II. EXISTENCE, UNIQUENESS, AND ALL THAT

In this chapter we will study the

INITIAL VALUE PROBLEM (IVP): Given an open subset D of $\mathbb{R}\times\mathbb{R}^n$, a continuous function $f : D \to \mathbb{R}^n$, and a point $(t_0, x^0) \in D$, find a solution $x(t)$ to the equations

$$x'(t) = f(t, x) \tag{2.1a}$$
$$x(t_0) = x^0. \tag{2.1b}$$

To completely specify the problem, we formulate more precisely what we mean by a solution:

Definition 2.1: Suppose that $I \subset \mathbb{R}$ is an open interval with $t_0 \in I$. A function $x : I \to \mathbb{R}^n$ is a solution of the IVP (2.1) if

(i) $x(t_0) = x^0$;

(ii) $(t, x(t)) \in D$ for all $t \in I$;

(iii) x is differentiable and $x'(t) = f(t, x(t))$ for all $t \in I$.

We will ask such questions as :

- Do any solutions of the IVP exist?
- If a solution exists, is it unique?
- How does the solution depend on t_0, x_0, and any parameters which may be present in f? Continuously? Differentiably?
- Under what circumstances can a solution defined on an interval I be extended to a larger interval of definition?

The precise answers to these questions will depend on what further hypotheses we put on the IVP, in particular, on the function f.

Let us first emphasize that we always assume that f is continuous, although it is possible to develop some of the theory under weaker hypotheses than this. We will see in section 2.2 that existence of solutions may be proved with no additional hypotheses beyond this continuity. We will frequently assume that f also satisfies a *Lipschitz condition*, a relatively mild restriction which yields uniqueness of solutions and a simpler proof of existence. In general, we will see that continuity of the solution in initial conditions and parameters is a consequence of existence and uniqueness, while to prove differentiability in these variables we will (not unnaturally) have to make differentiability assumptions on f.

2.1 Existence and Uniqueness under a Lipschitz condition

We will take, as our fundamental measure of the length of a vector $z \in \mathbb{R}^m$, the ℓ^2 norm $|z| \equiv |z|_2 = \left(\sum_{i=1}^m |z_i|^2\right)^{1/2}$, although other norms such as $|z|_1 = \sum_{i=1}^m |z_i|$ or $|z|_\infty = \max_{1\le i\le m} |z_i|$ would do as well.

Definition 2.2: (a) Suppose that $S \subset \mathbb{R}^n$. A function $g : S \to \mathbb{R}^m$ *satisfies a Lipschitz condition* or is *Lipschitzian* in S if there is a constant $k \ge 0$ such that, for any $x^1, x^2 \in S$,

$$|g(x^1) - g(x^2)| \le k|x^1 - x^2|.$$

k is called a *Lipschitz constant.*

(b) Suppose that $T \subset \mathbb{R}\times\mathbb{R}^n$. A function $f : T \to \mathbb{R}^n$ which satisfies a Lipschitz condition in x for each fixed t, with a single Lipschitz constant k independent of t, is said to satisfy a Lipschitz condition *with respect to* x in T.

(c) Suppose that $D \subset \mathbb{R} \times \mathbb{R}^n$ is open. A function $f : D \to \mathbb{R}^n$ which satisfies a Lipschitz condition in x on each compact subset $K \subset D$ is said to satisfy a *local* Lipschitz condition, or be *locally Lipschitzian*, with respect to x in D. (In this case the Lipschitz constant k may depend on the compact set K.)

Remark 2.1: (a) A function satisfying a Lipschitz condition in x is necessarily continuous in x, but may not be jointly continuous in x and t. It is easy to verify, however, that continuity in t together with a Lipschitz condition in x guarantees joint continuity in t and x.

(b) A continuous function is not necessarily Lipschitzian or even locally Lipschitzian. To see this, consider $f(x) = x^{1/3}$ for $x \in \mathbb{R}$. Since

$$|f(x) - f(0)| = |f(x)| = |x|^{1/3} = |x|^{-2/3}|x - 0|$$

and $|x|^{-2/3}$ is arbitrarily large for sufficiently small x, f cannot satisfy a Lipschitz condition on any interval containing 0.

(c) We will call a set $R \subset \mathbb{R} \times \mathbb{R}^n$ of the form

$$R = \{(t, x) \mid |t - t_0| \le a, |x - x^0| \le b\}$$

a *rectangle.* It is easy to see that for a continuous function $f : D \to \mathbb{R}^n$ to be locally Lipschitzian with respect to x it suffices that f satisfy a Lipschitz condition with respect to x on every rectangle $R \subset D$, or even that each point of D be contained in some rectangle on which f is Lipschitzian with respect to x.

(d) If $g(x)$ has continuous first derivatives then it is locally Lipschitzian; see Remark 2.7(c).

Our goal in this section is to prove

Theorem 2.1: *Suppose that D and f are as in the statement of IVP, and that f is locally Lipschitzian in x on D. Then for any $(t_0, x^0) \in D$:*

(a) There exists a solution $x : I \to \mathbb{R}^n$ of (2.1) defined on some open interval I containing t_0;

(b) If x and $\tilde{x}$ are two solutions of (2.1) defined on intervals I and $\tilde{I}$, respectively, then $x(t) = \tilde{x}(t)$ for $t \in I \cap \tilde{I}$.

Remark 2.2: Although existence can be proved with no hypotheses on f beyond continuity, some assumption such as a Lipschitz condition is necessary for uniqueness. For example, the IVP $x' = x^{1/3}$, $x(0) = 0$ (with $x \in \mathbb{R}$) has both $x(t) \equiv 0$ and

$$x(t) = \begin{cases} 0, & \text{if } t \le 0, \\ (2t/3)^{3/2}, & \text{if } t > 0, \end{cases}$$

as solutions.

We begin the proof by reformulating IVP as an integral equation.

Lemma 2.2: *Suppose that I is an open interval of $\mathbb{R}$ containing t_0 and that $x : I \to \mathbb{R}^n$ is continuous and satisfies $(t, x(t)) \in D$ for all $t \in I$. Then x is a solution of IVP if and only if, for all $t \in I$,*

$$x(t) = x^0 + \int_{t_0}^{t} f(s, x(s))ds. \tag{2.2}$$

Proof: If x satisfies (2.2) then $x(t_0) = x^0$ and by the Fundamental Theorem of Calculus (first form), x is differentiable and $x'(t) = f(t, x(t))$ for all $t \in I$. Conversely, if x is a solution of IVP, then $x(t)$ is a primitive (antiderivative) for $f(t, x(t))$, so that again by the Fundamental Theorem (second form),

$$\int_{t_0}^{t} f(s, x(s))ds = x(t) - x(t_0) = x(t) - x^0,$$

which is (2.2). ■

Our next result is a fundamental existence theorem due to Picard. It is formulated in terms of a function f defined on a closed rectangle R rather than on an open set D.

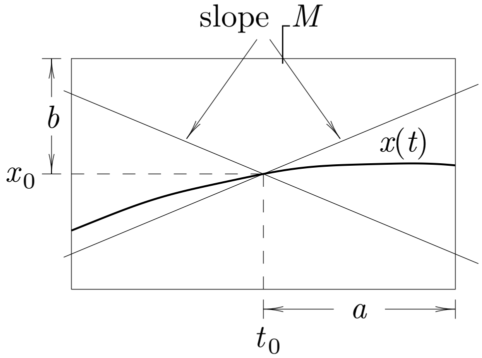
Theorem 2.3: *Let $R \subset \mathbb{R} \times \mathbb{R}^n$ be the closed rectangle*

$$R = \{(t, x) \mid |t - t_0| \le a, |x - x^0| \le b\}, \tag{2.3}$$

where $a, b > 0$, and suppose that $f : R \to \mathbb{R}^n$ is continuous and satisfies a Lipschitz condition in x with Lipschitz constant k. Let M be the maximum of $|f|$ on R and let $A = \min\{a, b/M\}$. Then there exists a function $x(t)$ defined on the (closed) interval $J = [t_0 - A, t_0 + A]$ and satisfying the integral equation (2.2) for all $t \in J$.

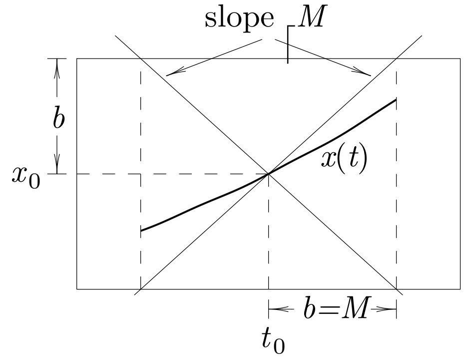
Remark 2.3: The restriction on the half-width A of the interval of definition J—that $A \le a$ and $A \le b/M$—follows from the requirement that $f(t, x(t))$ be defined, i.e., that $(t, x(t))$ lie in R. The first condition simply guarantees that $|t - t_0| \le a$, i.e., that our solution cannot leave the rectangle through the "end" where $|t - t_0| = a$ (see Figure 2.1(a) for the case $n = 1$). Moreover, since $|x'(t)| = |f(t, x(t))| \le M$, $x(t)$ must always satisfy $|x(t) - x^0| \le M|t - t_0|$, so that the second condition $|t - t_0| \le A \le b/M$ guarantees

$|x(t)-x^0| \le b$, i.e., that the solution does not leave the rectangle through the "side" where $|x - x^0| = b$ (see Figure 2.1(b)).



Case(a): $A = a < b/M$

Figure 2.1(a)



Case (b): $A = b/M < a$

Figure 2.1(b)

Proof: The idea is to define a sequence of functions $\{x_m(t)\}_{m=0}^{\infty}$, with $x_m : J \to \mathbb{R}^n$, by

$$\begin{aligned} x_0(t) &= x^0, \\ x_m(t) &= x^0 + \int_{t_0}^{t} f(s, x_{m-1}(s))\, ds, \qquad m > 0, \end{aligned} \tag{2.4}$$

and then to show that $x(t) = \lim_{m\to\infty} x_m(t)$ exists and satisfies the integral equation. We show that this works in a series of claims.

Claim 1. Suppose that we have defined $x_0, \ldots, x_m$ by (2.4) as continuous functions on J. Then $(t, x_m(t)) \in R$ for all $t \in J$, so that x_{m+1} is well defined. (x_{m+1} is then clearly continuous.)

Pf: For m=0 this is trivial: $(t, x^0) \in R$ for all $t \in J$ because $|t - t_0| \le A \le a$ on J. For general m we have by induction that for all $t \in J$,

$$|x_m(t) - x^0| = \Bigl|\int_{t_0}^{t} f(s, x_{m-1}(s))\, ds\Bigr| \le \Bigl|\int_{t_0}^{t} |f(s, x_{m-1}(s))|\, ds\Bigr| \le M|t - t_0| \le MA \le b,$$

from which the conclusion follows. (The extra set of absolute value signs in the third term in this chain of inequalities is needed because if $t < t_0$ then $\int_{t_0}^{t} |f(s, x_{m-1}(s))|\, ds \le 0$.)

Claim 2. The functions x_m satisfy, for all $m \ge 0$ and all $t \in J$,

$$|x_{m+1}(t) - x_m(t)| \le \frac{Mk^m}{(m+1)!}|t - t_0|^{m+1}. \tag{2.5}$$

Pf: For notational simplicity we verify (2.5) for $t \ge t_0$; this avoids the sort of extra absolute value signs which appeared in the proof of Claim 1. We argue by induction. When $m = 0$:

$$|x_1(t) - x_0(t)| = \Bigl|\int_{t_0}^{t} f(s, x^0)\, ds\Bigr| \le \int_{t_0}^{t} |f(s, x^0)|\, ds \le M(t - t_0),$$

which is (2.5) in this case. For general m,

$$\begin{aligned}
|x_{m+1}(t) - x_m(t)| &= \Big| \int_{t_0}^{t} \big[f(s, x_m(s)) - f(s, x_{m-1}(s)) \big]\, ds \Big| \\
&\le \int_{t_0}^{t} \big| f(s, x_m(s)) - f(s, x_{m-1}(s)) \big|\, ds \\
&\le \int_{t_0}^{t} k \big| x_m(s) - x_{m-1}(s) \big|\, ds \\
&\le k \frac{M k^{m-1}}{m!} \int_{t_0}^{t} (s - t_0)^m\, ds \\
&= \frac{M k^m}{(m+1)!} (t - t_0)^{m+1}.
\end{aligned}$$

This completes the inductive step.

Claim 3. The sequence $\{x_n(t)\}$ converges uniformly on J to some function $x(t)$.

Pf: We define

$$x(t) = x_0(t) + \sum_{i=0}^{\infty} \big(x_{i+1}(t) - x_i(t) \big).$$

This infinite series converges uniformly on J by the Weierstrass M-test; in fact, since $|t - t_0| < A$, the i^{th} term is bounded by $(Mk^i/(i+1)!)|A|^{i+1}$, and

$$\sum_{i=0}^{\infty} \frac{Mk^i}{(i+1)!} |A|^{i+1} = Mk^{-1} \big(\exp(k|A|) - 1 \big).$$

But the m^{th} partial sum of this telescoping series is just x_{m+1}, so that uniform convergence of the series to $x(t)$ is just uniform convergence of the sequence $\{x_m(t)\}$ to $x(t)$.

Claim 4. $(t, x(t))$ lies in R for all $t \in J$. Moreover, the function $x(t)$ is continuous and satisfies the integral equation (2.2) on J.

Pf: Since $|x_m(t) - x^0| \le b$ for $t \in J$ and $x_m(t) \to x(t)$, $|x(t) - x^0| \le b$ also, i.e., $(t, x(t)) \in R$. x is continuous because it is the uniform limit of continuous functions. Since f is continuous and hence uniformly continuous on the compact set R, the uniform convergence of x_m to x on J implies the uniform convergence of $f(t, x_m(t))$ to $f(t, x(t))$ on J: given $\epsilon > 0$ there is a $\delta > 0$ such that for $(\tilde{t}, \tilde{x}) \in R$ and $(t, x) \in R$, $|f(\tilde{t}, \tilde{x}) - f(t, x)| < \epsilon$ whenever $|(\tilde{t} - t, \tilde{x} - x)| < \delta$, and an $N > 0$ such that for $t \in J$, $|x(t) - x_m(t)| < \delta$ whenever $m \ge N$, so that for $t \in J$, $|f(t, x(t)) - f(t, x_m(t))| < \epsilon$ whenever $m \ge N$. Because of this uniform convergence we may, in taking the $m \to \infty$ limit of equation (2.4),

$$x_m(t) = x^0 + \int_{t_0}^{t} f(s, x_{m-1}(s))\, ds,$$

interchange the limit with the integral, yielding (2.2):

$$x(t) = x^0 + \int_{t_0}^{t} f(s, x(s))\, ds.$$

Specifically, given $\epsilon > 0$ and taking N as above, we have for $m \geq N$,

$$\Big|\int_{t_0}^t f(s,x(s))\,ds - \int_{t_0}^t f(s,x_m(s))\,ds\Big| \leq \Big|\int_{t_0}^t |f(s,x(s)) - f(s,x_m(s))|\,dt\Big| < \epsilon|t-t_0|,$$

so that $\lim_{m\to\infty}\int_{t_0}^t f(s,x_{m-1}(s))\,ds = \int_{t_0}^t f(s,x(s))\,ds$. ■

From the Picard Theorem we immediately have the

Proof of Theorem 2.1, part (a): For any point $(t_0,x^0) \in D$ we may find a closed rectangle $R \subset D$ centered at (t_0,x^0), i.e., of the form (2.3). Because R is compact, f satisfies a Lipschitz condition on R. Taking M, k, and A as in the Picard Theorem, we obtain from that theorem a solution of the integral equation (2.2) on $[t_0 - A, t_0 + A]$ and hence, by Lemma 2.2, a solution of IVP on the open interval $I = (t_0 - A, t_0 + A)$. ■

We now turn to the question of uniqueness, i.e., to part (b) of Theorem 2.1. One approach to the uniqueness question uses arguments similar to those of the Picard Theorem; see, for example, the text by Cronin. Here we use an alternate proof technique based on a useful inequality due to Gronwall (see, e.g., [Hartman]):

Lemma 2.4: *Let $u(t)$ and $h(t)$ be continuous functions defined on a closed interval $[\alpha,\beta]$, with $h \geq 0$, let C be a non-negative constant, and suppose that*

$$u(t) \leq C + \int_\alpha^t u(s)h(s)\,ds \tag{2.6}$$

for all t in the interval. Then

$$u(t) \leq Ce^{\int_\alpha^t h(s)\,ds} \tag{2.7}$$

for all t in the interval.

Note in particular that if $C = 0$, then $u(t) \leq 0$ for all t.

Proof: Let us write $U(t) = C + \int_\alpha^t u(s)h(s)\,ds$. By the Fundamental Theorem of Calculus and (2.6), U is differentiable and

$$U'(t) = u(t)h(t) \leq U(t)h(t). \tag{2.8}$$

Now if (2.8) were a differential equation rather than a differential inequality, we would solve it by multiplying by the integrating factor $\mu(t) = \exp\Big[-\int_\alpha^t h(s)\,ds\Big]$. In fact however, the same method works on the inequality; multiplying (2.8) by μ and rearranging leads to $(\mu U)'(t) \leq 0$, and integrating this inequality yields

$$\mu(s)U(s)\Big|_{s=\alpha}^{s=t} = \mu(t)U(t) - C \leq 0,$$

and hence

$$u(t) \leq U(t) \leq C[\mu(t)]^{-1},$$

which is (2.7). ∎

Proof of Theorem 2.1, part (b): Subtracting the integral equations (2.2) satisfied by x and $\tilde{x}$ leads to

$$x(t) - \tilde{x}(t) = \int_{t_0}^{t} \big(f(s, x(s)) - f(s, \tilde{x}(s))\big)\, ds, \tag{2.9}$$

for any $t \in I \cap \tilde{I}$. Now fix $t_1 \in I \cap \tilde{I}$, with $t_1 > t_0$ for definiteness (the proof for $t_1 < t_0$ is the same), and let k be a Lipschitz constant for f on the compact set which is given by $\{(t, x(t)) \mid t \in [t_0, t_1]\} \cup \{(t, \tilde{x}(t)) \mid t \in [t_0, t_1]\}$. With $u(t) = |x(t) - \tilde{x}(t)|$, (2.9) leads to

$$u(t) \le \int_{t_0}^{t} k u(s)\, ds$$

for all $t \in [t_0, t_1]$, and Gronwall's inequality then implies that u vanishes identically in that interval. Since t_1 is arbitrary, u must vanish in $I \cap \tilde{I}$. ∎

2.2 Existence without a Lipschitz condition

In this section we will show that the existence of solutions of IVP may in fact be established with no additional hypotheses on f beyond continuity. That is, we prove a version of Theorem 2.1 in which the Lipschitz hypothesis on f is removed and there is no conclusion of uniqueness:

Theorem 2.5: *Suppose that D and f are as in the statement of IVP. Then for any $(t_0, x^0) \in D$ there exists a solution $x : I \to \mathbb{R}^n$ of (2.1) defined on some open interval I containing t_0.*

Once again, this existence theorem will follow immediately from a fundamental Picard-like existence theorem in a rectangle. The proof of the latter uses one of the basic results of analysis, known variously as Ascoli's Theorem or the Ascoli-Arzela Theorem, which we now recall.

Definition 2.3: Suppose that $S \subset \mathbb{R}^p$. A sequence $\{f_m\}_{m=1}^{\infty}$ of functions, each mapping S to $\mathbb{R}^q$, is *equicontinuous* if for any $\epsilon > 0$ there is a $\delta > 0$ such that, for any m, $|f_m(x) - f_m(y)| < \epsilon$ whenever $x, y \in S$ satisfy $|x - y| < \delta$.

Note that the definition says that, given ϵ, the same δ works for all x, y and for all the functions f_m. In particular, a member of an equicontinuous sequence is uniformly continuous.

Theorem 2.6 (Ascoli-Arzela): *Let $K \subset \mathbb{R}^p$ be compact and let $\{f_m\}$ be an equicontinuous sequence of functions, each mapping K to $\mathbb{R}^q$. Suppose additionally that there is a constant M such that $|f_m(x)| \le M$ for all m and all $x \in K$. Then there exists a subsequence $\{f_{m_k}\}_{k=1}^{\infty}$ which converges uniformly on K to some limit function $f : K \to \mathbb{R}^q$.*

We can now state and prove the fundamental existence result.

Theorem 2.7: *Let $R \subset \mathbb{R} \times \mathbb{R}^n$ be the closed rectangle*

$$R = \{(t, x) \mid |t - t_0| \le a, |x - x^0| \le b\},$$

where $a, b > 0$, and suppose that $f : R \to \mathbb{R}^n$ is continuous. Let M be the maximum of $|f|$ on R and let $A = \min\{a, b/M\}$. Then there exists a function $x(t)$ defined on the (closed) interval $J = [t_0 - A, t_0 + A]$ and satisfying the integral equation (2.2) for all $t \in J$.

Proof: We will construct $x(t)$ for $t \in J_+ \equiv [t_0, t_0 + A]$; the construction for $t < t_0$ is similar. The method is due to Euler and is frequently mentioned in numerical analysis as a simple scheme to construct approximate solutions of an IVP.

For each $m \geq 1$ we subdivide J_+ into m subintervals of the form $[t_{k-1}^{(m)}, t_k^{(m)}]$, where $t_k^{(m)} = t_0 + Ak/m$ for $k = 1, \ldots, m$, and construct an approximate solution $x_m(t)$ on J_+ which is linear on each subinterval. The construction is by induction on the index k of the subinterval; we first define $x_m(t_0) = x^0$, and then, assuming that we have constructed x_m with $(t, x_m(t)) \in R$ on all intervals $[t_{j-1}^{(m)}, t_j^{(m)}]$ for $j \leq k$, we define

$$x_m(t) = x_m(t_k^{(m)}) + (t - t_k^{(m)}) f(t_k^{(m)}, x_m(t_k^{(m)})) \qquad \text{for} \qquad t_k^{(m)} \leq t \leq t_{k+1}^{(m)}.$$

Note that this definition is chosen so that (i) x_m is continuous at $t_k^{(m)}$, and (ii) on the interval $[t_k^{(m)}, t_{k+1}^{(m)}]$, x_m has derivative $f(t_k^{(m)}, x_m(t_k^{(m)}))$, our best guess at the correct derivative $f(t, x(t))$. In particular, it follows from (i) and (ii) that

$$x_m(t) = x^0 + \int_{t_0}^{t} f^{(m)}(s)\, ds \tag{2.10}$$

for $0 \leq t \leq t_{k+1}^{(m)}$, where

$$f^{(m)}(t) = f(t_j^{(m)}, x_m(t_j^{(m)})), \qquad t \in [t_j^{(m)}, t_{j+1}^{(m)}].$$

Equation (2.10) implies that

$$|x_m(t) - x^0| \leq M|t - t_0| \leq MA \leq b, \tag{2.11}$$

so that $(t, x_m(t)) \in R$ for $t_0 \leq t \leq t_{k+1}^{(m)}$, allowing us to continue the induction. Eventually we construct x_m on all of J_+.

Now from (2.10) it follows that the sequence $\{x_m\}$ is equicontinuous and uniformly bounded on J_+: for $t, t' \in J_+$,

$$|x_m(t) - x_m(t')| = \left| \int_t^{t'} f^{(m)}(t)\, dt \right| \leq M|t - t'|; \tag{2.12}$$

and

$$|x_m(t)| \leq |x_m(t_0)| + |x_m(t) - x_m(t_0)| \leq |x^0| + MA. \tag{2.13}$$

By the Ascoli-Arzela theorem there is thus a subsequence $\{x_{m_j}\}$ which converges uniformly on J_+ to some continuous function x. We claim that $x(t)$ is the desired solution of (2.2). This will follow immediately from (2.10) if we can show that $f^{m_j}(t)$ converges uniformly on J_+ to $f(t,x(t))$; the argument is the same as that given in the proof of Claim 4 of Theorem 2.3.

We now verify the uniform convergence of the sequence $\{f^{m_j}\}$. For notational simplicity we suppose that the sequence $\{x_m\}$ itself converges to x. Then given $\epsilon > 0$, uniform continuity of f on R implies that there exists a $\delta > 0$ such that $|f(t,x)-f(t',x')| < \epsilon$ whenever $|(t-t', x-x')| < \delta$. Now choose m so large that $A/m < \delta/3$, that $MA/m < \delta/3$, and, using the uniformity of convergence of $\{x_m\}$, that $|x_m(t) - x(t)| < \delta/3$ whenever $t \in J_+$. If $t \in J_+$ then $t_k^{(m)} \le t \le t_{k+1}^{(m)}$ for some k, so that by (2.12),

$$\begin{aligned}
|(t_k^{(m)} - t, x_m(t_k^{(m)}) - x(t))| &\le |t_k^{(m)} - t| + |x_m(t_k^{(m)}) - x(t)| \\
&\le |t_k^{(m)} - t| + |x_m(t_k^{(m)}) - x_m(t)| + |x_m(t) - x(t)| \\
&\le \frac{A}{m} + \frac{MA}{m} + \frac{\delta}{3} \le \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta,
\end{aligned}$$

and hence

$$|f^{(m)}(t) - f(t,x(t))| = |f(t_k^{(m)}, x_m(t_k^{(m)})) - f(t,x(t))| < \epsilon. \quad \blacksquare$$

As remarked above, the existence theorem Theorem 2.5 now follows as in the proof of Theorem 2.1.

It will in fact be convenient to have at our disposal a formally stronger statement of existence, showing that under appropriate restrictions the interval of definition of the solution may be chosen uniformly in the initial condition, and the pair $(t, x(t))$ may be required to remain in a compact subset of D.

Corollary 2.8: *Let $D \subset \mathbb{R} \times \mathbb{R}^n$ be open and let $f : D \to \mathbb{R}^n$ be continuous. Then for any compact subset $K \subset D$ there exist open sets U, V, with compact closures, satisfying*

$$K \subset U \subset \overline{U} \subset V \subset \overline{V} \subset D, \tag{2.14}$$

and an $\epsilon > 0$, such that for every $(t_0, x^0) \in U$ a solution $x(t)$ of (2.1) is defined on the interval $I = (t_0 - \epsilon, t_0 + \epsilon)$ and satisfies $(t, x(t)) \in \overline{V}$ for $t \in I$.

Proof: Let U, V be any open sets with compact closures satisfying (2.14). Let

$$\eta = d(\overline{U}, V^c) \equiv \inf_{x \in \overline{U}, y \in V^c} |x - y|,$$

where V^c denotes the complement of V and $d(A,B)$ is the distance between two sets. Because $\overline{U}$ is compact and V^c is closed, η is strictly positive. Let $M = \sup_{\overline{V}} |f(x)|$.

Now if $(t_0, x^0) \in U$, then the rectangle

$$\{(t,x) \mid |t - t_0| \le \eta/\sqrt{2}, |x - x^0| \le \eta/\sqrt{2}\}$$

is contained in $\overline{V}$. The stated result, with $\epsilon = \min\{\eta/\sqrt{2}, \eta/\sqrt{2}M\}$, now follows from Theorem 2.7, the fundamental existence result proved above. $\blacksquare$

2.3 Continuity in the initial conditions

The essence of the continuity in the initial conditions—continuity for small values of t—is quite simple, and may be established in various ways. Here we will obtain it twice: first from Gronwall's inequality and the assumption that f is Lipschitzian, and then as a consequence of the uniqueness of solutions, without using the Lipschitz condition directly. It is then a technical exercise to obtain continuity in the entire domain of the solution; in the course of doing so we will also show that this domain is open.

Throughout this section we take D to be an open subset of $\mathbb{R} \times \mathbb{R}^n$ and $f : D \to \mathbb{R}^n$ to be continuous. We will say that the o.d.e. $x' = f(t, x)$, (2.1a), *has unique solutions* if for any $(t_0, x^0) \in D$ and any two solutions x, $\tilde{x}$ of (2.1) defined on intervals I and $\tilde{I}$, respectively, we must have $x(t) = \tilde{x}(t)$ for $t \in I \cap \tilde{I}$. The following remark establishes further terminology.

Remark 2.4: Suppose that the o.d.e. (2.1a) has unique solutions. Then:

(a) There must exist, for each $(t_0, x^0) \in D$, an open interval $I_{(t_0,x^0)}$ containing t_0 and a solution $\hat{x}(t; t_0, x^0)$ of IVP defined for $t \in I_{(t_0,x^0)}$, with the following property: if $x(t)$ is a solution of IVP defined on some interval I, then $I \subset I_{(t_0,x^0)}$, and $x(t) = \hat{x}(t; t_0, x^0)$ for $t \in I$. Specifically, let $I_{(t_0,x^0)}$ be the set of all t such that a solution x of IVP exists on some interval I with $t \in I$, and for such t let $\hat{x}(t; t_0, x^0) = x(t)$. The uniqueness assumption implies that $\hat{x}$ is well defined by this rule, and Theorem 2.5 that $I_{(t_0,x^0)}$ is not empty.

(b) We will denote by E the domain of the function $\hat{x}$ in all its variables:

$$E = \{(t; t_0, x^0) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \mid t \in I_{(t_0,x^0)}\}.$$

We emphasize that the function $\hat{x}$ contains all solutions of the initial value problem for all possible choices of t_0 and x^0. Note that, if $K \subset D$ is compact, U, V, and ϵ are as in Corollary 2.8, and

$$E^0 \equiv \{(t; t_0, x^0) \mid (t_0, x^0) \in U, |t - t_0| < \epsilon\}, \tag{2.15}$$

then the conclusion of Corollary 2.8 may be stated as: $E^0 \subset E$, and $(t, \hat{x}(t; t_0, x^0)) \in \overline{V}$ for $(t; t_0, x^0) \in E^0$.

Remark 2.5: It is easy to verify, under the hypotheses of the above remark, that if $t > t_0$, $t \in I_{(t_0,x^0)}$, and s satisfies $t_0 < s < t$, then $s \in I_{(t_0,x^0)}$, $t \in I_{(s,\hat{x}(s;t_0,x^0))}$, and

$$\hat{x}(t; t_0, x^0) = \hat{x}(t; s, \hat{x}(s; t_0, x^0)). \tag{2.16}$$

Conversely, if the right hand side of (2.16) is defined for some s satisfying $t_0 < s < t$—that is, if $s \in I_{(t_0,x^0)}$ and $t \in I_{(s,\hat{x}(s;t_0,x^0))}$—then $t \in I_{(t_0,x^0)}$ and (2.16) holds. Results for $t < t_0$ are similar.

We now give our first proof of continuity for small times, using Gronwall's inequality and a Lipschitz hypothesis.

Lemma 2.9: *Suppose that D and f satisfy the hypotheses of Theorem 2.1, that is, that f is locally Lipschitzian in x on D. Take $K \subset D$ compact and let U, V, and ϵ be as in Corollary 2.8. Then $\hat{x}$ is continuous on the domain E^0 of (2.15).*

Proof: From Corollary 2.8 we know that if $(s; s_0, y^0) \in E^0$, then

$$|f(s, \hat{x}(s; s_0, y^0))| \le M \equiv \sup_{\overline{V}} |f|.$$

For (t_0, x^0) and $(\bar{t}_0, \bar{x}^0)$ in U, we write

$$\begin{aligned} x(s) &\equiv \hat{x}(s; t_0, x^0), \\ \bar{x}(s) &\equiv \hat{x}(s; \bar{t}_0, \bar{x}^0), \end{aligned}$$

so that

$$\begin{aligned} x(s) &= x^0 + \int_{t_0}^{s} f(\tau, x(\tau))\, d\tau, \\ \bar{x}(s) &= \bar{x}^0 + \int_{\bar{t}_0}^{s} f(\tau, \bar{x}(\tau))\, d\tau, \end{aligned}$$

for $|s - t_0| < \epsilon$ and $|s - \bar{t}_0| < \epsilon$, respectively. Now choose t and $\bar{t}$ so that $|t - t_0| < \epsilon$ and $|\bar{t} - \bar{t}_0| < \epsilon$; for convenience, take $t \ge t_0$. We will prove continuity at $(t; t_0, x^0)$ by showing that $|\bar{x}(\bar{t}) - x(t)| = |\hat{x}(\bar{t}; \bar{t}_0, \bar{x}^0) - \hat{x}(t; t_0, x^0)|$ is small when $|\bar{t} - t|$, $|\bar{t}_0 - t_0|$, and $|\bar{x}^0 - x^0|$ are all small; in doing so we may suppose that $|\bar{t}_0 - t_0| < \epsilon - |t - t_0|$, which guarantees that $\bar{x}(s)$ is defined whenever $s \in [t_0, t]$. Then

$$|\bar{x}(\bar{t}) - x(t)| \le |\bar{x}(\bar{t}) - \bar{x}(t)| + |\bar{x}(t) - x(t)|. \tag{2.17}$$

Now

$$|\bar{x}(\bar{t}) - \bar{x}(t)| = \left| \int_{\bar{t}}^{t} f(s, \bar{x}(s))\, ds \right| \le M|\bar{t} - t|. \tag{2.18}$$

On the other hand, setting $u(s) = |\bar{x}(s) - x(s)|$ and using the Lipschitz condition on $\overline{V}$, we have for $s \in [t_0, t]$

$$\begin{aligned} u(s) &= \left| \bar{x}^0 - x^0 + \int_{\bar{t}_0}^{t_0} f(\tau, \bar{x}(\tau))\, d\tau + \int_{t_0}^{s} [f(\tau, \bar{x}(\tau)) - f(\tau, x(\tau))]\, d\tau \right| \\ &\le |\bar{x}^0 - x^0| + M|\bar{t}_0 - t_0| + \int_{t_0}^{s} k u(\tau)\, d\tau, \end{aligned}$$

so that from Gronwall's inequality,

$$u(t) = |\bar{x}(t) - x(t)| \le \big(|\bar{x}^0 - x^0| + M|\bar{t}_0 - t_0|\big)e^{k\epsilon}. \tag{2.19}$$

Equations (2.17)–(2.19) show that $\bar{x}$ is continuous at $(t; t_0, x^0)$. ∎

We now give an alternate proof of continuity for small times which uses only the existence and uniqueness of solutions. The result is of interest for its own sake, and moreover we will apply it in the next section to obtain continuous dependence on parameters. The proof is based on the Ascoli-Arzela Theorem.

Lemma 2.10: *Suppose that the o.d.e. (2.1a) has unique solutions. Take $K \subset D$ compact and let U, V, and ϵ be as in Corollary 2.8. Then $\hat{x}$ is continuous on the domain E^0 of (2.15).*

Proof: Suppose that $(t; t_0, x^0) \in E^0$, that is, that $(t_0, x^0) \in U$ and $|t - t_0| < \epsilon$. We will show that, for any sequence $\{(t_m; t_{0m}, x^{0m})\}$ of points of E^0 with $\lim_{m\to\infty}(t_m; t_{0m}, x^{0m}) = (t; t_0, x^0)$, we have $\lim_{m\to\infty} \hat{x}(t_m; t_{0m}, x^{0m}) = \hat{x}(t; t_0, x^0)$.

Let J be the interval $J = \{s \mid |s - t_0| \le \epsilon'\}$, where $\epsilon' < \epsilon$ and ϵ' is large enough that J contains t. For m sufficiently large that $|t_{0m} - t_0| < (\epsilon - \epsilon')$, $x_m(s) \equiv \hat{x}(s; t_{0m}, x^{0m})$ will be defined for all $s \in J$. This sequence of functions is uniformly bounded because the points $(s, x_m(s))$ all lie in the compact set $\overline{V}$. From the integral equation,

$$x_m(s) = x^{0m} + \int_{t_{0m}}^{s} f(s, x_m(s))\, ds, \tag{2.20}$$

we obtain equicontinuity of the sequence: for $s, s' \in J$ with $s \le s'$,

$$|x_m(s') - x_m(s)| \le \int_{s}^{s'} |f(s'', x_m(s''))|\, ds'' \le M|s' - s|.$$

where $M = \sup_{\overline{V}} |f(x)|$. Hence, by Theorem 2.6, we may extract a subsequence $\{x_{m_k}\}$ which converges uniformly on J to a limit function $x(s)$. Each x_{m_k} satisfies (2.20); we take the $k \to \infty$ limit in this equation, using the uniform continuity of f on $\overline{V}$ to establish uniform convergence of $\{f(s, x_{m_k}(s))\}$ to $f(s, x(s))$ and thus justify an interchange of the limit with the integral, and conclude that x also satisfies the integral equation. By the uniqueness hypothesis, $x(s) = \hat{x}(s; t_0, x^0)$ for $|s - t_0| < \epsilon'$, and hence, by continuity, on J.

We have established that $\lim_{k\to\infty} x_{m_k}(s) = \hat{x}(s; t_0, x^0)$ on J. To see that the same result holds for the original sequence $\{x_m\}$, note that the above reasoning implies that for any $s \in J$ an arbitrary subsequence of $\{x_m(s)\}$ has itself a subsequence converging to $\hat{x}(s; t_0, x^0)$; this is possible only if $\{x_m\}$ itself converges to $\hat{x}$. In particular, $\lim_{k\to\infty} x_m(t) = \hat{x}(t; t_0, x^0)$. ■

Finally we give the complete (global) continuity result.

Theorem 2.11: *Suppose that the o.d.e. (2.1a) has unique solutions. Then E is open and $\hat{x}$ is continuous on E.*

Proof: Fix $(t_0, x^0) \in D$. For any $t \in I_{(t_0,x^0)}$ we must find a neighborhood of $(t; t_0, x^0)$ on which $\hat{x}$ is defined and continuous. We write $I_{(t_0,x^0)} \equiv (a, b)$.

For $t = t_0$, and hence for t near t_0, such a neighborhood exists by Lemma 2.10. To be concrete, we may take the compact set in this lemma to be $K_1 \equiv \{(t_0, x^0)\}$; the lemma then guarantees the existence of sets U_1 and V_1 and of some positive number ϵ_1. The desired neighborhood is the set E_1^0 defined by $E_1^0 \equiv \{(t; t_0, x^0) \mid (t_0, x^0) \in U_1, |t - t_0| < \epsilon_1\}$ (compare this with (2.15)).

We now consider the case of general $t > t_0$; the case $t < t_0$ is similar. Let S be the set of numbers $s \ge t_0$ such that, for some neighborhood $W = W_s$ of (t_0, x^0), $\hat{x}$ is defined and continuous on $[t_0, s] \times W_s$. S is not empty since, from the previous paragraph,

$[t_0, t_0 + \epsilon_1) \subset S$; let β be the supremum of S. Clearly $t_0 + \epsilon_1 \le \beta \le b$; we claim $\beta = b$. If this claim is true then the desired neighborhood of $(t; t_0, x^0)$ may be taken to be $(t_0, s) \times W_s$ for any s satisfying $t < s < b$, and we are done.

To prove the claim, suppose that $\beta < b$. Let us write $y(s; t_*, x^*) \equiv (s, \hat{x}(s; t_*, x^*))$. The function $y(s; t_*, x^*)$ is defined on E and for fixed $(t_*, x^*) \in D$ is a continuous function of s, since $\hat{x}(s; t_*, x^*)$ is differentiable and hence continuous in s for fixed (t_*, x^*). In particular, then, the set $K = \{y(s; t_0, x^0) \mid t_0 \le s \le \beta\}$ is the image of the compact set $[t_0, \beta]$ under a continuous map and hence is compact. An application of Corollary 2.8 to K provides a set U with $K \subset U \subset D$ and an $\epsilon > 0$ such that $\hat{x}(t; t_2, x^2)$ is defined for $(t_2, x^2) \in U$ and $0 \le t - t_2 < \epsilon$; without loss of generality we may assume that $\epsilon \le \epsilon_1$.

By definition of β, we know that $\hat{x}$ is defined and continuous on $[t_0, \beta - \epsilon/2] \times W_{\beta-\epsilon/2}$. Thus y is also defined and continuous on this set, and hence $y(\beta - \epsilon/2; t_1, x^1)$ is continuous as a function of (t_1, x^1) in $W_{\beta-\epsilon/2}$. Since $y(\beta - \epsilon/2; t_0, x^0) \in K \subset U$ and U is open, this continuity implies that there exists a neighborhood W of (t_0, x^0), with $W \subset W_{\beta-\epsilon/2}$, such that $y(\beta - \epsilon/2; t_1, x^1) = (\beta - \epsilon/2, \hat{x}(\beta - \epsilon/2; t_1, x^1)) \in U$ whenever $(t_1, x_1) \in W$.

Now by Lemma 2.10, $\hat{x}(t; t_2, x^2)$ is continuous for $(t_2, x^2) \in U$ and $0 \le t - t_2 < \epsilon$. Thus $\hat{x}$, already known to be defined and continuous on $[t_0, \beta] \times W$, is defined on $[\beta - \epsilon/2, \beta + \epsilon/2] \times W$ by

$$\hat{x}(s; t_1, x^1) = \hat{x}(s; (\beta - \epsilon/2), \hat{x}(\beta - \epsilon/2; t_1, x^1)) \tag{2.21}$$

with agreement on the overlap (see Remark 2.5). Moreover, (2.21) presents $\hat{x}$ on this domain as a composition of continuous functions, so that $\hat{x}$ is itself continuous there. Thus every point s in $[t_0, \beta + \epsilon/2)$ lies in the set S, contradicting $\beta \equiv \sup S$. ■

2.4 Dependence of the solution on parameters

We now extend our notion of an initial value problem to include the possibility that the right hand side of the o.d.e. depends on some additional variables, called *parameters*. Thus we assume that D is an open, connected subset of $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ and that $f : D \to \mathbb{R}^n$ is continuous. We write points in $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ as (t, x, μ); the variables $(\mu_1, \dots, \mu_m)$ are the parameters. For $(t_0, x^0, \mu) \in D$ we want to solve

$$x'(t) = f(t, x, \mu), \tag{2.22a}$$

$$x(t_0) = x^0. \tag{2.22b}$$

We will say that f is locally Lipschitzian with respect to x in D if for any compact $K \subset D$ there is a constant k such that

$$|f(t, x^1, \mu) - f(t, x^2, \mu)| \le k|x^1 - x^2|,$$

whenever $(t, x^1, \mu), (t, x^2, \mu) \in K$; note that the constant k is uniform in μ as well as in t.

The existence of solutions for (2.22) follows immediately from the existence proved in Section 2.2; the parameters are simply held constant and play no role. Moreover, the proof of the the uniform existence result Corollary 2.8 extends almost word for word to obtain an existence result uniform in the parameters as well as the initial conditions. Finally, if we assume that f is locally Lipschitzian in x, then the uniqueness of solutions for (2.22) follows directly from Theorem 2.1, again, the parameters play no role. We summarize in

Theorem 2.12: *Let D and f be as above.*

(a) Suppose that $K \subset D$ is compact. Then there exist open sets U, V, with compact closures, satisfying

$$K \subset U \subset \overline{U} \subset V \subset \overline{V} \subset D,$$

and an $\epsilon > 0$, such that a solution $x(t)$ of (2.22) exists and satisfies $(t, x(t), \mu) \in \overline{V}$ whenever $(t_0, x^0, \mu) \in U$ and $|t - t_0| < \epsilon$.

(b) Suppose now that f is locally Lipschitzian in x. If $(t_0, x^0, \mu) \in D$, and x and $\tilde{x}$ are solutions of (2.22) defined on intervals I and $\tilde{I}$, respectively, then $x(t) = \tilde{x}(t)$ for $t \in I \cap \tilde{I}$.

As in Remark 2.4, when (2.22a) has unique solutions we may for $(t_0, x^0, \mu) \in D$ define a solution $\hat{x}(t; t_0, x^0, \mu)$ of (2.22) on some maximal interval $I_{(t_0,x^0,\mu)}$, and introduce the domain

$$E = \{(t; t_0, x^0, \mu) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \mid t \in I_{(t_0,x^0,\mu)}\}. \tag{2.23}$$

We would like to prove a result corresponding to Theorem 2.11, that is, that E is open and $\hat{x}$ is continuous on E.

There is a clever transformation which reduces this to the previous problem: we may change the parameter dependence in (2.22) into a dependence on initial conditions. To do so, we introduce a new initial value problem with dependent variable $z = (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$\begin{aligned} x'(t) &= f(t, x, y), \\ y'(t) &= 0, \end{aligned} \qquad (2.24a) \qquad\qquad \begin{aligned} x(t_0) &= x^0, \\ y(t_0) &= \mu. \end{aligned} \qquad (2.24b)$$

It is easy to verify that

(i) $z(t) = (x(t), y(t))$ is a solution of (2.24) if and only if $x(t)$ is a solution of (2.22) and $y(t) \equiv \mu$, hence

(ii) if (2.22a) has unique solutions, then so does (2.24a), and

(iii) in this case, the natural domain (see Remark 2.4) of the solution $\hat{z}$ of (2.24) is precisely the domain E of (2.23), and $\hat{z}$ satisfies $\hat{z}(t; t_0, x^0, \mu) = \hat{x}(t; t_0, x^0, \mu)$.

Theorem 2.11 thus immediately implies

Theorem 2.13: *Suppose that $D \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ is open, that $f : D \to \mathbb{R}^n$ is continuous, and that the o.d.e. $x' = f(t, x, \mu)$ ((2.22a)) has unique solutions. Then the domain E (see (2.23)) of the solution $\hat{x}$ is open and $\hat{x}$ is continuous on E.*

Remark 2.6: Of course, Theorem 2.12 implies that the conclusions of Theorem 2.13 hold in particular if f is locally Lipschitzian in x. It is important to realize the role played by Lemma 2.10 in reaching this conclusion. We cannot deduce local continuity for the IVP (2.24) from Lemma 2.9 unless we assume that the function $f(t, x, y)$ is Lipschitzian in the variable $z = (x, y)$, so that the right hand side of (2.24a) will satisfy our usual Lipschitz hypothesis. A local Lipschitz condition in x alone, however, guarantees uniqueness through Theorem 2.12, and from this Lemma 2.10 implies local and hence global continuity.

2.5 Differentiability of the solution in the initial conditions

We begin by recalling some elementary facts about the differentiation of functions of several real variables. Suppose that U is an open subset of $\mathbb{R}^m$ and that $g : U \to \mathbb{R}^p$. We know how to define the partial derivatives: for $i = 1, \ldots, p,$, $j = 1, \ldots, m$, and $x \in U$,

$$\frac{\partial g_i}{\partial x_j}(x) = \lim_{h\to 0} \frac{g(x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_m) - g(x_1, \ldots, x_j, \ldots, x_m)}{h}.$$

More generally, we may define partial derivatives of order k:

$$\frac{\partial^k g_i}{\partial x_{j_1} \cdots \partial x_{j_k}}(x).$$

We say that g belongs to class $C^k(U)$ ($k \geq 0$) if g itself and all partial derivatives of g of order k', $k' \leq k$ exist and are continuous in U. In particular, $g \in C^0(U)$ if g is continuous.

We will write Dg for the $p \times m$ matrix of first partial derivatives of g, when these all exist:

$$Dg_x = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(x) & \ldots & \frac{\partial g_1}{\partial x_m}(x) \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial x_1}(x) & \ldots & \frac{\partial g_p}{\partial x_m}(x) \end{bmatrix}.$$

By convention, the point x at which the derivatives are evaluated is frequently written as a subscript, so that the notation $Dg_x u$ may be used to denote the matrix Dg_x acting on the (column) vector u:

$$\big[Dg_x u\big]_i = \sum_{j=1}^{m} \frac{\partial g_i}{\partial x_j}(x) u_j.$$

One advantage of this notation is the natural expression it gives for the chain rule as a matrix product: if $h : \mathbb{R}^p \to \mathbb{R}^q$ and $G(x) = h(g(x))$, then $DG_x = Dh_{g(x)} Dg_x$. If a function F depends on two sets of variables, say $F = F(x, y)$ with $x \in \mathbb{R}^m, y \in \mathbb{R}^n$, we will write $D_x F$ and $D_y F$, respectively, for the matrices of derivatives with respect to these two variables.

Recall that the function g is *differentiable* at a point $x \in U$ if there is an $p \times m$ matrix A, called the derivative of g at x, such that error in the linear approximation $g(x + v) \simeq g(x) + Av$ vanishes as $v \to 0$ faster than linearly:

$$\lim_{v\to 0} \frac{|g(x + v) - g(x) - Av|}{|v|} = 0.$$

A basic result, whose proof we omit, is

Theorem 2.14: *If g is differentiable at a point $x \in U$ then all partial derivatives $(\partial g_i/\partial x_j)(x)$ exist, and Dg_x is the derivative of g at x in the above sense. Conversely, if $g \in C^1(U)$ then g is differentiable at each $x \in U$.*

Finally, if A is any $p \times m$ matrix, we write $\|A\|$ for the *norm* (the *operator norm*) of A, the smallest real number such that $|Ax| \le \|A\||x|$ for all x:

$$\|A\| = \sup_{x \in \mathbb{R}^m, x \neq 0} \frac{|Ax|}{|x|}.$$

Remark 2.7: (a) It is easy to verify that $\|\cdot\|$ has the standard properties of a norm: $\|A\| \ge 0$, with $\|A\| = 0$ if and only if $A = 0$, $\|cA\| = |c|\|A\|$ for $c \in \mathbb{R}$, and $\|A + B\| \le \|A\| + \|B\|$.

(b) It is not obvious how to compute explicitly the norm $\|A\|$ of a given matrix A. On the other hand, since A has pm matrix elements A_{ij}, for $i = 1, \dots p$ and $j = 1 \dots m$, it may be regarded as a point in $\mathbb{R}^{pm}$ and hence has an ℓ^2 (Euclidean) norm $|A|_2 = \left(\sum_{ij} A_{ij}^2\right)^{1/2}$. Moreover, since using the Cauchy-Schwarz inequality we have

$$|Ax| = \left(\sum_{i=1}^{p}\Bigl(\sum_{j=1}^{m} A_{ij}x_j\Bigr)^2\right)^{1/2} \le \left(\sum_{i=1}^{p}\Bigl(\sum_{j=1}^{m} A_{ij}^2\Bigr)\Bigl(\sum_{j=1}^{m} x_j^2\Bigr)\right)^{1/2} = |A|_2|x|,$$

certainly $\|A\| \le |A|_2$. Since we can compute $|A|_2$, this is a useful bound. In particular, it implies that the norm $\|A\|$ is a continuous function of A (that is, a continuous function on $\mathbb{R}^{pm}$), since $\bigl|\|A\| - \|B\|\bigr| \le \|A - B\| \le |A - B|_2$.

(c) Any C^1 function is locally Lipschitzian; this is frequently the easiest way to check the Lipschitz property. For if $g \in C^1(U)$ and $x^0 \in U$, choose a compact ball $K = \{x \mid |x - x^0| \le \delta\} \subset U$ and let $k = \sup_{x \in K} \|Dg_x\|$. ($k$ is finite because the norm is a continuous function of the matrix entries, so that $\|Dg_x\|$ is continuous on K.) For $x, y \in K$ we have by the chain rule

$$|g(y) - g(x)| = \left|\int_0^1 \frac{d}{d\tau}[g(x + \tau(y - x))]\, d\tau\right| = \left|\int_0^1 Dg_{x+\tau(y-x)}(y - x)\, d\tau\right| \le k|y - x|.$$

Satisfaction of a Lipschitz condition on such a ball for each x^0 implies a local Lipschitz condition; this is essentially Remark 2.1(c). Similarly, a continuous function $f(t, x)$ which has (jointly) continuous derivatives with respect to x is locally Lipschitzian with respect to x.

Now we turn to our main result in this section. To avoid a profusion of subscripts and superscripts we will write the initial conditions as $(T, X) \in D$.

Theorem 2.15: *Let $D \subset \mathbb{R} \times \mathbb{R}^n$ be open and connected and let $f : D \to \mathbb{R}^n$ belong to $C^k(D)$, $k \ge 1$. Then the solution $\hat{x}(t; T, X)$ of (2.1) is of class C^k on its domain E. Moreover, each derivative of order k has a continuous t derivative.*

Remark 2.8: (a) Because the hypothesis $f \in C^k(D)$, $k \ge 1$, implies that f is locally Lipschitzian in D (and hence locally Lipschitzian with respect to x in D), we know that $\hat{x}$ is well defined and continuous on its (open) domain E.

(b) The existence of an extra derivative in the variable t should not be surprising; even if f is only continuous, we know that the solution $\hat{x}$, if it exists, is differentiable in t, by (2.1).

Note that, by applying this theorem to (2.24), we immediately extend the result to initial value problems involving parameters:

Corollary 2.16: *Let $D \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ be open and connected and let $f : D \to \mathbb{R}^n$ belong to $C^k(D)$, $k \geq 1$. Then the solution $\hat{x}(t; T, X, \mu)$ of (2.22) is of class C^k on its domain E. Moreover, each derivative of order k has a continuous t derivative.*

Remark 2.9: In fact, Theorem 2.15 for any particular value of k implies Corollary 2.16 for the same value, a fact that we will use in our inductive proof of the theorem.

Before starting on the formal proof of Theorem 2.15 we give a motivational discussion. Consider the case $k = 1$ and the problem of verifying continuous differentiability in the variable X. Suppose that in fact $D_X \hat{x}_{(t;T,X)}$ exists for all $(t; T, X)$ in E; it is an $n \times n$ matrix which for the moment we write as $J(t; T, X)$. By taking the X derivative of the initial condition $\hat{x}(T; T, X) = X$ we find that

$$J(T; T, X) = I, \tag{2.25a}$$

where I is the $n \times n$ identity matrix. Similarly, applying D_X to the differential equation $\hat{x}'(t; T, X) = f(t, \hat{x}(t; T, X))$ and using the chain rule yields

$$J'(t; T, X) = D_x f_{(t, \hat{x}(t;T,X))} J(t; T, X). \tag{2.25b}$$

(2.25) represents an initial value problem which has the matrix J as its solution. Note that the o.d.e. (2.25b) is linear in the variable J.

It turns out that the way to prove that $\hat{x}$ is differentiable in X is to find the derivative J by solving this initial value problem. We want to emphasize this point: from now on, until the conclusion of the proof of Theorem 2.15, we will regard J as defined to be the solution of (2.25). Once we have shown that this solution exists, we finish the proof by showing that the resulting function J is in fact the derivative $D_X \hat{x}$.

There are two possible approaches. The first is to regard the original problem (2.1) together with (2.25) as a single IVP for the pair of variables $\hat{x}$ and J. It follows from Corollary 2.8 that this new IVP has a solution $\big(x(t), J(t)\big)$. If it is applicable to the joint (x-J) problem, the Picard method of course likewise yields existence of a solution but, more importantly, also shows easily that $J = D_X \hat{x}$ (see Cronin). The drawback of this approach is that the joint problem is not locally Lipschitzian, and hence the Picard method is not applicable, without additional assumptions—in particular, we need $D_x f$ to be locally Lipschitzian in x. One way to achieve this is to suppose that f is actually of class C^2; that is, to tolerate the loss of one derivative in going from f to the solution $\hat{x}$.

To avoid losing a derivative we will take an alternative approach. We think of first solving the original IVP, so that $\hat{x}$ in (2.25b) may be considered a known function. Now (2.25) is easy to solve, and the problem is to prove that the solution really is the derivative $D_X \hat{x}$. To prepare for our proof, we first quote a fundamental existence and uniqueness theorem for *linear* differential equations (see Remark 1.2) which is needed to treat (2.25), and then give a technical lemma.

Theorem 2.17: *Suppose that $A(t)$ is an $n \times n$ matrix with entries $a_{ij}(t)$, $i, j = 1, \ldots, n$, defined and continuous on an open interval $I \subset \mathbb{R}$, and that $u : I \to \mathbb{R}^n$ is continuous. Then for $t_0 \in I$ and $x^0 \in \mathbb{R}^n$ the initial value problem*

$$x' = A(t)x + u(t), \qquad x(t_0) = x^0, \tag{2.26}$$

has a unique solution defined for all $t \in I$.

Proof: The right hand side of (2.26) is easily seen to be locally Lipschitzian, so our results on existence, uniqueness, and continuity apply. The point of the theorem is that for linear equations the interval of definition of the solution is independent of the initial conditions and is as large as possible. We will prove this theorem in Chapter III, when we study linear equations.

Lemma 2.18: *Suppose that $f \in C^1(D)$, and that $K \subset D$ is a compact set such that every cross section $K_t = \{x \mid (t,x) \in K\}$ is convex. Then there exists a continuous $n \times n$ matrix function $F(t,x,y)$ defined on $\hat{K} \equiv \{(t,x,y) \mid x, y \in K_t\}$ such that*

$$f(t,y) - f(t,x) = [D_x f_{(t,x)} + F(t,x,y)](y - x)$$

and $\|F(t,x,y)\| \le \eta(|x-y|)$, with η an increasing function on the positive reals which satisfies $\lim_{r\to 0} \eta(r) = 0$.

Proof: For $(t,x,y) \in \hat{K}$ we write

$$\begin{aligned} f(t,y) - f(t,x) - D_x f_{(t,x)}(y-x) &= \int_0^1 \frac{d}{d\tau}\big[f(t, x + \tau(y-x))\big]\, d\tau - D_x f_{(t,x)}(y-x) \\ &= \left[\int_0^1 \big[D_x f_{(t,x+\tau(y-x))} - D_x f_{(t,x)}\big]\, d\tau\right](y-x) \\ &\equiv F(t,x,y)(y-x), \end{aligned}$$

thus defining F (which is clearly continuous). Set $\eta(r) = \sup_{(t,x,y)\in\hat{K}, |x-y|\le r} \|F(t,x,y)\|$. $D_x f$ is continuous and hence uniformly continuous on K, so that for any $\epsilon > 0$ there exists a $\delta > 0$ such that $\|D_x f_{(t,z)} - D_x f_{(t,x)}\| < \epsilon$ whenever $|z - x| < \delta$. From this it follows immediately that $\eta(r) < \epsilon$ when $r < \delta$. ∎

Proof of Theorem 2.15: We begin with the case $k = 1$; note that $f \in C^1(D)$ implies that $f(t,x)$ is locally Lipschitzian with respect to x, so that $\hat{x}$ exists and is continuous on its natural domain E. We first show that $\hat{x}(t;T,X)$ is continuously differentiable in X. Consider the IVP (2.25); here X plays the role of a parameter and T is both a parameter (because $\hat{x}$ depends on T) and an initial time. As pointed out in the proof of Theorem 2.17 the right hand side of (2.25b) is locally Lipschitz with respect to J; hence, by Theorem 2.13, the solution $J(t;T,X)$ is continuous on its natural domain of definition. Theorem 2.17 then implies that this domain is in fact E (to apply the theorem directly one may think of (2.25) as n separate problems for the n columns of J). Thus we may finish the proof of this case by showing that $D_X\hat{x}$ exists and is equal to J on E. According to Theorem 2.14, it suffices to choose $(T,X) \in D$ and $[a,b] \in I_{(T,X)}$ and then show that, for all $t \in [a,b]$,

$$\lim_{Y\to 0} |Y|^{-1}|u(Y)| = 0, \tag{2.27}$$

where $u(Y) = u(t;T,X,Y)$ is given by

$$u(Y) = \hat{x}(t;T,X+Y) - \hat{x}(t;T,X) - J(t;T,X)Y.$$

In order to be able to apply Lemma 2.18 we choose ϵ sufficiently small that the compact set $K = \{(s,x) \mid s \in [a,b], |x - \hat{x}(s;T,X)| \le \epsilon\}$ is contained in D; then by the continuity of $\hat{x}$ we have $\hat{x}(s;T,X+Y) \in K$ for $s \in [a,b]$ and $|Y|$ sufficiently small. The integral equations

$$\hat{x}(t;T,X+Y) = X + Y + \int_T^t f(s,\hat{x}(s;T,X+Y))\,ds,$$

$$\hat{x}(t;T,X) = X + \int_T^t f(s,\hat{x}(s;T,X))\,ds,$$

and

$$J(t;T,X) = I + \int_T^t D_x f_{(s,\hat{x}(s;T,X))} J(s;T,X)\,ds,$$

yield, for $|Y|$ sufficiently small,

$$\begin{aligned} u(Y) &= \int_T^t \bigl\{ f(s,\hat{x}(s;T,X+Y)) - f(s,\hat{x}(s;T,X)) - (D_x f)JY \bigr\}\,ds \\ &= \int_T^t \bigl\{ [D_x f + F]\bigl[u(Y) + JY\bigr] - (D_x f)JY \bigr\}\,ds \\ &= \int_T^t \bigl\{ [D_x f + F]u(Y) + FJY \bigr\}\,ds\ , \end{aligned}$$

where we have introduced the shortened notation $D_x f \equiv D_x f_{(s,\hat{x}(s;T,X))}$, $J \equiv J(s;T,X)$, and $F \equiv F(s,\hat{x}(s;T,X),\hat{x}(s;T,X+Y))$. Setting

$$M_1 = \sup_{s\in[a,b]} \|D_x f_{(s,\hat{x}(s;T,X))}\|.$$

$$M_2 = \sup_{s\in[a,b]} \|J(s;T,X)\|,$$

and

$$r = \sup_{s\in[a,b]} |\hat{x}(s;T,X) - \hat{x}(s;T,X+Y)|,$$

and noting that $\|F(s,\hat{x}(s;T,X),\hat{x}(s;T,X+Y))\| \le \eta(r)$, we have

$$|u(Y)| \le \eta(r)M_2(b-a)|Y| + \left|\int_T^t (M_1 + \eta(r))|u(Y)|\,ds\right|.$$

Then Gronwall's inequality implies that

$$|u(Y)| \le \eta(r) M_2 (b-a)|Y| e^{(M_1+\eta(r))|t-T|},$$

which implies (2.27), since $\lim_{|Y|\to 0} \eta(r) = 0$.

Differentiability of $\hat{x}$ in T is proved similarly. We first establish an IVP for the derivative $A(t;T,X) \equiv D_T \hat{x}_{(t;T,X)}$. From the formulae

$$\begin{aligned} \hat{x}(T;T+S,X) &= X + \int_{T+S}^{T} f(s, \hat{x}(s;T+S,X))\, ds, \\ \hat{x}(T;T,X) &= X, \end{aligned}$$

we form a difference quotient and determine that

$$A(T;T,X) = -f(T,X). \tag{2.28a}$$

The o.d.e., derived as for J, is

$$A'(t;T,X) = D_x f_{(t,\hat{x}(t;T,X))} A(t;T,X). \tag{2.28b}$$

From this point the proof follows the lines of that above, that is, we regard A as defined by (2.28) and prove that it is in fact $D_T\hat{x}$. Details will be left as an exercise for the student.

It remains to verify that $\hat{x}$ is of class C^2 in the variable t and that J and A are differentiable in t. The first statement follows directly by differentiation of the formula

$$D_t \hat{x}_{(t;T,X)} = f(t, \hat{x}(t;T,X)) \tag{2.29}$$

and the fact that f and $\hat{x}$ are C^1, the second from equations (2.25b) and (2.28b), which give explicit formulae for the derivatives. This completes the case $k=1$ of Theorem 2.15.

The general case of Theorem 2.15 is proved by induction on k. Suppose that the theorem holds when k is replaced by $k-1$. The quantities $J = D_X\hat{x}$ and $A = D_T\hat{x}$ are solutions, defined on E, of the IVP (2.25) and (2.28), respectively. These are linear equations in which the coefficients of A and J are of class C^{k-1}: they are first derivatives of f, which is C^k, composed with the function $\hat{x}$, which by the induction hypothesis is C^{k-1}. Now the induction hypothesis implies the $k-1$ case of Corollary 2.16; application of this corollary to (2.25) and (2.28) shows that A and J are themselves of class C^{k-1} on E and that their derivatives of order $k-1$ are differentiable in t. This verifies the conclusions of the theorem for derivatives of $\hat{x}$ involving at least one derivative in X or T. Finally, again by differentiation of (2.29), all derivatives of $\hat{x}$ in t of order at most $k+1$ exist and are continuous. ■

2.6 Extension of solutions

In this section we investigate the problem of extending a solution of our IVP, defined on some interval I, to a larger interval—that is, we are interested in determining the maximal interval $I_{(t_0,x^0)}$. Our answers will be rather indirect, consisting of characterizations of how the solution $x(t)$ must behave as t approaches the endpoints of $I_{(t_0,x^0)}$, but they do have applications to certain physical systems, as we shall see. Throughout we will assume that $D \subset \mathbb{R} \times \mathbb{R}^n$ is open and connected, that $f : D \to \mathbb{R}^n$ is continuous, and that the IVP $x' = f(t, x)$ has unique solutions. Since we are not primarily interested in dependence on initial conditions we will usually fix an initial point $(t_0, x^0) \in D$ and write $x(t) \equiv \hat{x}(t; t_0, x^0)$; $x(t)$ has domain $(a, b) \equiv I_{(t_0,x^0)}$. We will not consider parameter dependence, although our results may easily be extended to the case in which parameters are present.

Theorem 2.19: *Let D and f be as above, let $K \subset D$ be compact, and take $(t_0, x^0) \in K$. Suppose that $b < \infty$. Then there is a b' with $t_0 < b' < b$ such that $(t, x(t)) \notin K$ for $b' < t < b$. A similar result holds if $a > -\infty$.*

Proof: Apply Corollary 2.8 to D, f, and K, producing an $\epsilon > 0$ such that $\hat{x}(t; t_1, x^1)$ is defined for $(t_1, x^1) \in K$ and $|t - t_1| < \epsilon$. Let $b' = b - \epsilon$. If $b' < t < b$, but $(t, x(t)) \in K$, then the interval of definition of $x(s)$ can be extended beyond (a, b), in fact to the interval $(a, t + \epsilon)$, by the formula

$$x(s) = \hat{x}(s; t, x(t)), \qquad\qquad t < s < t + \epsilon$$

(see Remark 2.5), contradicting the maximality of (a, b). ■

The theorem may be restated as follows: as t approaches a finite endpoint of the maximal interval $I_{(t_0,x^0)}$, $x(t)$ must leave and remain outside any compact subset of D. Another formulation is

Corollary 2.20: *Let D and f be as above, with $(t_0, x^0) \in D$, $x(t) = \hat{x}(t; t_0, x^0)$, and $(a, b) = I_{(t_0,x^0)}$. For $t \in (a, b)$ let*

$$p(t) = d((t, x(t)), D^c) \equiv \inf_{(t_1,x^1)\notin D} [|t - t_1| + |x(t) - x^1|]$$

be the distance from $(t, x(t))$ to the boundary of D. Then either $b = \infty$, $x(t)$ is unbounded on $[t_0, b)$, or $\lim_{t\to b} p(t) = 0$. A similar result holds with b replaced by a.

Proof: If none of the above alternatives hold, then $b < \infty$ and there exists an $N > 0$ such that $(t, x(t))$ lies in the compact set

$$K = \{(s, y) \mid d((s, y), D^c) \geq 1/N\} \cap \{(s, y) \mid t_0 \leq s \leq b, |y| \leq N\},$$

for points $t \in (a, b)$ arbitrarily close to b, contradicting Theorem 2.19. ■

To give an application of these theorems, we return to the model, considered in Example 1.2, of a system of point particles obeying Newton's laws. To simplify notation we will replace the $3N$ coordinates of the N particles by n undifferentiated coordinates, so that an index i here corresponds to a pair (k, α) in the earlier formulation. With each coordinate i is associated a mass M_i (corresponding to the previous m_k). Not all values of coordinates may be possible; we therefore assume that the configuration of our system at any time is described by a point $x \in U$, where U is some open subset of $\mathbb{R}^n$. We want to consider the case where the force acting on a particle depends only on its position and the position of the other particles, but not explicitly on time, and are thus led to the system of o.d.e.s

$$M_i x_i'' = F_i(x), \qquad i = 1, \ldots, n, \qquad x \in U.$$

Here $F : U \to \mathbb{R}^n$ (the force vector) is a continuous function. As discussed in Example 1.2, we may reformulate the problem as a first order system for $2n$ variables $(x, v) \in U \times \mathbb{R}^n$, the positions and velocities of the particles, which must satisfy

$$\begin{aligned} x_i' &= v_i, \\ v_i' &= F_i(x)/M_i, \end{aligned} \tag{2.30}$$

for $i = 1, \ldots, n$. We will always assume that (2.30) has unique solutions; it suffices, for example, to assume that F is locally Lipschitzian in U.

In classical mechanics (the study of models governed by Newton's laws) a distinguished role is played by conservative systems. The system described above is *conservative* if there exists a C^1 function V defined on U, called the *potential energy* of the system, such that F is the negative gradient of V:

$$F_i(x) = -\frac{\partial V}{\partial x_i}(x).$$

Conservative systems are of interest because we may introduce a total energy for the system, and this energy is constant in time, i.e., conserved. The energy, a function of the position coordinates x and velocity coordinates v defined on $U \times \mathbb{R}^n$, is the sum of the potential energy $V(x)$ and a *kinetic energy*, denoted $T(v)$:

$$E(x, v) = T(v) + V(x) \equiv \sum_{i=1}^{n} \frac{M_i v_i^2}{2} + V(x).$$

Conservation of energy is then expressed in

Lemma 2.21: *Let $(x(t), v(t)) \in U \times \mathbb{R}^n$, $t \in I$, be any solution of the o.d.e. (2.30) for a conservative system. Then*

$$\frac{d}{dt} E(x(t), v(t)) = 0, \qquad t \in I.$$

Proof: This is just a calculation:

$$\begin{aligned} \frac{d}{dt} E(x(t), v(t)) &= \frac{d}{dt} \sum_{i=1}^{n} \frac{M_i v_i^2(t)}{2} + V(x(t)) \\ &= \sum_{i=1}^{n} M_i v_i v_i' + \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}(x(t)) x_i'(t) \\ &= \sum_{i=1}^{n} v_i(t) \big[M_i v_i'(t) - F_i(x(t)) \big] \\ &= 0. \quad \blacksquare \end{aligned}$$

Now suppose that we start our physical system with some initial position x^0 and velocity v^0 at time t_0. The initial energy will be $E_0 = E(x^0, v^0)$, and by Lemma 2.21 this will be the energy at any future time; that is, the particle moves on an *energy surface*

$$S(E_0) \equiv \{(x, v) \in U \times \mathbb{R}^n \mid E(x, v) = E_0\}.$$

As an immediate consequence of either Theorem 2.19 or Corollary 2.20, we have

Theorem 2.22: *If the energy surface $S(E_0)$ is compact, then the solution $(x(t), v(t))$ of (2.30) satisfying the initial condition $x(t_0) = x^0$, $v(t_0) = v^0$ is defined for all time.*

That is, *a priori* knowledge that the trajectory is confined to a compact subset guarantees that any solution may be extended to all time.

Example 2.1: (a) The harmonic oscillator potential $V(x) = \sum_{i=1}^{n} k_i x_i^2/2$, where $k_1, \ldots, k_n$ are strictly positive constants, is defined on all of $\mathbb{R}^n$ and yields a total energy function

$$E(x, v) = \frac{1}{2} \sum_{i=1}^{n} \bigl[m_i v_i^2 + k_i x_i^2\bigr].$$

Here the energy surfaces are ellipsoids and therefore compact.

(b) More generally, if a potential energy function $V \in C^2(U)$ goes to $+\infty$ at the boundary of U (more precisely, if for any $N > 0$ there is a compact set $K \subset U$ with $V(x) \geq N$ for $x \in U \setminus K$) then the energy surfaces will be compact and the solutions will exist for all time. We leave the verification as an exercise.