# Lecture 15: Picard's theorem

updated 20/10/20

The same kind of "solution by iteration" that we saw in Newton's method and Banach's fixed point theorem can be used to solve ordinary differential equations; this method is usually referred to as Picard iteration. The technique actually predates Banach's fixed point theorem (for references see p. 181 of W. Cheney "Analysis for applied mathematics") but nowadays it is usually proved as a consequence of that theorem.

Recall that in Example L14-2 we rephrased the problem of finding a solution of an equation $g(v) = w$ ($g: V \to V$, $V$ a vector space, $w \in V$ fixed) as the problem of finding a fixed point of $f(v) = v - g(v) + w$. Suppose now we want to solve an ODE, which is an equation in an unknown continuously differentiable function $\varphi$ of a single real variable:

$$\frac{d}{dx}\varphi = h \circ \langle 1, \varphi \rangle \qquad (1.1)$$

where $h: \mathbb{R}^2 \to \mathbb{R}$ is continuous, and $\langle 1, \varphi \rangle : \mathbb{R} \to \mathbb{R}^2$ is $x \mapsto (x, \varphi(x))$. This is an equation of functions (so, an equation in $Cts(\mathbb{R}, \mathbb{R})$) but it is probably more familiar as an infinite family of equations in $\mathbb{R}$, i.e.

$$\varphi'(x) = h(x, \varphi(x)). \qquad (1.2)$$

In any case, the question is: what is the function $f$ whose fixed points are precisely the solutions of this ODE? That is

$$f(\varphi) = \varphi \iff \frac{d}{dx}\varphi - h \circ \langle 1, \varphi \rangle = 0.$$

so our $g(\varphi)$ is this expression, and $w = 0$.

But if $\varphi$ is a solution and $f(\varphi)=\varphi$ then

$$\frac{d}{dx} f(\varphi) = \frac{d}{dx}\varphi = h \circ \langle 1, \varphi \rangle$$

$$\therefore \quad f(\varphi) = C + \int h(x, \varphi(x))\, dx. \qquad (2.1)$$

This gives us a reasonable guess for $f$, although we need to fix $C$, which means imposing an initial condition, say $\varphi(x_0)=y_0$, on our ODE.

<u>Theorem L15-1</u> (Picard) Let $h: U \longrightarrow \mathbb{R}$ be continuous for some open set $U \subseteq \mathbb{R}^2$ containing $(x_0, y_0)$ and suppose there exists a constant $\alpha > 0$ with

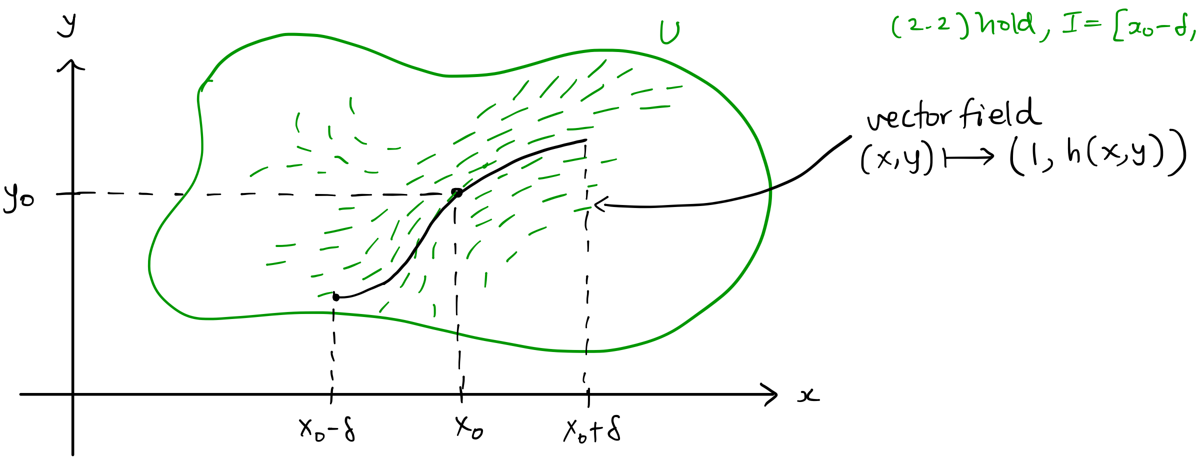
$$|h(x,y_1) - h(x,y_2)| \leq \alpha |y_1 - y_2|$$

for all $(x,y_1),(x,y_2) \in U$. Then there exists $\delta > 0$ such that the initial value problem

$$\varphi'(x) = h(x, \varphi(x)), \qquad \varphi(x_0) = y_0 \qquad (2.2)$$

has a unique solution on $[x_0 - \delta, x_0 + \delta]$.

[By a <u>solution</u> we mean a continuously differentiable function $\varphi: I \longrightarrow \mathbb{R}$ making (2.2) hold, $I = [x_0-\delta, x_0+\delta]$]



<u>Proof</u> Let $\delta > 0$, $b > 0$ be such that $I \times J \subseteq U$ where $I = [x_0 - \delta, x_0 + \delta]$ and $J = [y_0 - b, y_0 + b]$. Since $h$ is continuous and $I \times J$ is compact, $h|_{I \times J}$ is bounded, say $|h(x,y)| \leq M$ for all $(x,y) \in I \times J$. By shrinking $\delta$ we may assume $\alpha\delta < 1$ and $M\delta < b$. We look for solutions to the initial value problem in the space

$$\mathrm{Cts}(I, J) = \mathrm{Cts}([x_0 - \delta, x_0 + \delta], [y_0 - b, y_0 + b])$$

which by Corollary L13-6 is a complete metric space (since $J$ is a closed subspace of $\mathbb{R}$, and therefore complete by Ex. L13-9). Of course we need solutions which are not just continuous but <u>differentiable</u>, but this will work out, as we will see. Define following (2.1) the map

$$f : \mathrm{Cts}(I, J) \longrightarrow \mathrm{Cts}(I, J)$$

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

Now $\varphi$ is continuous, so $h \circ \langle 1, \varphi \rangle$ is continuous on $I$ and hence Riemann integrable on $I$. By the fundamental theorem of calculus $x \mapsto \int_{x_0}^{x} h \circ \langle 1, \varphi \rangle$ is differentiable on $I$ with derivative $h \circ \langle 1, \varphi \rangle$. The function $f(\varphi) : I \longrightarrow \mathbb{R}$ is therefore (continuously) differentiable:

$$f'(\varphi) = h \circ \langle 1, \varphi \rangle.$$

for any continuous $\varphi$. It remains to check that $f(\varphi)(I) \subseteq J$, and to show that $f$ is a contraction with respect to the $d_\infty$ metric. Suppose for a moment that we have done both of these things. Then the Banach fixed point theorem (L14-1) tells us that $f$ has a unique fixed point

Note that since $f(\varphi)$ always has a continuous derivative, a fixed point is necessarily continuously differentiable. And

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

$$\iff \varphi' = h \circ \langle 1, \varphi \rangle \quad \text{on } I, \text{ and } \varphi(x_0) = y_0$$

$$\iff \varphi \text{ is a solution of the initial value problem on } I.$$

The reverse implication ($\Leftarrow$) in the last step is by the second fundamental theorem of calculus (two antiderivatives differ by a constant). This completes the proof, once we have checked the two aforementioned items:

$f(\varphi)(I) \subseteq J$ for this we need (for $x > x_0$)

$$\left| \int_{x_0}^{x} h(t,\varphi(t))\,dt \right| \leq \int_{x_0}^{x} \left| h(t,\varphi(t)) \right| dt$$

$$\leq \int_{x_0}^{x} M\,dt \qquad (\text{since } |h(x,y)| \leq M \text{ on } I \times J)$$

$$= M(x - x_0) \leq M\delta < b$$

and similarly for $x < x_0$. This shows $f(\varphi) \in Cts(I,J)$.

$f$ is a contraction $\quad d_\infty(f\varphi, f\psi) = \sup\{ | f(\varphi)(x) - f(\psi)(x) | \mid x \in I \}$

$$= \sup\left\{ \left| \int_{x_0}^{x} h(t,\varphi(t))\,dt - \int_{x_0}^{x} h(t,\psi(t))\,dt \right| \;\middle|\; x \in I \right\}.$$

$$\leq \sup\left\{ \int_{x_0}^{x} \left| h(t,\varphi(t)) - h(t,\psi(t)) \right| dt \;\middle|\; x \in I \right\}$$

But for $t \in I$, $\varphi(t), \psi(t)$ are both in $J$, so by hypothesis

$$\left| h(t, \varphi(t)) - h(t, \psi(t)) \right| \leq \alpha \left| \varphi(t) - \psi(t) \right|$$

Hence

$$\int_{x_0}^{x} \left| h(t, \varphi(t)) - h(t, \psi(t)) \right| dt \leq \alpha \int_{x_0}^{x} \left| \varphi(t) - \psi(t) \right| dt$$

$$\leq \alpha \cdot |x - x_0| \cdot \sup\{ |\varphi(t) - \psi(t)| \mid t \in I \}$$

$$\leq \alpha \delta \cdot d_\infty(\varphi, \psi).$$

Then it follows that $d_\infty(f\varphi, f\psi) \leq \alpha\delta \cdot d_\infty(\varphi, \psi)$. We chose $\delta$ such that $\alpha\delta < 1$, this shows $f$ is a contraction. $\square$

Moreover the fixed point theorem tells us how to find a solution by iteration, namely, choose any continuous function $\varphi_0 : I \longrightarrow J$ ( $\varphi_0 \equiv y_0$ is a safe bet) and then take the limit of $\varphi_0, f\varphi_0, f^2\varphi_0, \ldots, f^n\varphi_0, \ldots$. According to W. Cheney (cited above, see p. 181) this method is "rarely used directly in the numerical solution of initial value problems because the step-by-step methods of numerical integration are superior", but nonetheless this limit is guaranteed to converge (eventually).

Remark L15-1 We have not given the strongest possible statement of Picard's theorem (see Cheney's book). The theorem is easily extended to systems of first order ODEs and thereby in the usual way to higher-order ODEs. However, PDEs are a completely different story.

Example L15-1 Let $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be $h(x,y) = y$, so the differential $eq^N$ is

$$\varphi' = \varphi.$$

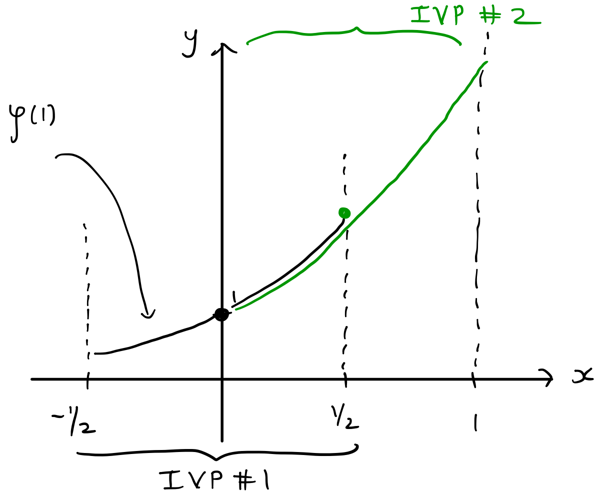
Then $|h(x,y_1) - h(x,y_2)| = |y_1 - y_2|$ so $\alpha \geq 1$ will do. Take the initial condition $\varphi(0) = 1$, and as the starting point of our iteration $\varphi_0 \equiv 1$. Then

$$\varphi_1 = f(\varphi_0) = 1 + \int_0^x \varphi_0(t)\,dt = 1 + x$$

$$\varphi_2 = f(\varphi_1) = 1 + \int_0^x \varphi_1(t)\,dt = 1 + \left[t + \tfrac{1}{2}t^2\right]_0^x = 1 + x + \tfrac{1}{2}x^2$$

One proves by induction that $\varphi_n = \sum_{i=0}^n \frac{1}{i!}x^i$. Note we are free to choose $I = [-\delta, \delta]$, $J = [1-b, 1+b]$ arbitrarily provided $\delta < 1$ (so that we may choose $\alpha < 1/\delta$) and $\delta < b/1+b$ (so just take $b$ large, and any $\delta < 1$ will do). So on $I$ we find that $\lim_{n\to\infty} \varphi_n = e^x$ is the unique solution.

This example shows how the requirement $\alpha\delta < 1$ is a bit... cheesy. We can extend our solution by first finding a solution $\varphi^{(1)}$ on $[-1/2, 1/2]$ and then applying the same method to the IVP $\varphi' = \varphi$ with $\varphi(1/2) = \varphi^{(1)}(1/2)$. We will get a unique solution $\varphi^{(2)}$ to this second problem on $[0,1]$, and by uniqueness it agrees with $\varphi^{(1)}$ on the overlap. We can repeat this in both directions to show that there is a unique $sol^N$ (namely $e^x$) on all of $\mathbb{R}$.




However if you look at the details you will see this extension relied on the particular nature of $h(x,y) = y$. In general if $h : [a,b] \times \mathbb{R} \longrightarrow \mathbb{R}$ we can always extend to all of $[a,b]$.

There is something quite remarkable about this example: the iteration constructs a sequence of polynomial functions converging to the unique solution. It is clear that this works more generally, to show that solutions of polynomial ODEs may be written as the uniform limit of polynomials: if $h : U \to \mathbb{R}$ is a polynomial function and we take $\varphi_0 \equiv y_0$ then every function $f^n(\varphi_0)$ constructed by the iteration will be a polynomial, because the integral of a polynomial is a polynomial. We already know that convergence in $Cts(I, J)$ means uniform convergence.

Remark L15-2 Why should we care about function spaces, say $Cts(X, \mathbb{R})$?

Beginning in Lecture 12 we have made the point that continuous maps $X \to \mathbb{R}$ can represent configurations of physical systems, with those configurations consistent with physical law typically being a subset $A \subseteq Cts(X, \mathbb{R})$ consisting of sol$^{N}$s to some differential equation. We certainly care about the set $A$ But why are we forced to care about the set of all continuous maps, and further, why must we care about the topology on this set?

Consider the expression $e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n$. We can view the right hand side as a kind of algorithm which constructs the real number $e$ beginning with integers, where the operations allowed in the construction are the usual arithmetic operations (here used are addition, division and iterated multiplication aka exponentiation) together with the limit. By definition $\mathbb{N}$ plus these operations "generate" $\mathbb{R}$. Moreover if we apply a continuous function $f : \mathbb{R} \to \mathbb{R}$ to $e$, it transforms the algorithm for constructing $e$ to an algorithm for constructing $f(e)$, provided $f$ itself is computed by some algorithm:

$$f(e) = f\left(\lim_{n\to\infty}\left(1+\tfrac{1}{n}\right)^n\right) = \lim_{n\to\infty} f\left(\left(1+\tfrac{1}{n}\right)^n\right).$$

For example, $e^2 = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{2n}$ presents $e^2$ as a limit of rational numbers.

It is the topology on $\mathbb{R}$ which provides the ambient structure that gives limits, and thus such algorithms, meaning. Incidentally, such considerations are developed at length in Turing's paper: A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem", 1936, which is more famous for introducing what are now called Universal Turing Machines.

Returning to differential equations and our set of solutions $A \subseteq \mathrm{Cts}(X, \mathbb{R})$, with say $\psi \in A$ expressed as a uniform limit $\psi = \lim_{n\to\infty} \varphi_n$ of polynomial functions $\varphi_n \in \mathrm{Cts}(X, \mathbb{R})$ (not themselves solutions), we see that polynomial functions play a role analogous to integers or rational numbers, as the "simple" functions which generate via limits (and thus the topology on $\mathrm{Cts}(X, \mathbb{R})$) other functions of interest. Moreover the "algorithm" $\psi = \lim_{n\to\infty} \varphi_n$ for constructing $\psi$ may be transformed to an algorithm for constructing quantities that are a continuous function of $\psi$, for example if $X = [a,b]$

$$\int_{[a,b]} \psi = \int_{[a,b]} \lim_{n\to\infty} \varphi_n = \lim_{n\to\infty} \int_{[a,b]} \varphi_n$$

and $\int_{[a,b]} \varphi_n$ is easily computed since $\varphi_n$ is polynomial. Here we have used that $\int_{[a,b]}(-)$ is continuous, see Exercise L15-3.

In conclusion: to compute with solutions $\psi$ we use constructions of such solutions $\psi = \lim_{n\to\infty} \varphi_n$ as limits of "approximate solutions" $\varphi_n$ taken from a class of "simple" functions (e.g. polynomials). It is the topology on $\mathrm{Cts}(X, \mathbb{R})$ which provides the ambient structure that gives such constructions meaning.

This all points to a natural question, which we will address in Lecture 16:

Question: which functions $\psi \in \mathrm{Cts}([a,b], \mathbb{R})$ may be written as a uniform limit of polynomials?

meaning, a limit in $(\mathrm{Cts}([a,b], \mathbb{R}), d_\infty)$

<u>Exercise 45-1</u>* This exercise walks you through the extension of the Theorem to systems of first-order ODEs, i.e.

$$\begin{aligned}
\varphi_1'(x) &= h_1(x, \varphi_1(x), \dots, \varphi_n(x)) & \varphi_1(x_0) &= y_0^{(1)} \\
\varphi_2'(x) &= h_2(x, \varphi_1(x), \dots, \varphi_n(x)) & \varphi_2(x_0) &= y_0^{(2)} \\
&\vdots & &\vdots \\
\varphi_n'(x) &= h_n(x, \varphi_1(x), \dots, \varphi_n(x)) & \varphi_n(x_0) &= y_0^{(n)}.
\end{aligned}$$

Let $h : U \longrightarrow \mathbb{R}^n$ be continuous where $U \subseteq \mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$ is open, then a <u>solution</u> of the above IVP on an interval $I \subseteq \mathbb{R}$ containing $x_0$ is a function $\varphi : I \longrightarrow \mathbb{R}^n$ (whose components are the $\varphi_i(x)$) which is continuously differentiable (meaning each $\varphi_i(x)$ is so) with the property that as functions (where $\varphi'(x) = (\varphi_1'(x), \dots, \varphi_n'(x))$)

$$\varphi' = h \circ \langle 1, \varphi \rangle, \quad \varphi(x_0) = y_0 = (y_0^{(1)}, \dots, y_0^{(n)}).$$

Suppose $\alpha > 0$ exists with

$$\| h(x, y_1) - h(x, y_2) \| \leq \alpha \| y_1 - y_2 \| \quad \forall (x, y_1), (x, y_2) \in U$$

with $\| - \| : \mathbb{R}^n \longrightarrow \mathbb{R}$ given by $\|y\| = \sum_{i=1}^n |y_i|$. Also assume that $(x_0, y_0) \in U$. <u>Prove</u> that there exists $\delta > 0$ such that the IVP has a unique solution on $[x_0 - \delta, x_0 + \delta]$. (Note: by Ex. 43-10 you can choose a metric on $\mathbb{R}^n$ which suits you)

<u>Exercise L15-2</u> Prove using the previous exercise that the IVP $\varphi'' = -\varphi$, $\varphi(0)=0$, $\varphi'(0)=1$ has a unique solution on $[-\delta, \delta]$ for some $\delta$, and use Picard iteration to give a powerseries expansion of the sol$^N$.
(for the second part you will need to know the appropriate $f$).

<u>Exercise L15-3</u> Prove that the function

$$\begin{aligned} C_b([a,b], \mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_{[a,b]} f \end{aligned}$$

is continuous, where $\int$ denotes the Riemann integral (see Tutorial 8 and T. Tao's book "Analysis" for a reminder), and $C_b([a,b], \mathbb{R})$ has the compact-open topology. (Hint: use Lemma L8-4 and recall the interaction of uniform convergence and integrals, from say Tao's book, Theorem 14.6.1 of Vol. 2).

<u>Remark</u> For Picard's theorem we need $\alpha > 0$ such that

$$|h(x,y_1) - h(x,y_2)| \leqslant \alpha |y_1 - y_2| \qquad (10.1)$$

for all $(x,y_1), (x,y_2) \in U$. Suppose for simplicity that whenever $(x,y_1) \in U$ and $(x,y_2) \in U$ with $y_2 \geqslant y_1$ then $(x,y) \in U$ for all $y_1 \leqslant y \leqslant y_2$. Then if $\partial h / \partial y$ exists on $U$ and $|\partial h/\partial y \,(x,y)| \leqslant \alpha$ for all $(x,y) \in U$ then given $(x,y_1), (x,y_2) \in U$ with $y_2 \geqslant y_1$ we have

$$\frac{h(x,y_2) - h(x,y_1)}{y_2 - y_1} = \frac{\partial h}{\partial y}(x,c) \qquad \text{some } c \in [y_1, y_2]$$

and hence (10.1) holds for the given $\alpha$. This is the most common source of the bound (10.1) necessary for Picard's theorem.