Math 555 Linear Analysis
Winter 2009
Lecture Notes

Contents

Ordinary Differential Equations (ODEs)

ODEs

Let $\mathbb{F}$ be $\mathbb{R}$ on $\mathbb{C}$. Throughout this discussion, $|\cdot|$ will denote the Euclidean norm (i.e ℓ^2-norm) on $\mathbb{F}^n$ (so $\|\cdot\|$ is free to be used for norms on function spaces). An ODE is an equation of the form

$$g(t, x, x', \ldots, x^{(m)}) = 0$$

where g maps a subset of $\mathbb{R} \times (\mathbb{F}^n)^{m+1}$ into $\mathbb{F}^n$. A *solution* of this ODE on an interval $I \subset \mathbb{R}$ is a function $x : I \to \mathbb{F}^n$ for which $x', x'', \ldots, x^{(m)}$ exist at each $t \in I$, and

$$(\forall\, t \in I) \qquad g(t, x(t), x'(t), \ldots, x^{(m)}(t)) = 0\ .$$

We will focus on the case where $x^{(m)}$ can be solved for explicitly; i.e., the equation takes the form

$$x^{(m)} = f(t, x, x', \ldots, x^{(m-1)}),$$

and where the function f mapping a subset of $\mathbb{R} \times (\mathbb{F}^n)^m$ into $\mathbb{F}^n$ is continuous. This equation is called an m^{th}*-order* $n \times n$ system of ODE's. Note that if x is a solution defined on an interval $I \subset \mathbb{R}$ then the existence of $x^{(m)}$ on I (including one-sided limits at the endpoints of I) implies that $x \in C^{m-1}(I)$, and then the equation implies $x^{(m)} \in C(I)$, so $x \in C^m(I)$.

Reduction to First-Order Systems

Every m^{th}-order $n \times n$ system of ODE's is equivalent to a first-order $mn \times mn$ system of ODE's. Defining

$$y_j(t) = x^{(j-1)}(t) \quad \in \mathbb{F}^n \quad \text{for} \quad 1 \le j \le m$$

and

$$y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix} \in \mathbb{F}^{mn},$$

the system

$$x^{(m)} = f(t, x, \ldots, x^{(m-1)})$$

is equivalent to the first-order $mn \times mn$ system

$$y' = \begin{bmatrix} y_2 \\ y_3 \\ \vdots \\ y_m \\ f(t, y_1, \ldots, y_m) \end{bmatrix}.$$

Relabeling if necessary, we will focus on first-order $n \times n$ systems of the form $x' = f(t, x)$ where f maps a subset of $\mathbb{R} \times \mathbb{F}^n$ into $\mathbb{F}^n$ and f is continuous.

Example. Consider the $n \times n$ system $x'(t) = f(t)$ where $f : I \to \mathbb{F}^n$ is continuous on an interval $I \subset \mathbb{R}$. (Here f is independent of x.) Then calculus shows that for a fixed $t_0 \in I$, the general solution of the ODE (i.e., a form representing all possible solutions) is

$$x(t) = c + \int_{t_0}^{t} f(s)ds,$$

where $c \in \mathbb{F}^n$ is an arbitrary constant vector (i.e., $c_1, \ldots, c_n$ are n arbitrary constants in $\mathbb{F}$).

Provided f satisfies a Lipschitz condition (to be discussed soon), the general solution of a first-order system $x' = f(t, x)$ involves n arbitrary constants in $\mathbb{F}$ [or an arbitrary vector in $\mathbb{F}^n$] (whether or not we can express the general solution explicitly), so n scalar conditions [or one vector condition] must be given to specify a particular solution. For the example above, clearly giving $x(t_0) = x_0$ (for a known constant vector x_0) determines c, namely, $c = x_0$. In general, specifying $x(t_0) = x_0$ (these are called *initial conditions* (IC), even if t_0 is not the left endpoint of the t-interval I) determines a particular solution of the ODE.

Initial-Value Problems (IVP's) for First-order Systems

An IVP for the first-order system is the differential equation

$$DE: \qquad x' = f(t, x),$$

together with initial conditions

$$IC: \qquad x(t_0) = x_0 \ .$$

A solution to the IVP is a solution $x(t)$ of the DE defined on an interval I containing t_0, which also satisfies the IC, i.e., for which $x(t_0) = x_0$.

Examples:

(1) Let $n = 1$. The solution of the IVP:

$$\begin{aligned} DE: &\qquad x' = x^2 \\ IC: &\qquad x(1) = 1 \end{aligned}$$

is $x(t) = \frac{1}{2-t}$, which blows up as $t \to 2$. So even if f is C^∞ on all of $\mathbb{R} \times \mathbb{F}^n$, solutions of an IVP do not necessarily exist for all time t.

(2) Let $n = 1$. Consider the IVP:

$$\begin{aligned} DE: \quad & x' = 2\sqrt{|x|} \\ IC: \quad & x(0) = 0\ . \end{aligned}$$

For any $c \geq 0$, define $x_c(t) = 0$ for $t \leq c$ and $x_c(t) = (t-c)^2$ for $t \geq c$. Then every $x_c(t)$ for $c \geq 0$ is a solution of this IVP. So in general for continuous $f(t, x)$, IVP's may have non-unique solutions. (The difficulty here is that $f(t, x) = 2\sqrt{|x|}$ is not Lipschitz near $x = 0$.)

An Integral Equation Equivalent to an IVP

Suppose $x(t) \in C^1(I)$ is a solution of the IVP:

$$\begin{aligned} DE: \quad & x' = f(t, x) \\ IC: \quad & x(t_0) = x_0 \end{aligned}$$

defined on an interval $I \subset \mathbb{R}$ with $t_0 \in I$. Then for all $t \in I$,

$$\begin{aligned} x(t) &= x(t_0) + \int_{t_0}^{t} x'(s)ds \\ &= x_0 + \int_{t_0}^{t} f(s, x(s))ds, \end{aligned}$$

so $x(t)$ is also a solution of the *integral equation*

(IE) $$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds \qquad (t \in I).$$

Conversely, suppose $x(t) \in C(I)$ is a solution of the integral equation (IE). Then $f(t, x(t)) \in C(I)$, so

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds \in C^1(I)$$

and $x'(t) = f(t, x(t))$ by the Fundamental Theorem of Calculus. So x is a C^1 solution of the DE on I, and clearly $x(t_0) = x_0$, so x is a solution of the IVP. We have shown:

Proposition. On an interval I containing t_0, x is a solution of the IVP: $DE: x' = f(t, x)$; $IC: x(t_0) = x_0$ (where f is continuous) with $x \in C^1(I)$ if and only if x is a solution of the integral equation (IE) on I with $x \in C(I)$.

The integral equation (IE) is a useful way to study the IVP. We can deal with the function space of continuous functions on I without having to be concerned about differentiability: continuous solutions of (IE) are automatically C^1. Moreover, the initial condition is built into the integral equation.

We will solve (IE) using a fixed-point formulation.

Definition. Let (X, d) be a metric space, and suppose $g : X \to X$. We say that g is a *contraction* [on X] if there exists $c < 1$ such that

$$(\forall\, x, y \in X) \qquad d(g(x), g(y)) \le cd(x, y)$$

(c is sometimes called the contraction constant). A point $x_* \in X$ for which

$$g(x_*) = x_*$$

is called a *fixed point* of g.

Theorem.(**The Contraction Mapping Fixed-Point Theorem**)
Let (X, d) be a *complete* metric space and $g : X \to X$ be a contraction (with contraction constant $c < 1$). Then g has a unique fixed point $x_* \in X$. Moreover, for any $x_0 \in X$, if we generate the sequence $\{x_k\}$ iteratively by *functional iteration*

$$x_{k+1} = g(x_k) \quad \text{for} \quad k \ge 0$$

(sometimes called *fixed-point iteration*), then $x_k \to x_*$.

Proof. Fix $x_0 \in X$, and generate $\{x_k\}$ by $x_{k+1} = g(x_k)$ Then for $k \ge 1$,

$$d(x_{k+1}, x_k) = d(g(x_k), g(x_{k-1})) \le cd(x_k, x_{k-1}).$$

By induction

$$d(x_{k+1}, x_k) \le c^k d(x_1, x_0).$$

So for $n < m$,

$$\begin{aligned} d(x_m, x_n) &\le \sum_{j=n}^{m-1} d(x_{j+1}, x_j) \le \left(\sum_{j=n}^{m-1} c^j\right) d(x_1, x_0) \\ &\le \left(\sum_{j=n}^{\infty} c^j\right) d(x_1, x_0) = \frac{c^n}{1-c} d(x_1, x_0). \end{aligned}$$

Since $c^n \to 0$ as $n \to \infty$, $\{x_k\}$ is Cauchy. Since X is complete, $x_k \to x_*$ for some $x_k \in X$. Since g is a contraction, clearly g is continuous, so

$$g(x_*) = g(\lim x_k) = \lim g(x_k) = \lim x_{k+1} = x_*,$$

so x_* is a fixed point. If x and y are two fixed points of g in X, then

$$d(x, y) = d(g(x), g(y)) \le cd(x, y),$$

so $(1 - c)d(x, y) \le 0$, and thus $d(x, y) = 0$ and $x = y$. So g has a unique fixed point. □ □

Applications.

(1) Iterative methods for linear systems.

(2) *The Inverse Function Theorem* If $\Phi : N \to \mathbb{R}^n$ is a C^1 mapping on a neighborhood $N \subset \mathbb{R}^n$ of $x_0 \in \mathbb{R}^n$ satisfying $\Phi(x_0) = y_0$ and $\Phi'(x_0) \in \mathbb{R}^{n\times n}$ is invertible, then there exist neighborhoods $N_0 \subset N$ of x_0 and M_0 of y_0 and a C^1 mapping $\Psi : M_0 \to N_0$ for which $\Phi[N_0] = M_0$ and $\Phi \circ \Psi$ and $\Psi \circ \Phi$ are the identity mappings on M_0 and N_0, respectively.

Remark. Applying the Contraction Mapping Fixed-Point Theorem (C.M.F.-P.T.) to a function usually requires two steps:

(1) showing there is a topologically complete set S for which $g(S) \subset S$, and

(2) showing that g is a contraction on S.

To apply the C.M.F.-P.T. to the integral equation (IE), we need a further condition on the function $f(t, x)$.

Definition. Let $I \subset \mathbb{R}$ be an interval and $\Omega \subset \mathbb{F}^n$. We say that $f(t, x)$ mapping $I \times \Omega$ into $\mathbb{F}^n$ is *uniformly Lipschitz continuous with respect to* x if there is a constant L (called the *Lipschitz constant*) for which

$$(\forall\, t \in I)(\forall\, x, y \in \Omega) \qquad |f(t, x) - f(t, y)| \le L|x - y|\ .$$

We say that f is in (C, Lip) on $I \times \Omega$ if f is continuous on $I \times \Omega$ and f is uniformly Lipschitz continuous with respect to x on $I \times \Omega$.

For simplicity, we will consider intervals $I \subset \mathbb{R}$ for which t_0 is the left endpoint. Virtually identical arguments hold if t_0 is the right endpoint of I, or if t_0 is in the interior of I. (See Coddington & Levinson.)

Theorem (Local Existence and Uniqueness for (IE) for Lipschitz f)
Let $I = [t_0, t_0 + \beta]$ and $\Omega = \overline{B_r(x_0)} = \{x \in \mathbb{F}^n : |x - x_0| \le r\}$, and suppose $f(t, x)$ is in (C, Lip) on $I \times \Omega$. Then there exists $\alpha \in (0, \beta]$ for which there is a unique solution of the integral equation

(IE) $$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

in $C(I_\alpha)$ where $I_\alpha = [t_0, t_0 + \alpha]$. Moreover, we can choose α to be any positive number satisfying

$$\alpha \le \beta,\ \alpha \le \frac{r}{M}, \quad \text{and} \quad \alpha < \frac{1}{L}, \quad \text{where} \quad M = \max_{(t,x)\in I\times\Omega} |f(t, x)|$$

and L is the Lipschitz constant for f in $I \times \Omega$.

Proof. For any $\alpha \in (0, \beta]$, let $\|\cdot\|_\infty$ denote the max-norm on $C(I_\alpha)$:

$$\text{for} \quad x \in C(I_\alpha), \quad \|x\|_\infty = \max_{t_0 \le t \le t_0+\alpha} |x(t)|\ .$$

Although this norm clearly depends on α, we do not include α in the notation. Let $\widetilde{x}_0$ denote the constant function $\widetilde{x}_0(t) \equiv x_0$ in $C(I_\alpha)$. For $\rho > 0$ let

$$X_{\alpha,\rho} = \{x \in C(I_\alpha) : \|x - \widetilde{x}_0\|_\infty \le \rho\}.$$

Then $X_{\alpha,\rho}$ is a complete metric space since it is a closed subset of the Banach space $(C(I_\alpha), \|\cdot\|_\infty)$. For any $\alpha \in (0, \beta]$, define $g : X_{\alpha,r} \to C(I_\alpha)$ by

$$(g(x))(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds :$$

g is well-defined on $X_{\alpha,r}$ and $g(x) \in C(I_\alpha)$ for $x \in X_{\alpha,r}$ since f is continuous on $I \times \overline{B_r(x_0)}$. Fixed points of g are solutions of the integral equation (IE).

Claim. Suppose $\alpha \in (0, \beta]$, $\alpha \le \frac{r}{M}$, and $\alpha < \frac{1}{L}$. Then g maps $X_{\alpha,r}$ into itself and g is a contraction on $X_{\alpha,r}$.

Proof of Claim: If $x \in X_{\alpha,r}$, then for $t \in I_\alpha$,

$$|(g(x))(t) - x_0| \le \int_{t_0}^{t} |f(s, x(s))|ds \le M\alpha \le r,$$

so $g : X_{\alpha,r} \to X_{\alpha,r}$. If $x, y \in X_{\alpha,r}$, then for $t \in I_\alpha$,

$$\begin{aligned} |(g(x))(t) - (g(y))(t)| &\le \int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))|ds \\ &\le \int_{t_0}^{t} L|x(s) - y(s)|ds \\ &\le L\alpha\|x - y\|_\infty, \end{aligned}$$

so

$$\|g(x) - g(y)\|_\infty \le L\alpha\|x - y\|_\infty, \quad \text{and} \quad L\alpha < 1.$$

So by the C.M.F.-P.T., for α satisfying $0 < \alpha \le \beta$, $\alpha \le \frac{r}{M}$, and $\alpha < \frac{1}{L}$, g has a unique fixed point in $X_{\alpha,r}$, and thus the integral equation (IE) has a unique solution $x_*(t)$ in $X_{\alpha,r} = \{x \in C(I_\alpha) : \|x - \widetilde{x}_0\|_\infty \le r\}$. This is *almost* the conclusion of the Theorem, except we haven't shown x_* is the only solution in all of $C(I_\alpha)$. This uniqueness is better handled by techniques we will study soon, but we can still eke out a proof here. (We could say that f is only given on $I \times \overline{B_r(x_0)}$, but f can have a continuous extension to $I \times \mathbb{F}^n$.) Fix such an α. Then clearly for $0 < \gamma \le \alpha$, $x_*|_{I_\gamma}$ is the unique fixed point of g on $X_{\gamma,r}$. Suppose $y \in C(I_\alpha)$ is a solution of (IE) on I_α (using perhaps an extension of f) with $y \not\equiv x_\alpha$ on I_α. Let

$$\gamma_1 = \inf\{\gamma \in (0, \alpha] : y(t_0 + \gamma) \ne x_*(t_0 + \gamma)\}.$$

By continuity, $\gamma_1 < \alpha$. Since $y(t_0) = x_0$, continuity implies

$$\exists\, \gamma_0 \in (0, \alpha] \ni y|_{I_{\gamma_0}} \in X_{\gamma_0,r},$$

and thus $y(t) \equiv x_*(t)$ on I_{γ_0}. So $0 < \gamma_1 < \alpha$. Since $y(t) \equiv x_*(t)$ on I_{γ_1}, $y|_{I_{\gamma_1}} \in X_{\gamma_1,r}$. Let $\rho = M_{\gamma_1}$; then $\rho < M\alpha \leq r$. For $t \in I_{\gamma_1}$,

$$|y(t) - x_0| = |(g(y))(t) - x_0| \leq \int_{t_0}^{t} |f(s, y(s))|ds \leq M\gamma_1 = \rho,$$

so $y|_{I_{\gamma_1}} \in X_{\gamma_1,\rho}$. By continuity, there exists $\gamma_2 \in (\gamma_1, \alpha] \ni y|_{I_{\gamma_2}} \in X_{\gamma_1,r}$. But then $y(t) \equiv x_*(t)$ on I_{γ_2}, contradicting the definition of γ_1. □ □

The Picard Iteration

Although hidden in a few too many details, the main idea of the proof above is to study the convergence of functional iteration of g. If we choose the initial iterate to be $x_0(t) \equiv x_0$, we obtain the classical Picard iteration:

$$\begin{cases} x_0(t) & \equiv & x_0 \\ x_{k+1}(t) & = & x_0 + \int_{t_0}^{t} f(s, x_k(s))ds \quad \text{for} \quad k \geq 0 \end{cases}$$

The argument in the proof of the C.M.F.-P.T. gives only *uniform* estimates of, e.g., $x_{k+1} - x_k$: $\|x_{k+1} - x_k\|_\infty \leq L\alpha \|x_k - x_{k+1}\|_\infty$, leading to the condition $\alpha < \frac{1}{L}$. For the Picard iteration (and other iterations of similar nature, e.g., for Volterra integral equations of the second kind), we can get better results using *pointwise* estimates of $x_{k+1} - x_k$. The condition $\alpha < \frac{1}{L}$ turns out to be unnecessary (we will see another way to eliminate this assumption when we study continuation of solutions). For the moment, we will set aside the uniqueness question and focus on existence.

Theorem (Picard Global Existence for (IE) for Lipschitz f) *Let $I = [t_0, t_0 + \beta]$, and suppose $f(t, x)$ is in (C, Lip) on $I \times \mathbb{F}^n$. Then there exists a solution $x_*(t)$ of the integral equation (IE) in $C(I)$.*

Theorem (Picard Local Existence for (IE) for Lipschitz f) *Let $I = [t_0, t_0 + \beta]$ and $\Omega = \overline{B_r(x_0)} = \{x \in \mathbb{F}^n : |x - x_0| \leq r\}$, and suppose $f(t, x)$ is in (C, Lip) on $I \times \Omega$. Then there exists a solution $x_*(t)$ of the integral equation (IE) in $C(I_\alpha)$ where $I_\alpha = [t_0, t_0 + \alpha]$, $\alpha = \min\left(\beta, \frac{r}{M}\right)$, and where $M = \max_{(t,x) \in I \times \Omega} |f(t, x)|$.*

Proofs. We prove the two theorems together. For the global theorem, let $X = C(I)$ (i.e., $C(I, \mathbb{F}^n)$), and for the local theorem, let

$$X = X_{\alpha,r} \equiv \{x \in C(I_\alpha) : \|x - x_0\|_\infty \leq r\}$$

as before (where $x_0(t) \equiv x_0$). Then the map

$$(g(x))(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

maps X into X in both cases, and X is complete. Let

$$x_0(t) \equiv x_0, \quad \text{and} \quad x_{k+1} = g(x_k) \quad \text{for} \quad k \geq 0.$$

Let

$$\begin{aligned} M_0 &= \max_{t\in I} |f(t,x_0)| \qquad \text{(global thm)},\\ M_0 &= \max_{t\in I_\alpha} |f(t,x_0)| \qquad \text{(local thm)}. \end{aligned}$$

Then for $t \in I$ (global) or $t \in I_\alpha$ (local),

$$\begin{aligned} |x_1(t) - x_0| &\le \int_{t_0}^t |f(s,x_0)|ds \le M_0(t-t_0)\\ |x_2(t) - x_1(t)| &\le \int_{t_0}^t |f(s,x_1(s)) - f(s,x_0(s))|ds\\ &\le L\int_{t_0}^t |x_1(s) - x_0(s)|ds\\ &\le M_0 L \int_{t_0}^t (s-t_0)ds = \frac{M_0L(t-t_0)^2}{2!} \end{aligned}$$

By induction, suppose $|x_k(t) - x_{k-1}(t)| \le M_0 L^{k-1}\frac{(t-t_0)^k}{k!}$. Then

$$\begin{aligned} |x_{k+1}(t) - x_k(t)| &\le \int_{t_0}^t |f(s,x_k(s)) - f(s,x_{k-1}(s))|ds\\ &\le L\int_{t_0}^t |x_k(s) - x_{k-1}(s)|ds\\ &\le M_0L^k \int_{t_0}^t \frac{(s-t_0)^k}{k!}ds = M_0L^k\frac{(t-t_0)^{k+1}}{(k+1)!}. \end{aligned}$$

So

$$\begin{aligned} \sum_{k=0}^\infty |x_{k+1}(t) - x_k(t)| &\le \frac{M_0}{L}\sum_{k=0}^\infty \frac{(L(t-t_0))^{k+1}}{(k+1)!}\\ &= \frac{M_0}{L}(e^{L(t-t_0)} - 1)\\ &\le \frac{M_0}{L}(e^{L\gamma} - 1) \end{aligned}$$

where $\gamma = \beta$ (global) or $\gamma = \alpha$ (local). Hence the series $x_0 + \sum_{k=0}^\infty (x_{k+1}(t) - x_k(t))$, which has x_{N+1} as its N^{th} partial sum, converges absolutely and uniformly on I (global) or I_α (local) by the Weierstrass M-test. Let $x_*(t) \in C(I)$ (global) or $\in C(I_\alpha)$ (local) be the limit function. Since

$$|f(t,x_k(t)) - f(t,x_*(t))| \le L|x_k(t) - x_*(t)|,$$

$f(t,x_k(t))$ converges uniformly to $f(t,x_*(t))$ on I (global) or I_α (local), and thus

$$\begin{aligned} g(x_*)(t) &= x_0 + \int_{t_0}^t f(s,x_*(s))ds\\ &= \lim_{k\to\infty} (x_0 + \int_{t_0}^t f(s,x_k(x))ds)\\ &= \lim_{k\to\infty} x_{k+1}(t) = x_*(t), \end{aligned}$$

for all $t \in I$ (global) or I_α (local). Hence $x_*(t)$ is a fixed point of g in X, and thus also a solution of the integral equation (IE) in $C(I)$ (global) or $C(I_\alpha)$ (local.) □

Corollary. The solution $x_*(t)$ of (IE) satisfies

$$|x_*(t) - x_0| \leq \frac{M_0}{L}(e^{L(t-t_0)} - 1)$$

for $t \in I$ (global) or $t \in I_\alpha$ (local), where $M_0 = \max_{t\in I} |f(t, x_0)|$ (global), $M_0 = \max_{t=I_\alpha} |f(t, x_0)|$ (local).

Proof. This is established in the proof above. □ □

Remark. In each of the statements of the last three Theorems, we could replace "solution of the integral equation (IE)" with "solution of the IVP: $DE : x' = f(t, x)$; $IC : x(t_0) = x_0$" because of the equivalence of these two problems.

Examples.

(1) Consider a *linear* system $x' = A(t)x + b(t)$, where $A(t) \in \mathbb{C}^{n\times n}$ and $b(t) \in \mathbb{C}^n$ are in $C(I)$ (where $I = [t_0, t_0 + \beta]$). Then f is in (C, Lip) on $I \times \mathbb{F}^n$:

$$|f(t,x) - f(t,y)| \leq |A(t)x - A(t)y| \leq \left(\max_{t\in I} \|A(t)\|\right) |x - y|.$$

Hence there is a solution of the IVP: $x' = A(t)x + b(t)$, $x(t_0) = x_0$ in $C^1(I)$.

(2) $(n = 1)$ Consider the IVP: $x' = x^2$, $x(0) = x_0 > 0$. Then $f(t, x) = x^2$ is not in (C, Lip) on $I \times \mathbb{R}$. It is, however, in (C, Lip) on $I \times \Omega$ where $\Omega = \overline{B_r(x_0)} = [x_0 - r, x_0 + r]$ for each fixed r. For a given $r > 0$, $M = (x_0 + r)^2$, and $\alpha = \frac{r}{M} = \frac{r}{(x_0+r)^2}$ in the local theorem is maximized for $r = x_0$, where $\alpha = \frac{1}{4x_0}$. So the local theorem guarantees a solution in $\left[0, \frac{1}{4x_0}\right]$. The actual solution $x_*(t) = (x_0^{-1} - t)^{-1}$ exists in $\left[0, \frac{1}{x_0}\right)$.

Local Existence for Continuous f

A condition similar to the Lipschitz condition is needed to guarantee that the Picard iterates converge; it is also needed for uniqueness, which we will return to shortly. It is, however, still possible to prove a local existence theorem assuming only that f is continuous, without assuming the Lipschitz condition. We will need the following form of Ascoli's Theorem:

Theorem (Ascoli) *Let X and Y be metric spaces with X compact. Let $\{f_k\}$ be an* equicontinuous *sequence of functions $f_k : X \to Y$, i.e.,*

$$(\forall \varepsilon > 0)(\exists \delta > 0) \quad \textit{such that} \quad (\forall k \geq 1)(\forall x_1, x_2 \in X)$$
$$d_X(x_1, x_2) < \delta \Rightarrow d_Y(f_k(x_1), f_k(x_2)) < \varepsilon$$

(in particular, each f_k is continuous), and suppose for each $x \in X$, $\overline{\{f_k(x) : k \geq 1\}}$ is a compact subset of Y. Then there is a subsequence $\{f_{k_j}\}_{j=1}^\infty$ and a continuous $f : X \to Y$ such that

$$f_{k_j} \to f \quad \textit{uniformly on } X.$$

Remark. If $Y = \mathbb{F}^n$, the condition $(\forall\, x \in X)$ $\overline{\{f_k(x) : k \geq 1\}}$ is compact is equivalent to the sequence $\{f_k\}$ being *pointwise bounded*, i.e.,

$$(\forall\, x \in X)(\exists\, M_x) \quad \text{such that} \quad (\forall\, k \geq 1) \quad |f_k(x)| \leq M_x.$$

Example. Suppose $f_k : [a, b] \to \mathbb{R}$ is a sequence of C^1 functions, and suppose there exists $M > 0$ such that

$$(\forall\, k \geq 1) \quad \|f_k\|_\infty + \|f_k'\|_\infty \leq M$$

(where $\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|$). Then for $a \leq x_1 < x_2 \leq b$,

$$|f_k(x_2) - f_k(x_1)| \leq \int_{x_1}^{x_2} |f_k'(x)| dx \leq M|x_2 - x_1|,$$

so $\{f_k\}$ is equicontinuous (take $\delta = \frac{\varepsilon}{M}$), and $\|f_k\|_\infty \leq M$ certainly implies $\{f_k\}$ is pointwise bounded. So by Ascoli's Theorem, some subsequence of $\{f_k\}$ converges uniformly to a continuous function $f : [a, b] \to \mathbb{R}$.

Theorem. The Cauchy-Peano Existence Theorem
Let $I = [t_0, t_0 + \beta]$ and $\Omega = \overline{B_r(x_0)} = \{x \in \mathbb{F}^n : |x - x_0| \leq r\}$, and suppose $f(t, x)$ is continuous on $I \times \Omega$. Then there exists a solution $x_*(t)$ of the integral equation

$$\text{(IE)} \qquad x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) ds$$

in $C(I_\alpha)$ where $I_\alpha = [t_0, t_0 + \alpha]$, $\alpha = \min\left(\beta, \frac{r}{M}\right)$, and $M = \max_{(t,x) \in I \times \Omega} |f(t, x)|$ (and thus $x_*(t)$ is a C^1 solution of the IVP: $x' = f(t, x)$; $x(t_0) = x_0$ in I_α).

Proof. The idea of the proof is to construct continuous approximate solutions explicitly (we will use the piecewise linear interpolants of grid functions generated by Euler's method), and use Ascoli's Theorem to take the uniform limit of some subsequence. For each integer $k \geq 1$, define $x_k(t) \in C(I_\alpha)$ as follows: partition $[t_0, t_0 + \alpha]$ into k equal subintervals (for $0 \leq \ell \leq k$, let $t_\ell = t_0 + \ell \frac{\alpha}{k}$ (note: t_ℓ depends on k too)), set $x_k(t_0) = x_0$, and for $\ell = 1, 2, \ldots, k$ define $x_k(t)$ in $(t_{\ell-1}, t_\ell]$ inductively by $x_k(t) = x_k(t_{\ell-1}) + f(t_{\ell-1}, x_k(t_{\ell-1}))(t - t_{\ell-1})$. For this to be well-defined we must check that $|x_k(t_{\ell-1}) - x_0| \leq r$ for $2 \leq \ell \leq k$ (it is obvious for $\ell = 1$); inductively, we have

$$\begin{aligned}
|x_k(t_{\ell-1}) - x_0| &\leq \sum_{i=1}^{\ell-1} |x_k(t_i) - x_k(t_{i-1})| \\
&= \sum_{i=1}^{\ell-1} |f(t_{i-1}, x_k(t_{i-1}))| \cdot |t_i - t_{i-1}| \\
&\leq M \sum_{i=1}^{\ell-1} (t_i - t_{i-1}) \\
&= M(t_{\ell-1} - t_0) \leq M\alpha \leq r
\end{aligned}$$

by the choice of α. So $x_k(t) \in C(I_\alpha)$ is well defined. A similar estimate shows that for $t, \tau \in [t_0, t_0 + \alpha]$,

$$|x_k(t) - x_k(\tau)| \leq M|t - \tau|.$$

This implies that $\{x_k\}$ is equicontinuous; it also implies that

$$(\forall\, k \geq 1)(\forall\, t \in I_\alpha) \quad |x_k(t) - x_0| \leq M\alpha \leq r,$$

so $\{x_k\}$ is pointwise bounded (in fact, uniformly bounded). So by Ascoli's Theorem, there exists $x_*(t) \in C(I_\alpha)$ and a subsequence $\{x_{k_j}\}_{j=1}^\infty$ converging uniformly to $x_*(t)$. It remains to show that $x_*(t)$ is a solution of (IE) on I_α. Since each $x_k(t)$ is continuous and piecewise linear on I_α,

$$x_k(t) = x_0 + \int_{t_0}^t x_k'(s)ds$$

(where $x_k'(t)$ is piecewise constant on I_α and is defined for all t except t_ℓ $(1 \leq \ell \leq k-1)$, where we define it to be $x_k'(t_\ell^+)$). Define

$$\Delta_k(t) = x_k'(t) - f(t, x_k(t)) \quad \text{on} \quad I_\alpha$$

(note that $\Delta_k(t_\ell) = 0$ for $0 \leq \ell \leq k-1$ by definition). We claim that $\Delta_k(t) \to 0$ uniformly on I_α as $k \to \infty$. Indeed, given k, we have for $1 \leq \ell \leq k$ and $t \in (t_{\ell-1}, t_\ell)$ (including t_k if $\ell = k$), that

$$|x_k'(t) - f(t, x_k(t))| = |f(t_{\ell-1}, x_k(t_{\ell-1})) - f(t, x_k(t))|.$$

Noting that $|t - t_{\ell-1}| \leq \frac{\alpha}{k}$ and

$$|x_k(t) - x_k(t_{\ell-1})| \leq M|t - t_{\ell-1}| \leq M\frac{\alpha}{k},$$

the uniform continuity of f (being continuous on the compact set $I \times \Omega$) implies that

$$\max_{t \in I_\alpha} |\Delta_k(t)| \to 0 \quad \text{as} \quad k \to \infty.$$

Thus, in particular, $\Delta_{k_j}(t) \to 0$ uniformly on I_α. Now

$$\begin{aligned} x_{k_j}(t) &= x_0 + \int_{t_0}^t x_{k_j}'(s)ds \\ &= x_0 + \int_{t_0}^t f(s, x_{k_j}(s))ds + \int_{t_0}^t \Delta_{k_j}(s)ds. \end{aligned}$$

Since $x_{k_j} \to x_*$ uniformly on I_α, the uniform continuity of f on $I \times \Omega$ now implies that $f(t, x_{k_j}(t)) \to f(t, x_*(t))$ uniformly on I_α, so taking the limit as $j \to \infty$ on both sides of this equation for each $t \in I_\alpha$, we obtain that x_* satisfies (IE) on I_α □ □

Remark. In general, the choice of a subsequence of $\{x_k\}$ is necessary: there are examples where the sequence $\{x_k\}$ does not converge. (See Problem 12, Chapter 1 of Coddington & Levinson.)

Uniqueness

Uniqueness theorems are typically proved by comparison theorems for solutions of scalar differential equations, or by inequalities. The most fundamental of these inequalities is Gronwall's inequality, which applies to real first-order linear scalar equations.

Recall that a first-order linear scalar initial value problem

$$IVP: \qquad u' = a(t)u + b(t), \quad u(t_0) = u_0$$

can be solved by multiplying by the integrating factor $e^{-\int_{t_0}^t a}$ (i.e., $e^{-\int_{t_0}^t a(s)ds}$), and then integrating from t_0 to t. That is,

$$\frac{d}{dt}\left(e^{-\int_{t_0}^t a}u(t)\right) = e^{-\int_{t_0}^t a}b(t),$$

implies that

$$\begin{aligned} e^{-\int_{t_0}^t a}u(t) - u_0 &= \int_{t_0}^t \frac{d}{ds}\left(e^{-\int_{t_0}^s a}u(s)\right)ds \\ &= \int_{t_0}^t e^{-\int_{t_0}^s a}b(s)ds \end{aligned}$$

which in turn implies that

$$u(t) = u_0 e^{\int_{t_0}^t a} + \int_{t_0}^t e^{\int_s^t a}b(s)ds.$$

Since $f(t) \leq g(t)$ on $[c, d]$ implies $\int_c^d f(t)dt \leq \int_c^d g(t)dt$, the identical argument with "=" replaced by "≤" gives

Theorem (Gronwall's Inequality - differential form) *Let $I = [t_0, t_1]$. Suppose $a : I \to \mathbb{R}$ and $b : I \to \mathbb{R}$ are continuous, and suppose $u : I \to \mathbb{R}$ is in $C^1(I)$ and satisfies*

$$u'(t) \leq a(t)u(t) + b(t) \quad \textit{for} \quad t \in I, \quad \textit{and} \quad u(t_0) = u_0.$$

Then

$$u(t) \leq u_0 e^{\int_{t_0}^t a} + \int_{t_0}^t e^{\int_s^t a}b(s)ds.$$

Remarks.

(1) Thus a solution of the differential inequality is bounded above by the solution of the equality (i.e., differential equation $u' = au + b$).

(2) The result clearly still holds if u is only continuous and piecewise C^1, and $a(t)$ and $b(t)$ are only piecewise continuous.

(3) There is also an integral form of Gronwall's inequality (i.e., the hypothesis is an integral inequality): if $\varphi, \psi, \alpha \in C(I)$ are real-valued with $\alpha \geq 0$ on I, and

$$\varphi(t) \leq \psi(t) + \int_{t_0}^{t} \alpha(s)\varphi(s)ds \quad \text{for} \quad t \in I,$$

then

$$\varphi(t) \leq \psi(t) + \int_{t_0}^{t} e^{\int_s^t \alpha} \alpha(s)\psi(s)ds.$$

In particular, if $\psi(t) \equiv c$ (a constant), then $\varphi(t) \leq c e^{\int_{t_0}^t \alpha}$. (The differential form is applied to the C^1 function $u(t) = \int_{t_0}^t \alpha(s)\varphi(s)ds$ in the proof — see problem 4 on Prob. Set. 9.)

(4) For $a(t) \geq 0$, the differential form is also a consequence of the integral form: integrating

$$u' \leq a(t)u + b(t) \quad \text{from} \quad t_0 \quad \text{to} \quad t$$

gives

$$u(t) \leq \psi(t) + \int_{t_0}^{t} a(s)u(s)ds,$$

where

$$\psi(t) = u_0 + \int_{t_0}^{t} b(s)ds,$$

so integration by parts gives

$$\begin{aligned} u(t) &\leq \psi(t) + \int_{t_0}^{t} e^{\int_s^t a} a(s)\psi(s)ds \\ &= \cdots = u_0 e^{\int_{t_0}^t a} + \int_{t_0}^{t} e^{\int_s^t a} b(s)ds. \end{aligned}$$

(5) Caution: a differential inequality implies an integral inequality, but *not* vice versa: $f \leq g \not\Rightarrow f' \leq g'$.

(6) The integral form doesn't require $\varphi \in C^1$ (just $\varphi \in C(I)$), but is restricted to $\alpha \geq 0$. The differential form has no sign restriction on $a(t)$, but it requires a stronger hypothesis (in view of (5) and the requirement that u be continuous and piecewise C^1).

Uniqueness for Locally Lipschitz f

We start with a one-sided local uniqueness theorem for the initial value problem

$$IVP: \qquad x' = f(t, x); \quad x(t_0) = x_0.$$

Theorem. *Suppose for some $\alpha > 0$ and some $r > 0$, $f(t,x)$ is in (C, Lip) on $I_\alpha \times \overline{B_r(x_0)}$, and suppose $x(t)$ and $y(t)$ both map I_α into $\overline{B_r(x_0)}$ and both are C^1 solutions of (IVP) on $I_\alpha = [t_0, t_0+\alpha]$. Then $x(t) = y(t)$ for $t \in I_\alpha$.*

Proof. Set

$$u(t) = |x(t) - y(t)|^2 = \langle x(t) - y(t), x(t) - y(t)\rangle$$

(in the Euclidean inner product on $\mathbb{F}^n$). Then $u : I_\alpha \to [0,\infty)$ and $u \in C^1(I_\alpha)$ and for $t \in I_\alpha$,

$$\begin{aligned} u' &= \langle x-y, x'-y'\rangle + \langle x'-y', x-y\rangle \\ &= 2\mathcal{R}e\langle x-y, x'-y'\rangle \le 2|\langle x-y, x'-y'\rangle| \\ &= 2|\langle x-y, (f(t,x) - f(t,y))\rangle| \\ &\le |x-y| \cdot |f(t,x) - f(t,y)| \\ &\le 2L|x-y|^2 = 2Lu \ . \end{aligned}$$

Thus $u' \le Lu$ on I_α and $u(t_0) = x(t_0) - y(t_0) = x_0 - x_0 = 0$. By Gronwall's inequality, $u(t) \le u_0 e^{Lt} = 0$ on I_α, so since $u(t) \ge 0$, $u(t) \equiv 0$ on I_α. □ □

Corollary.

(i) The same result holds if $I_\alpha = [t_0 - \alpha, t_0]$.

(ii) The same result holds if $I_\alpha = [t_0 - \alpha, t_0 + \alpha]$.

Proof. For (i), let $\widetilde{x}(t) = x(2t_0 - t)$, $\widetilde{y}(t) = y(2t_0 - t)$, and $\widetilde{f}(t,x) = -f(2t_0 - t, x)$. Then $\widetilde{f}$ is in (C, Lip) on $[t_0, t_0+\alpha] \times \overline{B_r(x_0)}$, and $\widetilde{x}$ and $\widetilde{y}$ both satisfy the IVP

$$x' = \widetilde{f}(t,x); \quad x'(t_0) = x_0 \quad \text{on} \quad [t_0, t_0+\alpha].$$

So by the Theorem, $\widetilde{x}(t) = \widetilde{y}(t)$ for $t \in [t_0, t_0+\alpha]$, i.e., $x(t) = y(t)$ for $t \in [t_0 - \alpha, t_0]$. Now (ii) follows immediately by applying the Theorem in $[t_0, t_0+\alpha]$ and applying (ii) in $[t_0 - \alpha, t_0]$.
□ □

Remark. The idea used in the proof of (i) is often called "time-reversal." The important part is that $\widetilde{x}(t) = x(c - t)$, for some constant c, so that $\widetilde{x}'(t) = -x'(c-t)$, etc. The choice of $c = 2t_0$ is convenient but not essential.

The main uniqueness theorem is easiest to state in its two-sided version (i.e., where t_0 is in the interior of the interval of definition of a solution of the IVP). One-sided versions (where t_0 is the left endpoint or right endpoint of the interval of definition of a solution of the IVP) are true and have the same proof, but require a more delicate statement. (Exercise: State one-sided theorems corresponding to the upcoming theorem precisely.)

Definition. Let $\mathcal{D}$ be an open set in $\mathbb{R} \times \mathbb{F}^n$. We say that $f(t,x)$ mapping $\mathcal{D}$ into $\mathbb{F}^n$ is *locally Lipschitz continuous with respect to x* if for each $(t_1, x_1) \in \mathcal{D}$ there exists

$$\alpha > 0, \quad r > 0, \quad \text{and} \quad L > 0$$

for which $[t_1 - \alpha, t_1 + \alpha] \times \overline{B_r(x_1)} \subset \mathcal{D}$ and

$$(\forall\, t \in [t_1 - \alpha, t_1 + \alpha])(\forall\, x, y \in \overline{B_r(x_1)}) \quad |f(t,x) - f(t,y)| \le L|x - y|$$

(i.e., f is uniformly Lipschitz continuous with respect to x in $[t_1 - \alpha, t_1 + \alpha] \times \overline{B_r(x_1)}$). We will say $f \in (C, \mathrm{Lip}_{\mathrm{loc}})$ (not a standard notation) on $\mathcal{D}$ if f is continuous on $\mathcal{D}$ and locally Lipschitz continuous with respect to x on $\mathcal{D}$.

Example. Let $\mathcal{D}$ be an open set of $\mathbb{R} \times \mathbb{F}^n$. Suppose $f(t,x)$ maps $\mathcal{D}$ into $\mathbb{F}^n$, f is continuous on $\mathcal{D}$, and for $1 \le i, j \le n$, $\frac{\partial f_i}{\partial x_j}$ exists and is continuous in $\mathcal{D}$. (Briefly, we say f is continuous on $\mathcal{D}$ and C^1 with respect to x on $\mathcal{D}$.) Then $f \in (C, \mathrm{Lip}_{\mathrm{loc}})$ on $\mathcal{D}$. (Exercise.)

Main Uniqueness Theorem. *Let $\mathcal{D}$ be an open set in $\mathbb{R} \times \mathbb{F}^n$, and suppose $f \in (C, \mathrm{Lip}_{\mathrm{loc}})$ on $\mathcal{D}$. Suppose $(t_0, x_0) \in \mathcal{D}$, $I \subset \mathbb{R}$ is some interval containing t_0 (which may be open or closed at either end), and suppose $x(t)$ and $y(t)$ are both solutions of the initial value problem*

$$IVP: \qquad x' = f(t,x): \quad x(t_0) = y_0$$

in $C^1(I)$ which satisfy $(t, x(t)) \in \mathcal{D}$ and $(t, y(t)) \in \mathcal{D}$ for $t \in I$. Then $x(t) \equiv y(t)$ on I.

Proof. We first show $x(t) \equiv y(t)$ on $\{t \in I : t \ge t_0\}$. If not, let

$$t_1 = \inf\{t \in I : t \ge t_0 \quad \text{and} \quad x(t) \ne y(t)\}.$$

Then $x(t) = y(t)$ on $[t_0, t_1)$ so by continuity $x(t_1) = y(t_1)$ (if $t_1 = t_0$, this is obvious). By continuity and the openness of $\mathcal{D}$ (as $(t_1, x(t_1)) \in \mathcal{D}$), there exist $\alpha > 0$ and $r > 0$ such that $[t_1 - \alpha, t_1 + \alpha] \times \overline{B_r(x_1)} \subset \mathcal{D}$, f is uniformly Lipschitz continuous with respect to x in $[t_1 - \alpha, t_1 + \alpha] \times \overline{B_r(x_1)}$, and $x(t) \in \overline{B_r(x_1)}$ and $y(t) \in \overline{B_r(x_1)}$ for all t in $I \cap [t_1 - \alpha, t_1 + \alpha]$. By the previous theorem, $x(t) \equiv y(t)$ in $I \cap [t_1 - \alpha, t_1 + \alpha]$, contradicting the definition of t_1. Hence $x(t) \equiv y(t)$ on $\{t \in I : t \ge t_0\}$. Similarly, $x(t) \equiv y(t)$ on $\{t \in I : t \le t_0\}$. Hence $x(t) \equiv y(t)$ on I. □ □

Remark. t_0 is allowed to be the left or right endpoint of I.

Comparison Theorem for Nonlinear Real Scalar Equations

Theorem. *Let $n = 1$, $\mathbb{F} = \mathbb{R}$. Suppose $f(t,u)$ is continuous in t and u and Lipschitz continuous in u. Suppose $u(t)$, $v(t)$ are C^1 for $t \ge t_0$ (or some interval $[t_0, b)$ or $[t_0, b]$) and satisfy*

$$u'(t) \le f(t, u(t)), \qquad\qquad v(t) = f(t, v(t))$$

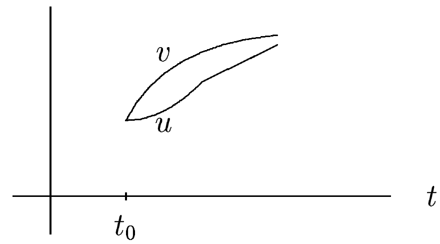
and $u(t_0) \le v(t_0)$. Then $u(t) \le v(t)$ for $t \ge t_0$.

Proof. By contradiction. If $u(T) > v(T)$ for some $T > t_0$, then set $t_1 = \sup\{t : t_0 \le t < T$ and $u(t) \le v(t)\}$. Then $t_0 \le t_1 < T$, $u(t_1) = v(t_1)$, and $u(t) > v(t)$ for $t > t_1$ (by continuity of $u - v$). For $t_1 \le t \le T$, $|u(t) - v(t)| = u(t) - v(t)$, so we have

$$(u - v)' \le f(t,u) - f(t,v) \le L|u - v| = L(u - v).$$

By Gronwall's inequality (applied to $u - v$ on $[t_1, T]$, with $(u-v)(t_1) = 0$, $a(t) \equiv L$, $b(t) \equiv 0$), $(u - v)(t) \le 0$ on $[t_1, T]$, a contradiction. □

Remarks. (1) as with the differential form of Gronwall's inequality a solution of the differential inequality $u' \leq f(t, u)$ is bounded above by the solution of the equality (i.e., the DE $v' = f(t, v)$). (2) It can be shown under the same hypotheses that if $u(t_0) < v(t_0)$, then $u(t) < v(t)$ for $t \geq t_0$. (3) Caution: It may happen that $u'(t) > v'(t)$ for some $t \geq t_0$: $u(t) \leq v(t) \not\Rightarrow u'(t) \leq v'(t)$.



Corollary. Let $n = 1$, $\mathbb{F} = \mathbb{R}$. Suppose $f(t, u) \leq g(t, u)$ are continuous in t and u, and one of them is Lipschitz continuous in u. Suppose also that $u(t)$, $v(t)$ are C^1 for $t \geq t_0$ (or some interval $[t_0, b)$ or $[t_0, b]$) and satisfy $u' = f(t, u)$, $v' = g(t, v)$, and $u(t_0) \leq v(t_0)$. Then $u(t) \leq v(t)$ for $t \geq t_0$.

Proof. Suppose first that g satisfies the Lipschitz condition. Then $u' = f(t, u) \leq g(t, u)$. Now apply the theorem. If f satisfies the Lipschitz condition, apply the first part of this proof to $\widetilde{u}(t) \equiv -v(t)$, $\widetilde{v}(t) \equiv -u(t)$, $\widetilde{f}(t, u) = -g(t, -u)$, $\widetilde{g}(t, u) = -f(t, -u)$. $\square$

Remark. Again, if $u(t_0) < v(t_0)$, then $u(t) < v(t)$ for $t \geq t_0$.

Continuation of Solutions

We consider two kinds of results

- local continuation (continuation at a point — no Lipschitz condition assumed)
- global continuation (for locally Lipschitz f)

Continuation at a Point

Suppose $x(t)$ is a solution of the DE $x' = f(t, x)$ on an interval I and that f is continuous on some subset $\mathcal{S} \subset \mathbb{R} \times \mathbb{F}^n$ containing $\{(t, x(t)) : t \in I\}$. (Note: no Lipschitz condition is assumed.)

Case 1. I is closed at the right end, i.e., $I = (-\infty, b]$, $[a, b]$, or $(a, b]$. $\longrightarrow$ Assume further that $(b, x(b))$ is in the interior of $\mathcal{S}$. Then the solution can be extended (by the Cauchy-Peano Existence Theorem) to an interval with right end $b + \beta$ for some $\beta > 0$. (Solve the IVP $x' = f(t, x)$ with initial value $x(b)$ at $t = b$ on some interval $[b, b + \beta]$ by Cauchy-Peano. To show that the connection is C^1 at $t = b$, note that the extended $x(t)$ satisfies the integral equation $x(t) = x(b) + \int_b^t f(s, x(s))ds$ on the extended interval $I \bigcup [b, b + \beta]$.)

Case 2. I is open at the right end, i.e., $I = (-\infty, b)$, $[a, b)$, or (a, b) with $b < \infty$. $\longrightarrow$ Assume further that $f(t, x(t))$ is *bounded* on $[t_0, b)$ for some $t_0 < b$ with $[t_0, b) \subset I$, say $|f(t, x(t))| \leq M$ on $[t_0, b)$.

[Remarks. (1) If this is true for any $\widetilde{t_0} \in I$, it is true for all $t_0 \in I$ (where of course M depends on t_0): for $t_0 < \widetilde{t_0}$, $f(t, x(t))$ is cont. on $[t_0, \widetilde{t_0}]$. So this assumption is a condition on the behavior of $f(t,x(t))$ near $t=b$. (2) This assumption can be restated with a slightly different emphasis: for some $t_0 \in I$, $\{(t,x(t)) : t_0 \le t < b\}$ stays within a subset of $\mathcal{S}$ on which f is bounded. For example, if $\{(t,x(t)) : t_0 \le t < b\}$ stays within a compact subset of $\mathcal{S}$, this condition is satisfied.] Then the integral equation $(*)$ $x(t) = x(t_0) + \int_{t_0}^{t} f(s,x(s))ds$ holds for $t \in I$. In particular, for $t_0 \le \tau \le t < b$,

$$|x(t) - x(\tau)| = \left| \int_{\tau}^{t} f(s,x(s))ds \right| \le \int_{\tau}^{t} |f(s,x(s))|ds \le M|t-\tau|.$$

Thus, for any sequence $t_n \uparrow b$, $\{x(t_n)\}$ is Cauchy. This implies $\lim_{t\to b^-} x(t)$ exists; call it $x(b^-)$. So $x(t)$ has a *continuous* extension from I to $I \cup \{b\}$. If in addition $(b, x(b^-))$ is in $\mathcal{S}$, then $(*)$ holds on $I \cup \{b\}$ as well, so $x(t)$ is a C^1 *solution* of $x' = f(t,x)$ on $I \cup \{b\}$. (Of course, if now in addition $(b, x(b^-))$ is in the interior of $\mathcal{S}$, we are back in Case 1 and can extend the solution $x(t)$ a little beyond $t = b$.)

Case 3. I is closed at the left end — similar to Case 1.

Case 4. I is open at the left end — similar to Case 2.

Global Continuation

Now suppose $f(t,x)$ is continuous on an open set $\mathcal{D} \subset \mathbb{R} \times \mathbb{F}^n$ and suppose f is locally Lipschitz continuous with respect to x on $\mathcal{D}$. (For example, if f is C^1 with respect to x in $\mathcal{D}$, i.e., $\frac{\partial f_i}{\partial x_j}$ exists and is continuous in $\mathcal{D}$ for $1 \le i,j \le n$, then f is locally Lipschitz continuous with respect to x on $\mathcal{D}$.) For brevity, we will write $f \in (C, \mathrm{Lip}_{\mathrm{loc}})$ on $\mathcal{D}$. Let $(t_0, x_0) \in \mathcal{D}$. We want to continue in t solutions of the IVP $x' = f(t,x)$, $x(t_0) = x_0$. Part of being a solution is that $(t, x(t)) \in \mathcal{D}$ (we are only assuming f is defined in $\mathcal{D}$). We know local existence of solutions and uniqueness of solutions on any interval.

Define $T_+ = \sup\{t > t_0 : \exists$ a solution of the IVP on $[t_0, t)\}$. By uniqueness, two solutions must agree on their common interval of definition, so $\exists$ a solution on $[t_0, T_+)$. Define T_- similarly. So (T_-, T_+) is the maximal interval of existence of the solution of the IVP. It is possible that $T_+ = \infty$ and/or $T_- = -\infty$. Note that the maximal interval (T_-, T_+) is open: if the solution could be extended to T_+ (or T_-), then since $\mathcal{D}$ is open, the results above on continuation at a point imply that the solution could be extended beyond T_+ (or T_-), contradicting the definition of T_+ (or T_-).

The ideal situation would be $T_+ = +\infty$ and $T_- = -\infty$, where the solution exists for all time t. Another "good" situation is if $f(t,x)$ is not defined for $t \ge T_+$. For example, if $a(t) = \frac{1}{1-t}$ (which blows up at $t = 1$), and $x'(t) = a(t)$, we don't expect the solution to exist beyond $t = 1$. Here, if $t_0 = 0$ and $\mathcal{D} = [0,1) \times \mathbb{R}$, then $T_+ = 1$.

Other less desirable behavior occurs for $x' = x^2$, $x(0) = x_0 > 0$, $t_0 = 0$, and $\mathcal{D} = \mathbb{R} \times \mathbb{R}$. The solution $x(t) = \frac{x_0}{1-x_0 t} = \frac{1}{\frac{1}{x_0} - t}$ blows up at $T_+ = \frac{1}{x_0}$ (note that $T_- = -\infty$). Observe that $x(t) \to \infty$ as $t \to (T_+)^-$. So the solution does not just "stop" in the interior of $\mathcal{D}$. This is the general behavior in this situation.

Theorem. Suppose $f \in (C, \mathrm{Lip}_{\mathrm{loc}})$ on an open set $\mathcal{D} \subset \mathbb{R} \times \mathbb{F}^n$. Let $(t_0, x_0) \in \mathcal{D}$, and let (T_-, T_+) be the maximal interval of existence of the solution of the IVP $x' = f(t,x)$,

$x(t_0) = x_0$. Given a compact set $K \subset \mathcal{D}$, there exists a $T < T_+$ for which $(t, x(t)) \notin K$ for $t > T$.

Proof. If not, $\exists\, t_j \to T_+$ with $(t_j, x(t_j)) \in K$ for all j. By taking a subsequence, we may assume that $x(t_j)$ also converges, say to $x_+ \in \mathbb{F}^n$, and $(t_j, x(t_j)) \to (T_+, x_+) \in K \subset \mathcal{D}$. We can thus choose r, τ, $N \ni \bigcup_{j=N}^{\infty} \{(t,x) : |t - t_j| \le \tau, |x - x(t_j)| \le r\}$ is contained in a compact subset of $\mathcal{D}$. There is an M for which $|f(t,x)| \le M$ on this compact set. By the local existence theorem, the solution of $x' = f(t,x)$ starting at the initial point $(t_j, x(t_j))$ exists for a time interval of length $T' \equiv \min\left\{\tau, \frac{r}{M}\right\}$, *independent of* j. Choose j for which $t_j > t_+ - T'$. Then $(t, x(t))$ exists in $\mathcal{D}$ beyond time T_+, which is a contradiction. □

Autonomous Systems

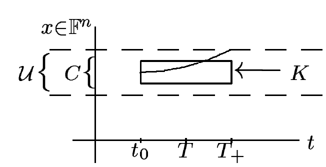
The system of ODE's $x'(t) = f(t,x)$ is called an *autonomous system* if $f(t,x)$ is independent of t, i.e., the ODE is of the form $x' = f(x)$.

Remarks.

(1) Time translates of solutions of an autonomous system are again solutions: if $x(t)$ is a solution, so is $x(t-c)$ for constant c.

(2) Any system of ODE's $x' = f(t,x)$ is equivalent to an autonomous system. Define "$x_{n+1} = t$" as follows: let $\widetilde{x} = (x_{n+1}, x) \in \mathbb{F}^{n+1}$, $\widetilde{f}(\widetilde{x}) = \widetilde{f}(x_{n+1}, x) = \begin{bmatrix} 1 \\ f(x_{n+1}, x) \end{bmatrix} \in \mathbb{F}^{n+1}$, and consider the autonomous IVP $\widetilde{x}' = \widetilde{f}(\widetilde{x})$, $\widetilde{x}(t_0) = \begin{bmatrix} t_0 \\ x_0 \end{bmatrix}$. This IVP is equivalent to the IVP $x' = f(t,x)$, $x(t_0) = x_0$.

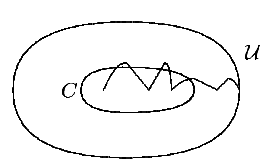
Special case for continuation for autonomous systems $x' = f(x)$



Suppose $f(x)$ is defined and locally Lipschitz continuous on an open set $\mathcal{U} \subset \mathbb{F}^n$. Take $\mathcal{D} = \mathbb{R} \times \mathcal{U}$. Suppose $T_+ < \infty$ and C is a compact subset of $\mathcal{U}$. Take $K = [t_0, T_+] \times C$ in the previous theorem.

It follows that $\exists\, T < T_+$ for which $x(t) \notin C$ for $T < t < T_+$.

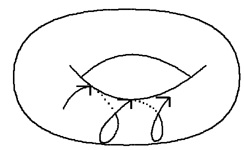
Picture in $\mathbb{F}^n$: We may say that $x(t) \to \partial\mathcal{U} \cup \{\infty\}$ as $t \to (T_+)^-$, meaning that $(\forall\, C^{\text{compact}} \subset \mathcal{U})(\exists\, T < T_+) \ni$ for $t \in (T, T_+)$, $x(t) \notin C$. Stated briefly, eventually $x(t)$ stays out of any given compact set.



Contrapositive

If $x(t)$ stays in a compact set $C \subset \mathcal{U}$, then $T_+ = \infty$.

Example. We may interpret $f(x)$ as a vector field. Suppose there is a smooth compact hypersurface ($n-1$ dim. manifold) $\mathcal{S} \subset \mathcal{U}$ for which $f(x)$ is tangent to $\mathcal{S}$ for all $x \in \mathcal{S}$. Then if $x_0 \in \mathcal{S}$, the solution of the IVP $x' = f(x)$, $x(t_0) = x_0$ stays on $\mathcal{S}$, and $T_+ = \infty$. (Details: exercise: locally patch together.) Generalization: compact manifolds.



Application of continuation theorem to linear systems

Consider the linear system $x'(t) = A(t)x(t) + b(t)$ for $a < t < b$ where $A(t) \in \mathbb{F}^{n\times n}$ and $b(t) \in \mathbb{F}^n$ are continuous on (a,b), with initial value $x(t_0) = x_0$ (where $t_0 \in (a,b)$). Let $\mathcal{D} = (a,b) \times \mathbb{F}^n$. Then $f(t,x) = A(t)x + b(t) \in (C, \mathrm{Lip}_{\mathrm{loc}})$ on $\mathcal{D}$. Moreover, for c,d satisfying $a < c \le t_0 \le d < b$, f is uniformly Lipschitz cont. with respect to x on $[c,d] \times \mathbb{F}^n$ (we can take $L = \max_{c\le t\le d} |A(t)|$). The Picard global existence theorem implies there is a solution of the IVP on $[c,d]$, which is unique by the uniqueness theorem for locally Lipschitz f. This implies that $T_- = a$ and $T_+ = b$. We now give an alternate proof using the continuation theorem.

Idea: prove an *a priori estimate* on the solution to show that $x(t)$ stays in a compact set in $\mathbb{F}^n$ for each compact subinterval of (a,b).

Given c,d satisfying $a < c \le t_0 \le d < b$, let $M = \max_{c\le t\le d}(2|A(t)| + |b(t)|)$. Let $u(t) = |x(t)|^2 = \langle x(t), x(t)\rangle$. Then by Schwarz,

$$\begin{aligned} u'(t) &= \langle x, x'\rangle + \langle x', x\rangle = 2\mathcal{R}e\langle x, x'\rangle \le 2|\langle x, x'\rangle| \le 2|x|\cdot|x'| \\ &= 2|x|\cdot|A(t)x + b(t)| \le 2|A(t)|\cdot|x|^2 + 2|b(t)|\cdot|x| \\ &\le 2|A(t)|\cdot|x|^2 + |b(t)|(|x|^2+1) \le M(|x|^2+1) = M(u+1) \end{aligned}$$

(since $2|x| \le |x|^2 + 1$).

By Gronwall's inequality (applied to $u' \le Mu + M$ with $a(t) \equiv M$, $b(t) \equiv M$), $u(t) \le u_0 e^{M(t-t_0)} + \int_{t_0}^t M e^{M(t-s)} ds = u_0 e^{M(t-t_0)} + e^{M(t-t_0)} - 1 \le R$ for $t_0 \le t \le d$, where $u_0 = u(t_0)$ and $R = (u_0+1)e^{M(b-t_0)} - 1$. So $x(t)$ must lie in the compact set $\{x : |x|^2 \le R\}$ for $t_0 \le t \le d$. So if $T_+ < b$, we would get a contradiction with the continuation theorem, taking $d = T_+$ and $K = [t_0, T_+] \times \{x : |x|^2 \le R\}$. (Strictly speaking, the logic is: for any $t \in [t_0, d]$ for which $x(t)$ exists, we must have $|x(t)|^2 \le R$.) A similar argument shows that $T_- = a$.

Continuity and Differentiability of Solutions

We now study the dependence of solutions of IVPs on (a) initial values, and (b) parameters in the DE. We begin with a fundamental estimate.

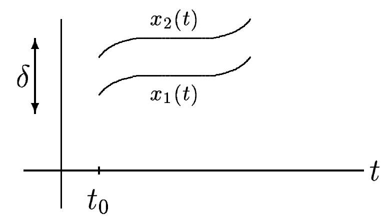
Definition. We say that $x(t)$ is an ε-*approximate solution* of the DE $x' = f(t,x)$ on an interval I if

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

Here, we 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-approx. soln. of $x' = f(t,x)$ and $x_2(t)$ is an ε_2-approx. soln. 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



$$|x_1(t) - x_2(t)| \le \delta e^{L|t-t_0|} + \frac{\varepsilon_1 + \varepsilon_2}{L}(e^{L|t-t_0|} - 1) \quad \forall\, t \in I.$$

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 approx. solutions. Note that this term is 0 at $t = t_0$.

(3) If $\varepsilon_1 = \varepsilon_2 = \delta = 0$, we can recover the uniqueness theorem for Lipschitz f.

Proof. We may assume $\varepsilon_1, \varepsilon_2, \delta > 0$ (otherwise, take limits as $\varepsilon_1 \to 0^+$, $\varepsilon_2 \to 0^+$, $\delta \to 0^+$). Also for simplicity, we may assume $t_0 = 0$ and we are considering $t \ge 0$ (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

$$\begin{aligned} u' &= 2\mathcal{R}e\langle x_1 - x_2, x_1' - x_2' \rangle \le 2|x_1 - x_2| \cdot |x_1' - x_2'| \\ &= 2|x_1 - x_2|\,|x_1' - f(t,x_1) - (x_2' - f(t,x_2)) + f(t,x_1) - f(t,x_2)| \\ &\le 2|x_1 - x_2|(\varepsilon_1 + \varepsilon_2 + L|x_1 - x_2|) = 2Lu + 2\varepsilon\sqrt{u}, \end{aligned}$$

where $\varepsilon = \varepsilon_1 + \varepsilon_2$.

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

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

But

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

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

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

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

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

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

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

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 $f_j \to f$ uniformly can be reduced: exercise.

Proof. Given $\varepsilon > 0$

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

will be $\leq \epsilon$ for all j sufficiently large, uniformly in $t \in [a, b]$. So $x(t)$ is an exact solution and $x_j(t)$ is an ε-approx. solution of $x' = f(t, x)$ on $[a, b]$. By the Fundamental Estimate,

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

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 $\|x\|_{C^1} = \|x\|_\infty + \|x'\|_\infty$).

Parameters in the DE

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, $f(t, x, \mu)$ 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: call it $x(t, \mu, y)$.

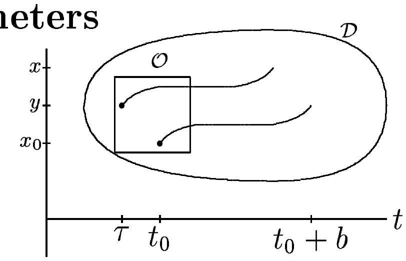
Theorem. If f is continuous in t, x, μ and Lipschitz in x with Lipschitz constant independent of t and μ, then $x(t, \mu, y)$ is continuous in (t, μ, y) jointly.

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 the map $(\mu, y) \mapsto x(t, \mu, y)$ mapping 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 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 $z \in \mathbb{F}^m$ which we think of as the solution to the IVP $z' = 0$, $z(t_0) = \mu$, so that $z(t) \equiv \mu$. Define

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

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

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

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 $\frac{\partial x}{\partial s}(t,s)$:

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

where

$$\left(\frac{\partial x}{\partial s}\right)' = \frac{d}{dt}\left(\frac{\partial x}{\partial s}\right),$$

$D_x f$ is the $n \times n$ Jacobian matrix having (i,j)-element $\frac{\partial f_i}{\partial x_j}$, and $D_s f$ is the $n \times 1$ derivative having ith element $\frac{\partial f_i}{\partial s}$. Evaluating at a fixed s_0, we get that $\frac{\partial x}{\partial s}(t,s_0)$ satisfies

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

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. Now, 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 we can now say is $\left.\frac{\partial x}{\partial s}\right|_{s_0}$. It then 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 of order $\mathcal{O}(|s-s_0|)^2$. 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 $|\Delta x| = \mathcal{O}(|\Delta s|)$ where $\Delta s = s - s_0$. Substitute into the DE, and formally drop terms of order $|\Delta s|^2$.

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

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

Initial Conditions for Equation of Variation

(1) Variable parameters, fixed IC:

$$x' = f(t,x,\mu), \quad x(t_0) = x_0 : \quad \text{take} \quad \frac{\partial x}{\partial \mu_k}(t_0) = 0.$$

(2) Variable IC (with t_0 fixed):

$$x' = f(t,x), \quad x(t_0) = y : \quad \text{take} \quad \frac{\partial x}{\partial y_k}(t_0) = e_k.$$

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 $B = \{y \in \mathbb{F}^n : |y - x_0| \leq r\}$). 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

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

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 previous 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 $y \in B^o$, assume h is small enough that $B_h(y) \subset B^o$], 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 $\varepsilon_1 = \varepsilon_2 = 0$),

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

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

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

Let

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

the modulus of continuity of $D_x f$ (with respect to x) on K. 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

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

since $|x(t,y+he_j) - x(t,,y)| \le |h|e^{L|b-a|}$; where

$$\varepsilon(h) = \omega(|h|e^{L|b-a|})e^{L|b-a|} \to 0 \quad \text{as} \quad h \to 0.$$

We have shown that $\theta(t,y,h)$ is an $\varepsilon(h)$-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,

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

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 $k = 1$ is 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 $x \in C^k$. □

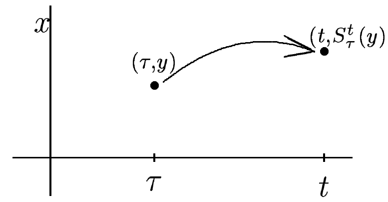
Nonlinear Solution Operator

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 $t \in [a,b]$, for some $\tau \in [a,b]$, 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 function S_τ^t mapping an open set $\mathcal{U} \subset \mathbb{F}^n$ into $\mathbb{F}^n$ by

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

so S_τ^t maps the initial value y (at time τ) into the solution $x(t,\tau,y)$ at time t.

Due to 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 $D_yS_\tau^t$ 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 $D_yS_\tau^t$ 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* operator.

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