horizontally in either direction without leaving R. But the vertical change isn't directly set. It depends on how fast the solution can possible move, i.e., $M = \max f(x, y)$.

$$\frac{dy}{dx} \leq M$$
$$dy \leq M dx$$
$$|y - y_0| \leq M|x - x_0|$$
$$|y - y_0| \leq Mh$$

But, $|y - y_0| \leq b$ vertically. Because of this hard bound, we choose h as something small enough such that

$$Mh \leq b$$
$$h \leq \frac{b}{M}$$

So essentially, h is either a, if the function allows for it, or is $\frac{b}{M}$, whichever is minimum.

$$h = \min\left(a, \frac{b}{M}\right)$$

Okay, so now we understand what we're working on, and what's on the table. We know a starting value. And we know how the function behaves. We know the solution could theoretically exist in that neighborhood. But we don't know the solution itself, or how it's going to look like, or if it even exists. Therefore, we take the initial condition and we try moving along the direction dictated by $f(x, y)$ to hopefully shape ourselves a solution $y(x)$.
Let $x = x_0 + h$ for some $h > 0$.

$$y' = \frac{dy}{dx} = f(x, y)$$
$$\int_{x_0}^{x} y(t)dt = \int_{x_0}^{x} f(t, y(t))dt$$
$$y(x) - y(x_0) = \int_{x_0}^{x} f(t, y(t))dt$$
$$y(x) = y(x_0) + \int_{x_0}^{x} f(t, y(t))dt$$

This equation is the integral form of the IVP, and it helps us construct a recursive sequence which we'll follow to approximate ourselves to the final solution $y(x)$. We'll go ahead and denote every iteration of this sequence as ϕ.
Our starting value will be

$$\phi_0(x) = y_0$$

Plugging this into consecutive iterations,

$$\phi_1(x) = y_0 + \int_{x_0}^{x} f(t, \phi_0(t))\, dt$$
$$\phi_2(x) = y_0 + \int_{x_0}^{x} f(t, \phi_1(t))\, dt$$
$$\vdots$$
$$\phi_n(x) = y_0 + \int_{x_0}^{x} f(t, \phi_{n-1}(t))\, dt \tag{1}$$

This sequence (1) defined for $[x_0, x_0 + h]$ is called Picard's Iterates, and it forms the heart of our proof. To validate ourselves that we're using the right setup, we'll first prove that the set ϕ_n is well defined and continuous, and subsequently that $f(x, \phi_n(x))$ is also well defined and continuous such that $|\phi_n(x) - y_0| \leq b$. Basically we prove that all values of ϕ_n lie within R

Definition 3 *Well-defined*
A function or operation is said to be well-defined if it assigns a unique output to every valid input in its domain. If $f : A \to B$, and $a \in A$, $b \in B$ then the value $f(a) = b$ is unique to a. In this context, (1) is said to be well-defined if the function $f(t, \phi_{n-1}(t))$ is integrable on $[x_0, x_0 + h]$, and the resulting integral yields a valid, unique real number.

We can use mathematical induction to do this, as it's a neat way to establish that a particular statement holds true for all $n \in \mathbb{N}$ iterations of a sequence. It consists of two steps - a *base case* where we show the statement holds true for a value like $n = 0$ or $n = 1$, and an *inductive case* where we assume it holds true for some arbitrary $n = k$ (a hypothesis), and use it to prove it holds true for $n = k + 1$.

Inductive Case: Assume $\phi_n(x)$ exists, has continuous derivative on $[x_0, x_0+h]$, and $|\phi_n(x) - y_0| \leq b$. This naturally implies $(x, \phi_n(x)) \in R$. Then, $f(x, \phi_n(x))$ is well-defined and continuous, as $f(x, \phi_n(x)) \leq M$. Now if,

$$\phi_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, \phi_n(t))\, dt$$

$$\phi_{n+1}(x) - y_0 + \int_{x_0}^{x} f(t, \phi_n(t))\, dt$$

$$|\phi_{n+1}(x) - y_0| = \left| \int_{x_0}^{x} f(t, \phi_n(t))\, dt \right| \leq \int_{x_0}^{x} |f(t, \phi_{n-1}(t))|\, dt$$

$$\leq \int_{x_0}^{x} M\, dt = M(x - x_0)$$

$$\leq Mh \quad \text{since } x - x_0 \leq h \leq b$$

Because of this, $(x, \phi_n(x))$ lies in R_1, so $f(x, \phi_{n+1}(x))$ is defined and continuous on $[x_0, x_0+h]$

Base case: When $n = 1$

$$\phi_1(x) = y_0 + \int_{x_0}^{x} f(t, \phi_0(t))\, dt$$

Since $\phi_0(x) = y_0$ and y_0 are well-defined and continuous, $\phi_1(x)$ is also well-defined & continuous.

$$|\phi_1(x) - y_0| \leq \int_{x_0}^{x} |f(t, y_0)|\, dt \leq M(x - x_0) \leq Mh \leq b$$

Thus, by method of induction, the sequence (1) possesses all desired properties in $[x_0, x_0+h]$.

What we have here is a series of approximations of what could be our solution, becoming consecutively more accurate with every successive iteration. Let's compare two successive iterates.

$$|\phi_{n+1}(x) - \phi_n(x)| = \left| \int_{x_0}^{x} f(t, \phi_n(t))\, dt - \int_{x_0}^{x} f(t, \phi_{n-1}(t))\, dt \right| \leq \int_{x_0}^{x} |f(t, \phi_n(t)) - f(t, \phi_{n-1}(t))|\, dt$$

Notice how the integrand terms are just the function $f(x, y)$ evaluated at two different points on the y-axis, keeping the x-axis value constant. This is precisely what Lipschitz continuity is, as defined previously. If we take L as the Lipschitz constant,

$$|\phi_{n+1}(x) - \phi_n(x)| \leq \int_{x_0}^{x} L|\phi_n(t) - \phi_{n-1}(t)|\, dt$$

Let's now compute the starting few iterations and see what we get.
Taking $n = 0$:

$$|\phi_1(x) - \phi_0(x)| \leq \int_{x_0}^{x} |f(t, \phi_0(t)|\, dt \leq \int_{x_0}^{x} M dt = M(x - x_0) \tag{2}$$

Taking $n = 1$:

$$|\phi_2(x) - \phi_1(x)| \leq \int_{x_0}^{x} L|\phi_1(t) - \phi_0(t)|\, dt$$

From (2),

$$\begin{aligned}
\int_{x_0}^{x} L|\phi_1(t) - \phi_0(t)|\, dt &\leq \int_{x_0}^{x} LM(t - x_0)\, dt \\
&= LM \int_{x_0}^{x} (t - x_0) dt \\
&= LM \left[\frac{(t - x_0)^2}{2}\right]_{x_0}^{x} \\
&= LM \frac{(x - x_0)^2}{2} \qquad (3)
\end{aligned}$$

Taking $n = 2$:

$$|\phi_3(x) - \phi_2(x)| \leq \int_{x_0}^{x} L|\phi_2(t) - \phi_1(t)|\, dt$$

From (3),

$$\begin{aligned}
\int_{x_0}^{x} L|\phi_2(t) - \phi_1(t)|\, dt &\leq \int_{x_0}^{x} L\left(LM\frac{(t - x_0)^2}{2}\right) dt \\
&= L^2 M \int_{x_0}^{x} \frac{(t - x_0)^2}{2}\, dt \\
&= L^2 M \left[\frac{(t - x_0)^3}{6}\right]_{x_0}^{x} \\
&= L^2 M \frac{(x - x_0)^3}{6}
\end{aligned}$$

We see a pattern emerging. As $n \to \infty$,

$$|\phi_{n+1}(x) - \phi_n(x)| \leq L^{n-1} M \frac{(x - x_0)^n}{n!}$$

We know from previous definitions that $x - x_0 \leq h$. Rearranging the terms a bit,

$$\begin{aligned}
&\leq L^n L^{-1} M \frac{h^n}{n!} \\
&= \frac{L^n}{L} \frac{M h^n}{n!}
\end{aligned}$$

$$= \frac{M}{L}\frac{(Lh)^n}{n!} \tag{4}$$

Now what's left is for us to prove that the series converges such that the elements of $\phi_n(x)$ get closer to each other, to give us a final approximation. This is called Cauchy convergence, and is the modus operandi of the Picard's iterates.

Definition 4 *Cauchy Convergence*
A sequence $\{\phi_n\}$ in a metric space is said to be *Cauchy convergent* (or simply, a *Cauchy sequence*) if for every $\varepsilon > 0$, there exists an number $N \in \mathbb{N}$ such that for all $m, n \geq N$, we have

$$|x_n - x_m| < \varepsilon.$$

This basically means, beyond a certain iteration count N, any two iterations of that sequence (they don't even have to be consecutive), their difference will be less than ε. In this context, the sequence of functions $\{\phi_n(x)\}$ is Cauchy convergent in the space of continuous functions on $[x_0, x_0 + h]$ in the supremum norm, if

$$\sup_{x\in[x_0,x_0+h]} |\phi_n(x) - \phi_m(x)| < \varepsilon \quad \text{for all } m, n \geq N.$$

Although the difference is very subtle, a good question to ask is why supremum? Why not maximum?

Definition 5 *Supremum Norm*
Given a function f defined on a closed interval $[a, b]$, the *supremum norm* (also called the *uniform norm*) of f is defined as

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|.$$

It represents the maximum absolute value attained by the function on the interval $[a, b]$. For a sequence of functions $\{\phi_n(x)\}$, convergence in the supremum norm means that

$$\|\phi_n - \phi\|_\infty = \sup_{x\in[a,b]} |\phi_n(x) - \phi(x)| \to 0 \quad \text{as } n \to \infty.$$

The supremum (sup) and maximum (max) are very similar, but the key difference lies in whether the maximum value is actually attained. Maximum is a value that a function actually reaches on the interval. Supremum is the smallest possible value that we can classify as an upper bound, even if the function never actually hits that value. What we basically mean to say is, the minimum possible upper bound of $|\phi_n(x) - \phi_m(x)|$ is 0, even if it never truly reaches 0 and gets very close to it. Because the series has this nature, this possibility, of never truly reaching 0, we use supremum.

So, by definition of Cauchy convergence, let's take two iterations beyond a point $N \in \mathbb{N}$. Assuming $m > n$, we can break them down to their consecutive parts:

$$|\phi_m(x) - \phi_n(x)| \leq |\phi_m(x) - \phi_{m-1}(x)| + |\phi_{m-1}(x) - \phi_{m-2}(x)| + ... + |\phi_{n+1}(x) - \phi_n(x)|$$

We know 4 holds true for consecutive iterates. Applying it as a summation for the above equation,

$$|\phi_m(x) - \phi_n(x)| \leq \frac{M}{L}\sum_{k=n}^{m-1} \frac{(Lh)^k}{k!} \tag{5}$$

We make a key observation here, something that solves one half of our problem statement. (5) is the tail of a well-known power series of the form e^z. If $z = Lh$,

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

$$e^{Lh} = \sum_{k=0}^{\infty} \frac{Lh^k}{k!}$$

We know the function e^{Lh} converges because it is well known to do so. It can be verified through a ratio test, if needed. There exists a small non-zero value $\varepsilon > 0$, and $N \in \mathbb{N}$ such that, for any of $N_0 \geq N$,

$$\frac{M}{L} \sum_{k=N_0}^{\infty} \frac{Lh^k}{k!} < \varepsilon$$

As $n \to \infty$, $\phi_n(x)$ gets increasingly closer to being our hypothesized solution $\phi(x)$. To evaluate this, we take their difference.

$$|\phi_n(x) - \phi(x)| = \left| \int_{x_0}^{x} f(t, \phi_{n-1}(t))dt - \int_{x_0}^{x} f(t, y(t))dt \right|$$

$$\leq \int_{x_0}^{x} L|(\phi_{n-1}(t) - \phi(t))|dt$$

We've already established that $|\phi_{n-1} - \phi| \leq ||\phi_{n-1} - \phi||$, and that the supremum $||\phi_{n-1} - \phi||$ is a constant upper bound value.

$$\leq L \int_{x_0}^{x} ||\phi_{n-1} - \phi||dt = L||\phi_{n-1} - \phi|| \int_{x_0}^{x} dt$$

$$= L||\phi_{n-1} - \phi||(|x - x_0|)$$

$$< Lh||\phi_{n-1} - \phi|| \to 0$$

When n gets sufficiently large the difference between ϕ_n and ϕ converges further, making $||\phi_{n-1} - \phi|| \to 0$ within some $\varepsilon > 0$.

Hence we have proved that there indeed exists a solution.

To check uniqueness, an easy way is to assume there are multiple solutions, and work our way backwards. Suppose there are two solutions, $y_1(x)$ and $y_2(x)$. They both have the same IVP where,

$$y_1(x_0) = y_0$$

$$y_2(x_0) = y_0$$

$$y_1(x) = y_0 + \int_{x_0}^{x} f(t, y_1(t))dt$$

$$y_2(x) = y_0 + \int_{x_0}^{x} f(t, y_2(t))dt$$

Subtracting them from each other,

$$|y_1(x) - y_2(x)| = \left| \int_{x_0}^{x} f(t, y_1(t))dt - f(t, y_2(t))dt \right|$$

$$\leq \int_{x_0}^{x} L|(y_1(t) - y_2(t))|dt$$

If

$$z(x) := |y_1(x) - y_2(x)|$$

Then

$$z(x) \leq L \int_{x_0}^{x} z(t)dt$$

We observe that this form is similar to the Grönwall's Inequality definition.

Lemma 1 *(Grönwall's Inequality)*
Let $z : [a, b] \to \mathbb{R}$ be a continuous, nonnegative function. Suppose there exists a constant $C \geq 0$ and a continuous, nonnegative function $A : [a, b] \to \mathbb{R}$ such that

$$z(x) \leq C + \int_a^x A(t)\, z(t)\, dt, \quad \forall x \in [a, b].$$

Then $z(x)$ is bounded by

$$z(x) \leq C \exp\left(\int_a^x A(t)\, dt \right), \quad \forall x \in [a, b].$$

This lemma basically tells that if a function is bounded by a constant and an integral term that grows proportionally to itself, then the function can grow at most exponentially, with the rate determined by the function $A(t)$. In our case however, $A(t) = L$ and $C = 0$, which results in

$$z(x) \leq 0(e^{L \int_{x_0}^{x} dt}) = 0(e^{L(x - x_0)}) = 0$$

Because $|y_1(x) - y_2(x)| = z(x) = 0$, this implies that there **exists** only one **unique** solution $y(x)$, which finally completes the proof.