Picard's Theorem

Our major application of the theory we have been investigating is the proof of a key basic theorem in the theory of ordinary differential equations. We investigate the following problem (known as an *initial value problem*): find a real-valued function x defined on some interval $[t_0 - \beta, t_0 + \beta]$, satisfying the differential equation

$$x'(t) = f(t, x(t)) \tag{1}$$

and the initial condition

$$x(t_0) = x_0 \tag{2}$$

where $f(t, x)$ is some given expression in t and x.

Intuition dictates that there should be a unique function x satisfying this. The graph of x has to pass through the point (t_0, x_0), and the differential equation determines the slope of the graph at any given point. If one imagines a particle whose position at time t is $(t, x(t))$ then the differential equation is steering the particle, so to speak: it determines the direction of travel at any point. Given that the starting point is specified, the entire path the particle follows must surely be uniquely determined.

Although the above reasoning may be plausible, it is not precise enough to be accepted as a mathematical proof. The main problem is that it is not clear what precise mathematical conditions are implicitly assumed by the intuitive identification of a function with a graph that one could draw, with the slope of the graph corresponding to the derivative. For example, if the function f appearing on the right hand side of Eq. (1) is discontinuous everywhere, then the solution function x would have to be differentiable with a derivative that is everywhere discontinuous. Our intuition does not usually deal with graphs like that. So it will be no surprise if we have to assume at least that f is continuous before we can give a rigorous proof.

The precise theorem that we are able to prove is as follows.

Theorem. *Let $S = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$, a rectangle in $\mathbb{R}^2$, and let f be a continous real-valued function on S satisfying*
(i) $|f(t, x)| \leq c$ for all $(t, x) \in S$,
(ii) $|f(t, x) - f(t, y)| \leq k|x - y|$ for all $t \in [t_0 - a, t_0 + a]$ and $x, y \in [x_0 - b, x_0 + b]$,
where c and k are (positive) constants. Let $\beta \in \mathbb{R}$ satisfy $\beta \leq a$ and $\beta \leq b/c$. Then there exists a unique function $x\colon [t_0 - \beta, t_0 + \beta] \to \mathbb{R}$ such that $x(t_0) = x_0$ and $x'(t) = f(t, x(t))$ for all $t \in [t_0 - \beta, t_0 + \beta]$.

Comments

a) Condition (i) is automatically satisfied for some constant c, given that f is continuous: there is a theorem which says that a continuous real valued function on a closed and bounded subset of $\mathbb{R}^n$ (such as the set S) must be bounded. Condition (ii) is automatically satisfied for some constant k if the partial derivative $\partial f/\partial x$ exists and is continuous on S; this is explained below.

b) A real-valued function g defined on an interval $[p, q]$ in $\mathbb{R}$ is said to satisfy a *Lipschitz condition* on $[p, q]$ if there exists a constant k (the *Lipschitz constant*) such that $|g(x) - g(y)| \leq k|x - y|$ for all $x, y \in [p, q]$. So condition (ii) can be reformulated as follows: f satisfies a Lipschitz condition with respect to its second argument.

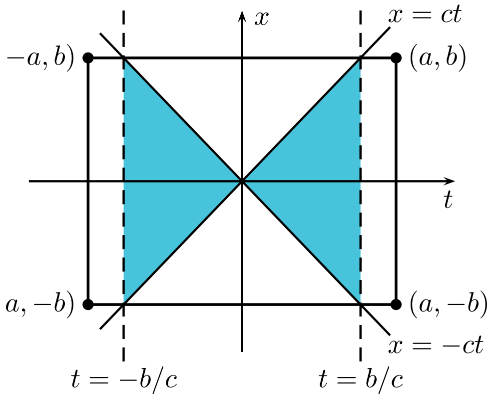
c) For the sake of simplifying the notation, we shall assume that $x_0 = 0$ and $t_0 = 0$. There is clearly no loss of generality involved in this; it just amounts to making a translation of the coordinate system, so that the new origin is the former point (t_0, x_0). In the new coordinates the rectangle S is $[-a, a] \times [-b, b]$, and the initial condition is $x(0) = 0$. So we are saved the trouble of writing a lot of x_0's and t_0's.

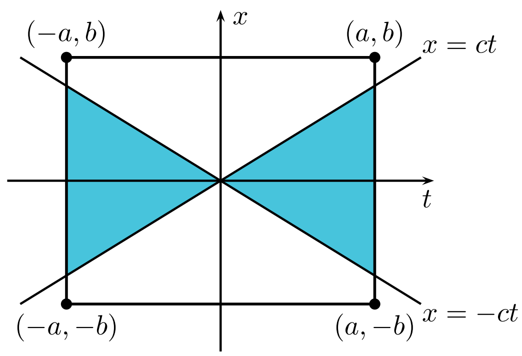
d) Although the Lipschitz condition made no explicit appearance in our intuitive explanation of why an initial value problem should have a solution, we need do it at a crucial point in our proof. Although it is conceivable that a more cunning proof than the one we shall give could replace condition (ii) with some weaker assumption, it is probably the case that intuition only deals with functions that satisfy a Lipschitz condition. For example, perhaps we are accustomed to thinking only of functions whose partial derivatives are continuous; this is a more restrictive assumption than the Lipschitz condition. Thus, suppose that $\partial f/\partial x$ exists and is continuous on S. Then $\partial f/\partial x$ is bounded on S (just as continuity of f implies that f is bounded). Let k be an upper bound, and for arbitrary $t \in [-a, a]$ and x, $y \in [-b, b]$, apply the Mean Value Theorem to the function $\phi\colon z \mapsto f(t, z)$ on the interval $[x, y]$. The conclusion is that

$$f(t, x) - f(t, y) = \phi(x) - \phi(y) = \phi'(z_0)(x - y)$$

for some $z_0 \in (x, y)$. But since $\phi' = \partial f/\partial x$ we have that $|\phi'(z_0)| \leq k$, and thus $|f(t, x) - f(t, y)| = |\phi'(z_0)|\,|x - y| \leq k|x - y|$, as required.

e) The theorem says that a solution $x(t)$ exists on $[-\beta, \beta]$ provided that β satisfies the conditions $\beta \leq a$ and $\beta \leq b/c$. These requirements are both quite natural. Since the differential equation (1) only makes sense when $(t, x(t))$ is in the domain of the function f, which is the rectangle S, we cannot hope to prove anything about $x(t)$ for $|t| > a$. Hence we get the condition $\beta \leq a$. The restriction $\beta \leq b/c$ is perhaps less clear at first, but it comes about for the same kind of reason. Since $|f(x, t)|$ is bounded by c, the differential equation implies that $|x'(t)| \leq c$. In an extreme case we could have $x'(t) = c$, making the graph of x a straight line of slope c through the origin $(0, 0)$ (since $x(0) = 0$ is assumed). If $b/c < a$ then this line exits the rectangle $[-a, a] \times [-b, b]$ at the point $(b/c, b)$; so in this case the point $(t, x(t))$ is outside the domain of f when $t > b/c$. So, in general, unless $\beta \leq b/c$ we cannot be sure that $(t, x(t))$ will remain in S.





In the diagrams above, illustrating the two cases $a > b/c$ and $a < b/c$, the shaded region is the part of the (x, t)-plane defined by $-\beta \leq t \leq \beta$ and $|x| \leq c|t|$. Since the graph of $x(t)$ must pass through $(0, 0)$ and have slope between $-c$ and c, it must lie in this shaded region. (Note that S is the rectangle with corners at $(\pm a, \pm b)$.)

Before proving Picard's Theorem, we recall a concept that was introduced in Lecture 14. (See also Exercise 4 of Tutorial 5 and Exercise 2 of Tutorial 8.)

Definition. An *isometry* from a metric space (X, d_X) to a metric space (Y, d_Y) is a function $\phi: X \to Y$ such that $d_Y(\phi(x_1), \phi(x_2)) = d_X(x_1, x_2)$.

It is easily seen that an isometry has to be injective. For, suppose that ϕ is an isometry, and suppose that x_1, $x_2 \in X$ with $\phi(x_1) = \phi(x_2)$. Then $d_Y(\phi(x_1), \phi(x_2)) = 0$, and since ϕ is an isometry it follows that $d_X(x_1, x_2) = 0$, and so $x_1 = x_2$. So we have shown, as required, that if x_1, $x_2 \in X$ with $\phi(x_1) = \phi(x_2)$ then $x_1 = x_2$. Furthermore, note that, as we saw in Question 4 of Tutorial 5, whenever ϕ is an injective function from an arbitrary set X to a set Y that is equipped with a metric d_Y, it is possible to define a metric d_X on X by means of the formula $d_X(x_1, x_2) = d_Y(\phi(x_1), \phi(x_2))$. In other words, we define d_X in such a way that ϕ becomes an isometry.

Metric spaces X and Y are said to be *isometric* if there is an isometry $X \to Y$ that is bijective. Isometric spaces can be regarded as copies of the same abstract space: as far as the metric structure is concerned, they are the same as each other. Properties of either space that can be formulated in terms of the distance function will be mirrored in the other space. For example, the concept of convergence of sequences is defined in terms of distance; so a sequence in one of the spaces will be convergent if and only if the corresponding sequence in the other space is convergent. Similarly, a sequence of points in one of the spaces will be a Cauchy sequence if and only if the corresponding sequence of points in the other space is a Cauchy sequence. And it follows that if X and Y are isometric spaces then X is complete if and only if Y is complete. All the concepts involved in the definition of completeness are ultimately expressible just in terms of the distance function; an isometric correspondence necessarily preserves such properties.

Returning now to Picard's Theorem, define $\mathcal{C}$ to be the set of all continuous real-valued functions on the interval $[-\beta, \beta]$, and define

$$\mathcal{C}^* = \{\, x \in \mathcal{C} \;\big|\; |x(t)| \leq c\beta \text{ for all } t \in [-\beta, \beta] \,\}.$$

The relevance of $\mathcal{C}^*$ derives in part from the following lemma.

Lemma. *If x is a solution of the initial value problem given by Eqs (1) and (2) above, then $x \in \mathcal{C}^*$.*

Proof. Since x is differentiable, it is certainly continuous on $[-\beta, \beta]$. Moreover, since f is bounded by c, and $x'(t) = f(t, x(t))$, it follows that for all $t \in [-\beta, \beta]$,

$$|x(t)| = |x(t) - x(0)| = \left| \int_0^t x'(s)\, ds \right| \leq \int_0^t |x'(s)|\, ds = \int_0^t |f(s, x(s))|\, ds \leq \int_0^\beta c\, ds = c\beta,$$

and therefore $x \in \mathcal{C}^*$. □

Observe that if $x \in \mathcal{C}^*$ and $s \in [-\beta, \beta]$ then $-a \leq s \leq a$ (since $\beta \leq a$) and $-b \leq x(s) \leq b$ (since $|x(s)| \leq c\beta \leq b$). So $(s, x(s)) \in S$, the domain of f, and it follows that we can define a function $Tx: [-\beta, \beta] \to \mathbb{R}$ by the formula

$$(Tx)(t) = \int_0^t f(s, x(s))\, ds \qquad (\text{for all } t \in [-\beta, \beta]).$$

The following simple observation is crucial for our strategy.

Lemma. *A function $x \in \mathcal{C}^*$ is a solution of the initial value problem if and only if it satisfies $Tx = x$.*

Proof. Assume first that x is a solution of Eqs (1) and (2). Integrating (1) gives

$$x(t) = \int_0^t f(s, x(s))\, ds + C$$

for some constant C. The initial condition (2) forces $C = 0$, and so

$$x(t) = \int_0^t f(s, x(s))\, ds \qquad \text{(for all } t \in [-\beta, \beta]); \tag{3}$$

That is, $Tx = x$.

Conversely, if we assume that $Tx = x$ then Eq. (3) above holds, and it follows from the Fundamental Theorem of Calculus that $x'(t) = f(t, x(t))$. So x satisfies the differential equation (1). Moreover, Eq. (3) also gives

$$x(0) = \int_0^0 f(s, x(s))\, ds = 0,$$

so that the initial condition (2) is also satisfied. □

The next lemma shows that $x \mapsto Tx$ defines a function from the set $\mathcal{C}^*$ to itself.

Lemma. *If $x \in \mathcal{C}^*$ then $Tx \in \mathcal{C}^*$.*

Proof. Given that $x \in \mathcal{C}^*$, we have, for all $t \in [-\beta, \beta]$,

$$\begin{aligned} |(Tx)(t)| &= \left| \int_0^t f(s, x(s))\, ds \right| \\ &\le \int_0^t c\, ds \qquad \text{(since } |f(s, x(s))| \le c \text{ for all } s) \\ &= c|t| \\ &\le c\beta. \end{aligned}$$

To complete the proof that $Tx \in \mathcal{C}^*$ it remains to show that Tx is continuous on $[-\beta, \beta]$. In fact it follows from the Fundamental Theorem of Calculus Tx is differentiable on $[-\beta, \beta]$, with $(Tx)'(t) = f(t, x(t))$. So Tx is certainly continuous, as required. □

It is, of course, our intention to apply the Banach Fixed Point Theorem, and to do this we shall show that, relative to a suitable metric, $\mathcal{C}^*$ is a complete metric space, and T a contraction mapping on $\mathcal{C}^*$.

We proved in Lecture 12 that the continuous real-valued functions on a closed interval form a complete metric space relative to the uniform metric d. So $\mathcal{C}$ is a complete metric space relative to d (which is defined by $d(x, y) = \sup_{t \in [-\beta, \beta]} |x(t) - y(t)|$ for all x, $y \in \mathcal{C}$). The set $\mathcal{C}^*$ is a subset of $\mathcal{C}$, but rather than regarding $\mathcal{C}^*$ as a subspace of $\mathcal{C}$ we shall define another metric on $\mathcal{C}^*$.

The hypotheses of Picard's Theorem involve a certain constant k (appearing in the Lipschitz condition that f is assumed to satisfy). Using this k, for each $x \in \mathcal{C}^*$ let ηx be the function $[-\beta, \beta] \to \mathbb{R}$ defined by

$$(\eta x)(t) = e^{-k|t|}x(t) \qquad (\text{for all } t \in [-\beta, \beta]).$$

Since the product of two continuous functions is continuous we see that $\eta x \in \mathcal{C}$ for all $x \in \mathcal{C}^*$, and so we may define a function $\eta: \mathcal{C}^* \to \mathcal{C}$ by $x \mapsto \eta x$ (for all $x \in \mathcal{C}^*$). This function is clearly injective, since $e^{-k|t|} \neq 0$ for all t. So, as explained in our discussion of isometries above, there is a metric D on $\mathcal{C}^*$ given by the formula

$$D(x, y) = d(\eta x, \eta y) = \sup_{t \in [-\beta, \beta]} \left| e^{-k|t|}x(t) - e^{-k|t|}y(t) \right|.$$

Since the condition

$$x(t) \leq c\beta \quad (\text{for all } t \in [-\beta, \beta])$$

is satisfied if and only if

$$e^{-k|t|}x(t) \leq e^{-k|t|}c\beta \quad (\text{for all } t \in [-\beta, \beta])$$

it follows that the image of η is the subspace $\widehat{\mathcal{C}}$ of $\mathcal{C}$ given as follows:

$$\widehat{\mathcal{C}} = \{\, y \in \mathcal{C} \mid y(t) \leq e^{-k|t|}c\beta \ \text{ for all } \ t \in [-\beta, \beta] \,\}.$$

In order to deduce that $(\mathcal{C}^*, D)$ is complete we need to know that $(\widehat{\mathcal{C}}, d)$ is complete; this is a consequence of the following lemma.

Lemma. *The set $\widehat{\mathcal{C}}$ is a closed subset of $\mathcal{C}$.*

Proof. By a proposition proved in Lecture 8, it suffices to show that if $(y_n)_{n=1}^{\infty}$ is a convergent sequence in $\mathcal{C}$ with $y_n \in \widehat{\mathcal{C}}$ then the limit y is also in $\widehat{\mathcal{C}}$. But $y_n \in \widehat{\mathcal{C}}$ gives $y_n(t) \leq e^{-k|t|}c\beta$ for all t, and it follows that $y(t) = \lim_{n\to\infty} y_n(t) \leq e^{-k|t|}c\beta$, whence $y \in \widehat{\mathcal{C}}$, as required. □

Being a closed subspace of a complete space, $\widehat{\mathcal{C}}$ must also be complete. And since $(\mathcal{C}^*, D)$ is isometric to the complete space $(\widehat{\mathcal{C}}, d)$ it follows that $(\mathcal{C}^*, D)$ is also a complete space. Now we at last come to the final ingredient of the proof of Picard's Theorem: we show that T is a contraction mapping on $(\mathcal{C}^*, D)$.

Completion of the proof. Let x, $y \in \mathcal{C}^*$, and let $t \in [-\beta, \beta]$. We need to treat the cases $t \geq 0$ and $t \leq 0$ separately, although they are nearly identical. In fact, we shall only do the case $t \leq 0$, leaving the (easier) case $t \geq 0$ as an exercise for the reader. Note that the definition of D yields that

$$e^{-k|s|}|x(s) - y(s)| \leq D(x, y) \qquad \text{for all } s \in [-\beta, \beta], \tag{1}$$

and thus we find that

$$\begin{aligned}
|(Tx)(t) - (Ty)(t)| &= \Big| \int_0^t f(s, x(s)) - f(s, y(s))\, ds \Big| \\
&\le \int_t^0 |f(s, x(s)) - f(s, y(s))|\, ds \qquad \text{(since } t \le 0\text{)} \\
&\le \int_t^0 k|x(s) - y(s)|\, ds \qquad \text{(by the Lipschitz condition } f \text{ satisfies)} \\
&\le \int_t^0 k e^{-ks} D(x, y)\, ds \qquad \text{(by Eq. (1), since } |s| = -s \text{ on } [t, 0]\text{)} \\
&= -ke^{-ks} D(x, y) \Big]_{s=t}^{s=0} \\
&= (-1 + e^{-kt}) D(x, y) \\
&= (e^{k|t|} - 1) D(x, y).
\end{aligned}$$

The same formula holds for $t \ge 0$. Thus

$$\begin{aligned}
D(Tx, Ty) &= \sup_{t \in [-\beta, \beta]} e^{-k|t|} |(Tx)(t) - (Ty)(t)| \\
&\le \sup_t e^{-k|t|} (e^{k|t|} - 1) D(x, y) \\
&= \alpha D(x, y)
\end{aligned}$$

where $\alpha = \sup_{t \in [-\beta, \beta]} (1 - e^{-k|t|}) = 1 - e^{-k\beta}$. As $e^{-k\beta} > 0$, it follows that $\alpha < 1$, and therefore T is a contraction mapping. By the Banach Fixed Point Theorem it follows that there is a unique $x \in \mathcal{C}^*$ such that $Tx = x$. But we have shown that any solution of the initial value problem must be an element of $\mathcal{C}^*$ and a fixed point of T, and conversely that every fixed point of T is a solution of the initial value problem. Thus the initial value problem has a unique solution on $[-\beta, \beta]$, as claimed.