Continuity and Differentiability of Solutions

We now study the dependence of solutions of initial value problems on the initial values and on parameters in the differential equation. We begin with a fundamental estimate.

Definition. We say that $x(t)$ is an ϵ-*approximate solution* of the DE $x' = f(t,x)$ on an interval I if

$$|x'(t) - f(t,x(t))| \le \epsilon \quad \forall\, t \in I.$$

Here we will consider only C^1 functions $x(t)$. See Coddington & Levinson to extend this concept to piecewise C^1 functions, etc.

Fundamental Estimate

Let $f(t,x)$ be continuous in t and x, and uniformly Lipschitz continuous in x with Lipschitz constant L. Suppose $x_1(t)$ is an ϵ_1-approximate solution of $x' = f(t,x)$ and $x_2(t)$ is an ϵ_2-approximate solution of $x' = f(t,x)$ on an interval I with $t_0 \in I$, and suppose $|x_1(t_0) - x_2(t_0)| \le \delta$. Then

$$|x_1(t) - x_2(t)| \le \delta e^{L|t-t_0|} + \frac{\epsilon_1 + \epsilon_2}{L}(e^{L|t-t_0|} - 1) \quad \forall\, t \in I.$$

Remarks.

(1) The first term on the RHS bounds the difference between the solutions of the IVPs with initial values $x_1(t_0)$ and $x_2(t_0)$ at t_0.

(2) The second term on the RHS accounts for the fact that $x_1(t)$ and $x_2(t)$ are only approximate solutions. Note that this term is 0 at $t = t_0$.

(3) If $\epsilon_1 = \epsilon_2 = \delta = 0$, we recover the uniqueness theorem for Lipschitz f.

Proof. We may assume $\epsilon_1, \epsilon_2, \delta > 0$ (otherwise, take limits as $\epsilon_1 \to 0^+$, $\epsilon_2 \to 0^+$, $\delta \to 0^+$). Also for simplicity, we may assume that $t_0 = 0$ and $t \ge 0$ (do time reversal for $t \le 0$). Set

$$u(t) = |x_1(t) - x_2(t)|^2 = \langle x_1 - x_2, x_2 - x_2 \rangle.$$

Then

$$\begin{aligned}
u' &= 2\mathcal{R}e\langle x_1 - x_2, x_1' - x_2' \rangle \le 2|x_1 - x_2| \cdot |x_1' - x_2'| \\
&= 2|x_1 - x_2|\, |x_1' - f(t,x_1) - (x_2' - f(t,x_2)) + f(t,x_1) - f(t,x_2)| \\
&\le 2|x_1 - x_2|(\epsilon_1 + \epsilon_2 + L|x_1 - x_2|) = 2Lu + 2\epsilon\sqrt{u},
\end{aligned}$$

where $\epsilon = \epsilon_1 + \epsilon_2$.

We want to use the Comparison Theorem to compare u to the solution v of

$$v' = 2Lv + 2\epsilon\sqrt{v}, \quad v(0) = \delta^2 > 0.$$

But

$$\widetilde{f}(v) \equiv 2Lv + 2\epsilon\sqrt{v}$$

is not Lipschitz for $v \in [0, \infty)$. However, for a fixed $\delta > 0$, it is uniformly Lipschitz for $v \in [\delta^2, \infty)$ since

$$\frac{d\widetilde{f}}{dv} = 2L + \frac{\epsilon}{\sqrt{v}} \text{ is bounded for } v \in [\delta^2, \infty),$$

and C^1 functions with bounded derivatives are uniformly Lipschitz:

$$|\widetilde{f}(v_1) - \widetilde{f}(v_2)| = \left| \int_{v_2}^{v_1} \frac{d\widetilde{f}}{dv} dv \right| \leq (\sup |\frac{d}{dv}\widetilde{f}(v)|)|v_1 - v_2|.$$

Although $u(t)$ may leave $[\delta^2, \infty)$, in the proof of the Comparison Theorem we only need $\widetilde{f}$ to be Lipschitz to conclude that $u > v$ cannot occur. Note that since $v' \geq 0$, $v(t)$ stays in $[\delta^2, \infty)$ for $t \geq 0$. So the Comparison Theorem does apply, and we conclude that $u \leq v$ for $t \geq 0$. To solve for v, let $v = w^2$. Then

$$2ww' = (w^2)' = v' = 2Lw^2 + 2\epsilon w.$$

Since $w > 0$, we get $w' = Lw + \epsilon$, $w(0) = \delta$, whose solution is

$$w = \delta e^{Lt} + \frac{\epsilon}{L}(e^{Lt} - 1).$$

Since $|x_1 - x_2| = \sqrt{u} \leq \sqrt{v} = w$, the estimate follows. □

Corollary. For $j \geq 1$, let $x_j(t)$ be a solution of $x_j' = f_j(t, x_j)$, and let $x(t)$ be a solution of $x' = f(t, x)$ on an interval $[a, b]$, where each f_j and f are continuous in t and x and f is Lipschitz in x. Suppose $f_j \to f$ uniformly on $[a, b] \times \mathbb{F}^n$ and $x_j(t_0) \to x(t_0)$ as $j \to \infty$ for some $t_0 \in [a, b]$. Then $x_j(t) \to x(t)$ uniformly on $[a, b]$.

Remark. The domain on which f_j is assumed to converge uniformly to f can be reduced: exercise.

Proof. We have

$$|x_j'(t) - f(t, x_j(t))| \leq |x_j'(t) - f_j(t, x_j(t))| + |f_j(t, x_j(t)) - f(t, x_j(t))|,$$

which can be made less than a given ϵ uniformly in $t \in [a, b]$ by choosing j sufficiently large. So $x(t)$ is an exact solution and $x_j(t)$ is an ϵ-approximate solution of $x' = f(t, x)$ on $[a, b]$. By the Fundamental Estimate,

$$|x_j(t) - x(t)| \leq |x_j(t_0) - x(t_0)|e^{L|t-t_0|} + \frac{\epsilon}{L}(e^{L|t-t_0|} - 1),$$

and thus $|x_j(t) - x(t)| \to 0$ uniformly in $[a, b]$. □

Remark. Also $f_j(t, x_j(t)) \to f(t, x(t))$ uniformly, so $x_j'(t) \to x'(t)$ uniformly. Thus $x_j \to x$ in $C^1([a, b])$ (with norm $\|x\|_{C^1} = \|x\|_\infty + \|x'\|_\infty$).

Continuity with Respect to Parameters and Initial Conditions

Now consider a family of IVPs

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

where $\mu \in \mathbb{F}^m$ is a vector of parameters and $y \in \mathbb{F}^n$. Assume for each value of μ that $f(t, x, \mu)$ is continuous in t and x and Lipschitz in x with Lipschitz constant L locally independent of μ. For each fixed μ, y, this is a standard IVP, which has a solution on some interval about t_0: call it $x(t, \mu, y)$.

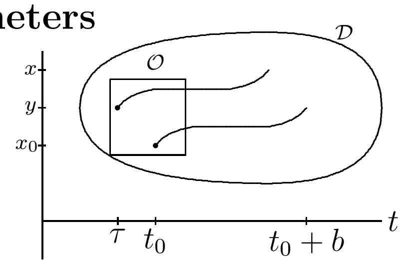
Theorem. If f is continuous in t, x, μ and Lipschitz in x with Lipschitz constant independent of t and μ, then $x(t, \mu, y)$ is continuous in (t, μ, y) jointly.

Remark: See Coddington & Levinson for results saying that if $x(t, \mu_0, y_0)$ exists on $[a, b]$, then $x(t, \mu, y)$ also exists on $[a, b]$ for μ, y near μ_0, y_0.

Proof. The argument of the Corollary shows that x is continuous in μ, y, uniformly in t. Since each $x(t, \mu, y)$ is continuous in t for given μ, y, we can restate this result as saying that the map $(\mu, y) \mapsto x(t, \mu, y)$ from a subset of $\mathbb{F}^m \times \mathbb{F}^n$ into $(C([a, b]), \|\cdot\|_\infty)$ is continuous. Standard arguments now show x is continuous in t, μ, y jointly. □

We have thus established *continuity* of solutions in their dependence on parameters and initial values. We now want to study *differentiability*. By transforming problems of one type into another type, we will be able to reduce our focus to a more restricted case. These transformations are useful for other purposes as well, so we will take a detour to study these transformations.

Transforming "initial conditions" into parameters



Suppose $f(t, x)$ maps an open subset $\mathcal{D} \subset \mathbb{R} \times \mathbb{F}^n$ into $\mathbb{F}^n$, where f is continuous in t and x and locally Lipschitz in x on $\mathcal{D}$. Consider the IVP $x' = f(t, x)$, $x(\tau) = y$, where $(\tau, y) \in \mathcal{D}$. Think of τ as a *variable initial time* t_0, and y as a *variable initial value* x_0. Viewing (τ, y) as parameters, let $x(t, \tau, y)$ be the solution of this IVP.

Remark. One can show that if $(t_0, x_0) \in \mathcal{D}$ and $x(t, t_0, x_0)$ exists in $\mathcal{D}$ on a time interval $[t_0, t_0 + b]$, then for (τ, y) in some sufficiently small open neighborhood $\mathcal{O}$ of (t_0, x_0), the solution $x(t, \tau, y)$ exists on $I_{\tau,t_0} \equiv [\min(\tau, t_0), \max(\tau, t_0) + b]$ (which contains $[t_0, t_0 + b]$ and $[\tau, \tau + b]$), and moreover $\{(t, x(t, \tau, y)) : t \in I_{\tau,t_0}, (\tau, y) \in \mathcal{O}\}$ is contained in some compact subset of $\mathcal{D}$.

Define

$$\widetilde{f}(t, x, \tau, y) = f(\tau + t, x + y) \quad \text{and} \quad \widetilde{x}(t, \tau, y) = x(\tau + t, \tau, y) - y.$$

Then $\widetilde{x}(t, \tau, y)$ is a solution of the IVP

$$\widetilde{x}' = \widetilde{f}(t, \widetilde{x}, \tau, y), \quad \widetilde{x}(0) = 0$$

with $n+1$ parameters (τ, y) and fixed initial conditions. This IVP is equivalent to the original IVP $x' = f(t,x)$, $x(\tau) = y$.

Remarks.

(1) $\widetilde{f}$ is continuous in t, x, τ, y and locally Lipschitz in x in the open set

$$\mathcal{W} \equiv \{(t,x,\tau,y) : (\tau+t, x+y) \in \mathcal{D},\ (\tau,y) \in \mathcal{D}\} \subset \mathbb{R} \times \mathbb{F}^n \times \mathbb{R} \times \mathbb{F}^n.$$

(2) If f is C^k in t, x in $\mathcal{D}$, then $\widetilde{f}$ is C^k in t, x, τ, y in $\mathcal{W}$.

Transforming Parameters into "Initial Conditions"

Suppose $f(t,x,\mu)$ is continuous in t, x, μ and locally Lipschitz in x on an open set $\mathcal{W} \subset \mathbb{R} \times \mathbb{F}^n \times \mathbb{F}^m$. Consider the IVP $x' = f(t,x,\mu)$, $x(t_0) = x_0$, with solution $x(t,\mu)$. Introduce a new variable $z \in \mathbb{F}^m$. Set

$$\widetilde{x} = \begin{bmatrix} x \\ z \end{bmatrix} \in \mathbb{F}^{n+m} \quad \text{and} \quad \widetilde{f}(t,\widetilde{x}) = \begin{bmatrix} f(t,x,z) \\ 0 \end{bmatrix} \in \mathbb{F}^{n+m}.$$

Consider the IVP

$$\widetilde{x}' = \widetilde{f}(t,\widetilde{x}), \quad \widetilde{x}(t_0) = \begin{bmatrix} x_0 \\ \mu \end{bmatrix},$$

(i.e., $x' = f(t,x,z)$, $z' = 0$, $x(t_0) = x_0$, $z(t_0) = \mu$), with solution $\widetilde{x}(t,\mu)$. Then $z(t) \equiv \mu$, so the solution is

$$\widetilde{x}(t,\mu) = \begin{bmatrix} x(t,\mu) \\ \mu \end{bmatrix},$$

and the two IVPs are equivalent.

Remarks.

(1) If f is continuous in t, x, μ and locally Lipschitz in x and μ (jointly), then $\widetilde{f}$ is continuous in $t, \widetilde{x}$ and locally Lipschitz in $\widetilde{x}$. (However, for this specific $\widetilde{f}$, Lipschitz continuity in z is not needed for uniqueness.)

(2) If f is C^k in t, x, μ, then $\widetilde{f}$ is C^k in $t, \widetilde{x}$.

(3) One can show that if $(t_0, x_0, \mu_0) \in \mathcal{W}$ and the solution $x(t,\mu_0)$ exists in $\mathcal{W}$ on a time interval $[t_0, t_0+b]$, then for μ in some sufficiently small open neighborhood $\mathcal{U}$ of μ_0 in $\mathbb{F}^m$, the solution $x(t,\mu)$ exists on $[t_0, t_0+b]$, and, moreover, the set

$$\{(t, x(t,\mu), \mu) : t \in [t_0, t_0+b], \mu \in \mathcal{U}\}$$

is contained in some compact subset of $\mathcal{W}$.

(4) An IVP $x' = f(t,x,\mu)$, $x(\tau) = y$ with parameters $\mu \in \mathbb{F}^m$ and variable initial values $\tau \in \mathbb{R}$, $y \in \mathbb{F}^n$ can be transformed similarly into either IVPs with variable IC and no parameters in the DE or IVPs with fixed IC and variable parameters in the DE.

The Equation of Variation

A main tool in proving the differentiability of x when f is C^k is the *equation of variation*, commonly also called the *linearization* of the DE (or the *linearized* DE), or the *perturbation equation*, etc. This is a *linear* DE for the (leading term of the) perturbation in the solution x due to, e.g., a perturbation in the parameters. Or it can be viewed as a linear DE for the derivative of x with respect, e.g., to the parameter(s).

The easiest case to describe is when there is one real parameter s in the DE; we will also allow the initial value $x(t_0)$ to depend on s. Let $x(t, s)$ be the solution of the IVP

$$x' = f(t, x, s), \quad x(t_0) = x_0(s),$$

where f is continuous in t, x, s and C^1 in x, s, and $x_0(s)$ is C^1 in s. If $x(t, s)$ is differentiable in s, then (formally) differentiating the DE and the IC with respect to s gives the following IVP for $\frac{\partial x}{\partial s}(t, s)$:

$$\begin{aligned} \left(\frac{\partial x}{\partial s}\right)' &= D_x f\left(t, x(t, s), s\right) \frac{\partial x}{\partial s} + D_s f\left(t, x(t, s), s\right) \\ \frac{\partial x}{\partial s}(t_0) &= \frac{dx_0}{ds} \end{aligned}$$

where $'$ denotes $\frac{d}{dt}$, $D_x f$ is the $n \times n$ Jacobian matrix $\frac{\partial f_i}{\partial x_j}$, and $D_s f$ is the $n \times 1$ derivative $\frac{\partial f_i}{\partial s}$. Evaluating at a fixed s_0, we get that $\frac{\partial x}{\partial s}(t, s_0)$ satisfies

$$\left(\left.\frac{\partial x}{\partial s}\right|_{s_0}\right)' = D_x f(t, x(t, s_0), s_0) \left.\frac{\partial x}{\partial s}\right|_{s_0} + D_s f(t, x(t, s_0), s_0), \quad \left.\frac{\partial x}{\partial s}\right|_{s_0}(t_0) = \left.\frac{dx_0}{ds}\right|_{s_0}.$$

This is a *linear* DE for $\left.\frac{\partial x}{\partial s}\right|_{s_0}$ of the form $z' = A(t)z + b(t)$, where both the coefficient matrix $A(t) = D_x f(t, x(t, s_0), s_0)$ and the inhomogeneous term $b(t) = D_s f(t, x(t, s_0), s_0)$ are known if f and $x(t, s_0)$ are known.

The theoretical view is this: *if* $x(t, s)$ is C^1 in s, then $\left.\frac{\partial x}{\partial s}\right|_{s_0}$ satisfies this linear DE. We want to prove that $x(t, s)$ is C^1. We *start* from this linear DE, which has a solution by our previous theory. This "gets our hands on" what ought to be $\left.\frac{\partial x}{\partial s}\right|_{s_0}$. We then prove (see theorem below) that

$$\frac{x(t, s_0 + \Delta s) - x(t, s_0)}{\Delta s}$$

converges as $\Delta s \to 0$ to this solution, which therefore must be $\left.\frac{\partial x}{\partial s}\right|_{s_0}$. It follows (from continuity with respect to parameters) that $\frac{\partial x}{\partial s}$ is continuous in t and s. The original DE implies $\frac{\partial x}{\partial t}$ is continuous in t and s. We conclude then that $x(t, s)$ is C^1 with respect to t and s jointly.

An alternate view of the equation of variation comes from the "tangent line approximation": formally, for s near s_0, we expect

$$x(t, s) \approx x(t, s_0) + (s - s_0)\frac{\partial x}{\partial s}(t, s_0),$$

with error $O(|s-s_0|^2)$. Setting $\Delta s = s - s_0$ and $\Delta x = x(t,s) - x(t,s_0)$, we expect $\Delta x \approx \frac{\partial x}{\partial s}(t,s_0)\Delta s$. We could either multiply the linear DE above by Δs, or proceed formally as follows: suppose $x(t,s) = x(t,s_0) + \Delta x(t,s)$ (which we abbreviate as $x = x_{s_0} + \Delta x$), and suppose $|\Delta x| = O(|\Delta s|)$ where $\Delta s = s - s_0$. Substitute into the DE, and formally drop terms of order $|\Delta s|^2$:

$$\begin{aligned}(x_{s_0} + \Delta x)' &= f(t, x_{s_0} + \Delta x, s_0 + \Delta s) \\ &= f(t, x_{s_0}, s_0) + D_x f(t, x_{s_0}, s_0)\Delta x + D_s f(t, x_{s_0}, s_0)\Delta s + \underbrace{\boxed{O(|\Delta s|^2)}}_{\text{neglect}}\end{aligned}$$

so, since $x_{s_0}' = f(t, x_{s_0}, s_0)$,

$$(\Delta x)' = D_x f(t, x_{s_0}, s_0)\Delta x + D_s f(t, x_{s_0}, s_0)\Delta s.$$

(This is equivalent to Δs times $\left(\frac{\partial x}{\partial s}\right)' = D_x f \left(\frac{\partial x}{\partial s}\right) + D_s f$, when we take Δx to mean the "tangent line approximation" $\frac{\partial x}{\partial s}\Delta s$.)

Example. Consider the IVP $x' = f(t,x,\mu)$ where $\mu \in \mathbb{F}^m$ with fixed IC $x(t_0) = x_0$. Then for $1 \le k \le m$,

$$\left(\frac{\partial x}{\partial \mu_k}\right)' = D_x f(t, x(t,\mu), \mu)\frac{\partial x}{\partial \mu_k} + D_{\mu_k} f(t, x(t,\mu), \mu)$$

is the equation of variation with respect to μ_k, with IC $\frac{\partial x}{\partial \mu_k}(t_0) = 0$. Put together in matrix form,

$$(D_\mu x)' = (D_x f)(D_\mu x) + D_\mu f, \quad D_\mu x(t_0) = 0.$$

Here, $D_\mu x$ is the $n \times m$ Jacobian matrix $\frac{\partial x_i}{\partial \mu_j}$, $D_\mu f$ is the $n \times m$ Jacobian matrix $\frac{\partial f_i}{\partial \mu_j}$, and as above $D_x f$ is the $n \times n$ Jacobian matrix $\frac{\partial f_i}{\partial x_j}$.

In the above example, the initial condition $x(t_0) = x_0$ was independent of the parameter μ, so the initial condition for the derivative was homogeneous: $\frac{\partial x}{\partial \mu_k}(t_0) = 0$. If the initial condition depends on the parameter, then of course it must be differentiated as well. For example, in the case in which the initial value is the parameter, so that one is solving $x' = f(t,x), \quad x(t_0) = y$, the initial condition for the derivative becomes $\frac{\partial x}{\partial y_k}(t_0) = e_k$.

Differentiability

We can now prove differentiability. From our discussion showing that dependence on parameters can be transformed into IC, it will suffice to prove the following.

Theorem. Suppose f is continuous in t, x and C^1 in x, and $x(t,y)$ is the solution of the IVP $x' = f(t,x)$, $x(t_0) = y$ (say on an interval $[a,b]$ containing t_0 for y in some closed ball $B = \{y \in \mathbb{F}^n : |y - x_0| \le r\}$). Then x is a C^1 function of t and y on $[a,b] \times B$.

Proof. By the previous theorem, $x(t,y)$ is continuous in $[a,b] \times B$, so

$$K \equiv \{(t, x(t,y)) : t \in [a,b], y \in B\}$$

is compact, and thus f is uniformly Lipschitz in x on K, say with Lipschitz constant L. From the DE, $\frac{\partial x}{\partial t}(t,y) = f(t, x(t,y))$, and so $\frac{\partial x}{\partial t}(t,y)$ is continuous on $[a,b] \times B$. Now fix j with $1 \leq j \leq n$. If $\frac{\partial x}{\partial y_j}$ exists, it must satisfy the linear IVP

$$(*) \qquad z' = A(t,y)z \text{ on } [a,b], \quad z(t_0) = e_j,$$

where $A(t,y) = D_x f(t, x(t,y))$. Let $z(t,y)$ be the solution of the IVP $(*)$. Since $A(t,y)$ is continuous on the compact set $[a,b] \times B$, it is bounded on $[a,b] \times B$. Let $M > 0$ be such a bound, i.e. $|A(t,y)| \leq M$ for all $(t,y) \in [a,b] \times B$. The DE in $(*)$ is linear, with RHS uniformly Lipschitz in z with Lipschitz constant M. By the global existence theorem for linear systems and the continuity theorem, $z(t,y)$ exists and is continuous on $[a,b] \times B$. For $h \in \mathbb{R}$ with $|h|$ small [strictly speaking, for fixed $y \in \text{int}(B)$, assume $|h| < r - |y - x_0|$ so that $B_h(y) \subset \text{int}(B)$], set

$$\theta(t,y,h) = \frac{x(t, y + he_j) - x(t,y)}{h}.$$

By the Fundamental Estimate (applied to $x' = f(t,x)$ with $\delta = |h|$ and $\epsilon_1 = \epsilon_2 = 0$),

$$|x(t, y + he_j) - x(t,y)| \leq |x(t_0, y + he_j) - x(t_0, y)|e^{L|b-a|} = |h|e^{L|b-a|}$$

so $|\theta(t,y,h)| \leq e^{L|b-a|}$. Also by the DE,

$$\theta'(t,y,h) = \frac{f(t, x(t, y + he_j)) - f(t, x(t,y))}{h}.$$

Denote by $\omega(\delta)$ the modulus of continuity of $D_x f$ (with respect to x) on K, so that

$$\omega(\delta) = \sup\{|D_x f(t, x_1) - D_x f(t, x_2)| : (t, x_1) \in K, (t, x_2) \in K, |x_1 - x_2| \leq \delta\}.$$

Since $D_x f$ is continuous on the compact set K, it is uniformly continuous on K, so $\omega(\delta) \to 0$ as $\delta \to 0^+$. Clearly $\omega(\delta)$ is an increasing function of δ. Also, whenever the line segment from (t, x_1) to (t, x_2) stays in K,

$$\begin{aligned} &|f(t,x_2) - [f(t,x_1) + D_x f(t,x_1)(x_2 - x_1)]| \\ &\quad = \left| \int_0^1 (D_x f(t, x_1 + s(x_2 - x_1)) - D_x f(t, x_1))(x_2 - x_1) ds \right| \leq \omega(|x_2 - x_1|) \cdot |x_2 - x_1|. \end{aligned}$$

We apply this bound with $x_1 = x(t,y)$ and $x_2 = x(t, y + he_j)$, for which the line segment is in K if $|h|$ is small enough, to obtain

$$\begin{aligned} &|\theta'(t,y,h) - A(t,y)\theta(t,y,h)| \\ &\quad = \frac{1}{|h|}|f(t, x(t, y + he_j)) - f(t, x(t,y)) - D_x f(t, x(t,y))(x(t, y + he_j) - x(t,y))| \\ &\quad \leq \frac{1}{|h|}\omega(|x(t, y + he_j) - x(t,y)|)|x(t, y + he_j) - x(t,y)| \leq \omega(|h|e^{L|b-a|})e^{L|b-a|} \end{aligned}$$

(since $|x(t, y + he_j) - x(t,y)| \leq |h|e^{L|b-a|}$). Set

$$\epsilon(h) = \omega(|h|e^{L|b-a|})e^{L|b-a|};$$

then $\epsilon(h) \to 0$ as $h \to 0$. We have shown that $\theta(t, y, h)$ is an $\epsilon(h)$-approximate solution to $z' = A(t, y)z$. Moreover, $\theta(t_0, y, h) = e_j$. So by the Fundamental Estimate applied to $(*)$, with Lipschitz constant M,

$$|\theta(t, y, h) - z(t, y)| \le \frac{\epsilon(h)}{M}(e^{M|b-a|} - 1).$$

This shows that $\lim_{h\to 0} \theta(t, y, h) = z(t, y)$ (including the existence of the limit). Thus $\frac{\partial x}{\partial y_j}(t, y) = z(t, y)$, which is continuous in $[a, b] \times B$. We conclude that $x(t, y)$ is C^1 in t, y on $[a, b] \times B$. □

We obtain as a corollary the main differentiability theorem.

Theorem. Suppose $f(t, x, \mu)$ is C^k in (t, x, μ) for some $k \ge 1$, and $x(t, \mu, \tau, y)$ is the solution of the IVP $x' = f(t, x, \mu)$, $x(\tau) = y$. Then $x(t, \mu, \tau, y)$ is a C^k function of (t, μ, τ, y).

Proof. By the transformations described previously, it suffices to consider the solution $x(t, y)$ to the IVP $x' = f(t, x)$, $x(t_0) = y$. The case $k = 1$ follows from the previous theorem. Suppose $k > 1$ and the result is true for $k - 1$. Then $\frac{\partial x}{\partial y_j}$ satisfies $(*)$ above with $A(t, y) \in C^{k-1}$, and $\frac{\partial x}{\partial t}$ satisfies

$$w' = D_t f(t, x(t, y)) + D_x f(t, x(t, y)) f(t, x(t, y)), \quad w(t_0) = f(t_0, x(t_0, y)).$$

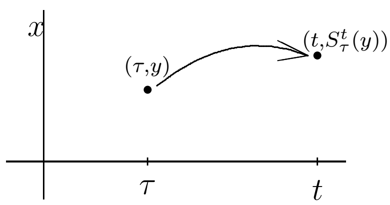
By induction, $\frac{\partial x}{\partial t}$ and $\frac{\partial x}{\partial y_j}$ (for $1 \le j \le n$) $\in C^{k-1}$; thus $x \in C^k$. □

Nonlinear Solution Operator

Consider the DE $x' = f(t, x)$ where f is continuous in t, x, Lipschitz in x. Let $x(t, \tau, y)$ be the solution of the IVP $x' = f(t, x)$, $x(\tau) = y$ (say, on an interval $t \in [a, b]$, where $\tau \in [a, b]$, and we assume all solutions we consider exist on $[a, b]$). For a *fixed* pair τ (the "initial time") and t (the "final time") in $[a, b]$, define a map S_τ^t from an open set $\mathcal{U} \subset \mathbb{F}^n$ into $\mathbb{F}^n$ by

$$S_\tau^t(y) = x(t, \tau, y),$$

so that S_τ^t maps the initial value y (at time τ) into the solution $x(t, \tau, y)$ at time t. By the continuity of $x(t, \tau, y)$ in (t, τ, y), S_τ^t is continuous on the open set $\mathcal{U} \subset \mathbb{F}^n$. By uniqueness, S_τ^t is invertible, and its inverse is S_t^τ (say, defined on $\mathcal{W} \equiv S_\tau^t(\mathcal{U})$), which is also continuous. So $S_\tau^t : \mathcal{U} \to \mathcal{W}$ is a *homeomorphism* (i.e., one-to-one, onto, continuous, with continuous inverse). [Note: the set $\mathcal{W}$ depends on t.] If f is C^k in t, x, then by our differentiability theorem, $S_\tau^t : \mathcal{U} \to \mathcal{W}$ is a C^k diffeomorphism (i.e., S_τ^t and its inverse S_t^τ are both bijective and C^k).



Remarks.

(1) If f is at least C^1, then the chain rule applied to $I = S_t^\tau \circ S_\tau^t$ (for fixed τ, t) implies that the Jacobian matrix $D_y S_\tau^t$ is invertible at each $y \in \mathcal{U}$. We will see another way to show this in the following material on linear systems. [Note: $S_\tau^t(y) = x(t, \tau, y)$, so for fixed τ, t, the ij^{th} element of $D_y S_\tau^t$ is $\frac{\partial x_i}{\partial y_j}(t, \tau, y)$.]

(2) Conversely, the inverse function theorem implies that any injective C^k mapping on $\mathcal{U}$ whose Jacobian matrix is invertible at each $y \in \mathcal{U}$ is a C^k diffeomorphism.

(3) Caution: For nonlinear f, S_τ^t is generally a *nonlinear* map.

Group Property of the Nonlinear Solution Operator

Consider the two-parameter family of operators $\{S_\tau^t : \tau, t \in [a,b]\}$. For simplicity, assume they all are defined for all $y \in \mathbb{F}^n$. (Otherwise, some consistent choice of domains must be made, e.g., let $\mathcal{U}_a$ be an open subset of $\mathbb{F}^n$, and define $\mathcal{U}_\tau = S_a^\tau(\mathcal{U}_a)$ for $\tau \in [a,b]$. Choose the domain of S_τ^t to be $\mathcal{U}_\tau$. Then $S_\tau^t(\mathcal{U}_\tau) = \mathcal{U}_t$.) This two-parameter family of operators has the following "group properties":

(1) $S_\tau^\tau = I$ for all $\tau \in [a,b]$, and

(2) $S_t^\sigma \circ S_\tau^t = S_\tau^\sigma$ for all $\tau, t, \sigma \in [a,b]$.

Stated in words, mapping the value of a solution at time τ into its value at time t, and then mapping this value at time t into the value of the solution at time σ is equivalent to mapping the value at time τ directly into the value at time σ.

Special Case — Autonomous Systems

□

For an autonomous system $x' = f(x)$, if τ_1, t_1, τ_2, t_2 satisfy $t_1 - \tau_1 = t_2 - \tau_2$, then $S_{\tau_1}^{t_1} = S_{\tau_2}^{t_2}$ (exercise). So we can define a one-parameter family of operators S_σ where $S_\sigma = S_\tau^t$ (for any τ, t with $t - \tau = \sigma$). The single parameter σ here is "elapsed time" (positive or negative) $t - \tau$, as opposed to the two parameters τ ("initial time") and t ("final time") above. The two properties become

(1′) $S_0 = I$

(2′) $S_{\sigma_2} \circ S_{\sigma_1} = S_{\sigma_2 + \sigma_1}$.